SCHOLARLY PUBLICATIONS

A CURRENT AWARENESS BULLETIN
OF RESEARCH OUTPUT

@DTU

(19th Edition)

JULY 2014

BY: CENTRAL LIBRARY

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

PREFACE

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

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

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

This issue contains the information published during July 2014. The arrangement of the contents is alphabetical. The full text of the article which is either subscribed by the university or available in the web is provided in this bulletin.

Central Library

CONTENTS

10. Implementation and Result Analysis of Polyalphabetic Approach to Caesar Cipher, ***Prachi Patni***, Computer Science & Engineering Department, DTU

11. Implementation of Modified Synchronous Reference Theory for Control of Grid Connected Distributed Generation Systems, ***Alka Singh,*** Electrical Engineering, DTU

12. Ligand based virtual screening for identifying potent inhibitors against viral neuraminidase: An in silico approach, Sangeeta Kashyap, ***Vinita Mishra***, ***Yasha Hasija*** Department of Biotechnology, DTU

13. Micro encapsulated cardanol derived benzoxazines for self-healing applications, Prasun Kumar Roy, Swapnil Shukla, Bimlesh Locha, ***Pratibha Sharma*** and ***Devendra Kumar***, Department of Applied Chemistry and Polymer Technology, DTU

14. Modeling &Simulation of Nano film thickness of Gold deposited by Thermal Evaporation Process (TEP), ***R. S. Mishra*** and ***Shailendra Kumar Gaur,*** ME, DTU

15. Near Field Communication Technology benefitted for Metro Rides, ***Manju Khari*** and Chetna Bajaj

16. A Probabilistic Analysis of Path Duration Using Routing Protocol in VANETs, Ram Shringar Rao, ***Sanjay Kumar Soni***, Nanhay Singh, and Omprakash Kaiwartya, DTU

17. Rubber toughening of unsaturated polyester with core–shell poly(siloxane)-epoxy microspheres, Prasun Kumar Roy, Chitra Rajagopal, ***Nahid Iqbal*** and ***Devendra Kumar***, Department of Applied Chemistry and Polymer Technology, DTU

18. Study of diffuse phase transition in Pb(Cd1/3Nb2/3)O3 compound. M. Pastor, Nawnit Kumar, Bineet Kumar, K. Biswas, A.C. Pandey and ***A. Panwar***, Applied Physics, DTU

***Faculty @Research Scholar $PG Scholar #Alumni**
£Undergraduate Student ~ Ex Faculty <Ex Research Scholar
! Teaching-cum-Research Fellow

International Journal of Engineering Inventions
e-ISSN: 2278-7461, p-ISSN: 2319-6491
Volume 3, Issue 12 (July 2014) PP: 01-07

A Study of Significant Software Metrics

Neha Saini[1], Sapna Kharwar[2], Anushree Agrawal[3]
[1]*Department of software Engineering, Delhi technological University, Delhi*
[2]*Department of software Engineering, Delhi technological University, Delhi*
[3]*Department of software Engineering, Delhi technological University, Delhi*

Abstract: *A software system continues to grow in size and complexity, it becomes increasing difficult to understand and manage. Software metrics are units of software measurement. As improvement in coding tools allow software developer to produce larger amount of software to meet ever expanding requirements. A method to measure software product, process and project must be used. In this article, we first introduce the software metrics including the definition of metrics and the history of this field. We aim at a comprehensive survey of the metrics available for measuring attributes related to software entities. Some classical metrics such as Lines of codes LOC, Halstead complexity metric (HCM), Function Point analysis and others are discussed and analyzed. Then we bring up the complexity metrics methods, such as McCabe complexity metrics and object oriented metrics(C&K method), with real world examples. The comparison and relationship of these metrics* are *also presented.*

Keywords: *Software metrics, software measurement scales, Function points, software attributes, categories of metrics, metrics for measuring internal attributes, software complexity.*

I. Introduction

In the field of science everything begins with quantification. All engineering disciplines have metric and some measurable quantity, so in the field of computer science also one needs some quantity to measure. Software measurement helps us in answering few questions like:-

- How good is the design
- How complex is the code
- How much efforts will be required

The above discussion makes it clears that software measure help in decision making in various life cycle phases. Software metrics can be defined as the continuous application of measurement based techniques to the software development process and its products, to supply meaning full and timely management information. Some definitions proposed for software metrics are [20]

Definition 1: Software Metrics provide a measurement for the software and the process of software production. It is giving quantitative values to the attributes involving in the product or the process.
Definition 2: Software metrics is to give the attributes some quantitative descriptions. These attributes are extracting from the software product, software development process and the related resources. They are product, process and resources.
Definition 3: Software measurement provides continuous measures for the software development process and its related products. It defines, collects and analyzes the data of measurable process, through which it facilitates the understanding, evaluating, controlling and improving the software product procedure.
Definition 4: According to IEEE "standard of software Quality Metrics Methodology", software metrics is a function, with input as the software data, and output is a value which could decide on how the given attribute affect the software.
Definition 5: According to J A McCall "The metrics are quantitative measure of the characteristics of the software which provide certain qualities. The hierarchical structure of the frame work provides relevance to management at one level and to the software developer at other level." [1, 20]

In 1979, Albrecht gave function points metrics, based on the requirement specification. Thereafter, in 1993 J-Y Chen and J-F Liu proposed a method, which used parameters like complexity, attributes, and reusability of class. to measure the Object- oriented software[18]. In 1994, Chidamber Kemerer based on inheritance tree gave a set of object-oriented metrics known as CK suite [6, 16,18]. In 1995, Brito also proposed, a set of metrics based on object-oriented attributes, called MOOD metrics [17]. In 2001, Victor and Daily based on software components gave a method called as SPECTRE, which is used to estimate time and modules level [20]. In 2003, Hastings and Sajeev proposed Vector Size Measure (VSM) metrics for early stage measurements in the software lifecycle. Subsequently, Arlene F. proposed a forecast metric based on the objects and attributes, it was used estimate work load and production power [19, 20].

1.1.Categories of metrics

Metrics can be classified into three categories [1]

- Product Metrics
- Process Metrics
- Project Metrics

1.1.1 Product Metric: Software Product metrics are measures of software products such as source code and design documents. Software product metrics that measure software product in different paradigms are also different. In procedural paradigm metr ics measure functions and how functions interact, In Object oriented paradigm classes and how classes interact is measured. [5]. Product Metrics can be further classified as shown in Fig 1[1, 5].

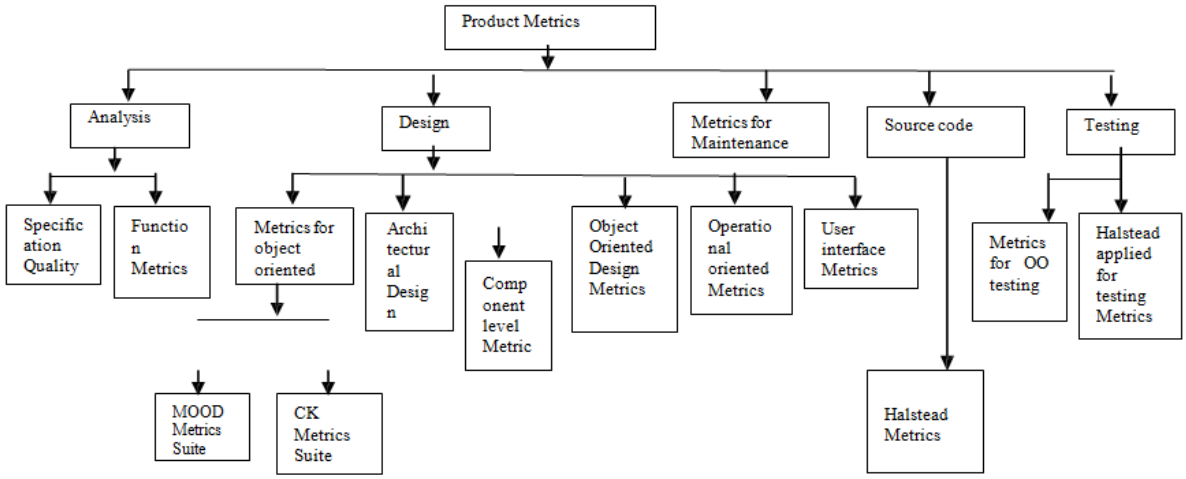


Fig 1. Classification of Product Metrics [2]

1.1.2 Process: Process metrics emphasize on the process of the software development. It mainly focus on how long a process last, what about the cost, whether the methods used are effective etc.. It includes the improvement of the process and the prediction for the future process. The main part of this metrics includes maturity, management, life cycle, product ratio, defect ratio, etc.. It is beneficial to the control and management of the whole development procedure.

1.1.3 Project: Project metrics are designed to control the project situation and status. The metrics includes scale, cost, workload, status, production power, risk, the degree of satisfaction from clients, etc.. Project metrics is used to analyze the project to avoid the risk factors and help to optimize the development plans. Afterward, project metrics improves the quality of the product via advancement in techniques, methods and management strategies.

II. Software Measurement

In order to measure, one needs to identify an entity and a specific attribute of it. It is very important to define in a clear way what one is measuring because otherwise one may not be able to perform the measure or the measures obtained can have differen t meaning to different people. Measurement can take place in all the different phases of software development: Requirement Analysis, Specification, Design, Coding and Verification. Undoubtedly, measurement is most useful if carried out in the early phases. Different internal and external attributes are shown in Table 1.

Table I. Entities and attributes [3]

Entity	**Internal Attributes**	**External Attributes**
A. **Product**		
Requirements	Size, Reuse, Modularity, Redundancy, Functionality	Understandability, Stability
Specification	Size, Reuse, Modularity, Redundancy, Functionality	Understandability, Maintainability
Design	Size, Reuse, Modularity, Coupling, Cohesion	Comprehensibility, Maintainability, Quality

Code	Size, Reuse, Modularity, Coupling, Cohesion, Control Flow Complexity	Reliability, Usability, Reusability, Maintainability
Test set	Size, Coverage level	Quality
B. **Process**		
Requirements Analysis	Time, Effort	Cost effectiveness
Specification	Time, Effort, Number of requirements Changes	Cost effectiveness
Design	Time, Effort, Number of specification Changes	Cost effectiveness
Coding	Time, Effort, Number of design Changes	Cost effectiveness
Testing	Time, Effort, Number of code changes	Cost effectiveness
C. **Resource**		
Personnel	Age, Cost	Productivity, Experience
Team	Size, Communication Level, Structure	Productivity
Software	Size, Price	Usability, Reliability
Hardware	Price, Speed, Memory size	Usability, Reliability

Many metrics have been proposed for structural complexity and they measure a number of internal attributes of software. Structural metrics can be divided in intra module metrics and inter module metrics. Module metrics are focused at the individual module level (subprogram or class) and comprehend: size metrics, control flow complexity metrics, data structure metrics and cohesion metrics. Inter module metrics measure the interconnections between modules of the system and are constituted by coupling metrics. Various measuring scales with example are written in table 2.

Table- II. Measurement scales [2]

Measurement scales	**Definition**	**Example**
Nominal	Items are assigned to group or categories, it is diffrentiative, no quantitative information is generated, no ordering is implied.	Recursive or non-recursive program Types of errors Binary executable program, or DLL
Ordinal	Measurements are ordered, higher no represent higher vales but the no are only for ordering	CMM maturing levels How often software fails
Interval	It separate classes based on vale you know exactly when the item crosses one class and goes to another. Classes are ordered. Additional and subtraction can be performed but	Logs of event on dates.
Ratio	Has ordering, interval sizes and ration are possible. Value	LOC (Length of code) as statement count.
Absolute	Counting entities in the entity set. All arithmetic operation are meaning full. There is only one	Lines of code, No of failures, no of project engineers.

III. Standard Metrics

Metrics chosen among the most widely used and are grouped according to the phase of software development in which they can be applied. The overview of various metrics is presented below.

a. Lines of code

The simplest software metric is the number of lines of code. But it is not clear whether one have to count the comments as well, even if they give a big contribution to the understand ability of the program, and the declarative parts, that give a contribution to the quality of the program [2,5]. Also some languages allow more than one instruction on the same line. If one counts the number of instructions, then we are still uncertain about comments and declarations. "A line of code is any line of program text that is not a comment or blank line, regardless of the number of statements or fragments of statements on the line. This specifically includes all lines containing program headers, declarations, and executable and non-executable statements." One of the main disadvantages with LOC is that it does not take into account the goodness of the

code: if one uses LOC to measure productivity, a short well designed program is "punished" by such a metric. Another disadvantage is that it does not allow comparing programs written in different languages.

In spite of these problems, lines of code are a widely used metric due to its simplicity, ease of application, inertia of tradition and absence of alternative size measures. Moreover, there are many empirical studies that demonstrate the usefulness of LOC:

LOC have been used for a variety of tasks in software development: planning, monitoring the progress of projects, predicting. From the point of view of measurement theory, LOC are a valid metric for the length attribute of a program because the empiri cal relation "is shorter than" is perfectly represented by the relation between lines of code [5]:

x "is shorter than" y -- LOC(x) < LOC(y)

b. Halstead's Software Science

M. Halstead is based on the assumption that a program is made only of operators and operands and that the knowledge of the numbers of distinct and repeated operators and operands is sufficient for determining a number of attributes of software such as program length, volume, level, programming effort. It is interesting to note that, even if the equation provides only estimates, they are exact, not statistical. Here let us assume

Operand: every variable or constant present in the program

Operator: every symbol or combination of symbols that influences the value or order of an operand.

Punctuation marks, arithmetic symbols (such as +,-,* and /), keywords (such as if, while, do, etc.), special symbols (such as:=, braces, parenthesis, ==, !=) and function names are operators. Some attributes are considered fundamental and used to derive all the other attributes of the model:

n1 the number of distinct operators n2 the number of distinct operands N1 is total number of operators

N2 is total number of operands

The length N of a program is the total number of symbols in the program and is given by

N=N1+N2 (1) Expected software length:

H= n1 log2(n1) + n2 log2 (n2)

Volume: V= N*log2(n)

Level: L=V*/V= (2/n1)*(n2/N2) Difficulty: D=V/V*= (n1/2)*(N2/n2)

Effort: E= V*D (2)

c. High-Level Design metrics

High-level designs is either traditional or object oriented. The goal of these metrics is to assess the quality of a software design with respect to its error proneness. The high level design of a system is seen as a collection of modules. A module is a provider of a computational service and is a collection of features, i.e. constants, type, variable and subroutine definition s. It is an object in object oriented systems, but can be present as well in traditional systems. Modules are composed of two 18 parts: an interface and a body (which may be empty). The interface contains the computational resources that the module makes visibl e for use to other modules. The body contains the implementation details that are not to be exported. The high-level design of a system consists only in the definitions of the interfaces of modules. A software parts is a collections of modules. The definitions of interactions.

Data declaration-Data declaration (DD) Interaction A data declaration A DD-interacts with another data declaration B if a change in A's declaration or use may cause the need for a change in B's declaration or use. Data declaration-Subroutine (DS) Interaction A data declaration DS-interacts with a subroutine if it DD-interacts with at least one of its data declarations.

Let us consider now the attribute cohesion. It is the extent to which features that are conceptually related belong to the same module. It is desirable to have a high cohesion because otherwise we would have features that depend on each other scattered all over the system, with the result that the software could be more error prone. The set of cohesive interactions in a module m is the union of the sets of DS interactions and DD-interactions, with the exception of those DD-interactions between a data declaration and a subroutine formal parameter.

d. McCabe's Cyclomatic Complexity

It concentrates on control flow complexity and does not take into account, the contribute to complexity that derives from data [9]. A program control flow can be represented by a graph which has a unique entry node and exit node, and in which all nodes are reachable from the entry and the exit is reachable from all nodes. Idea is to measure the complexity by considering the number of paths in the control graph of the program. But even for simple programs, if they contain at least one cycle, the number of paths is infinite. Therefore he considers only the number of independent paths: these are complete paths, (paths that go from the starting node to the end node of the graph), such that their linear combinations can produce all the set of complete path of a program.

e. Function Points

Function Points give a measure of the functionality of the system starting from a description, in natural language, of user's requirements. Thus they provide a technology independent estimate of the size of the final program and are probably the only measure of size that is not related to code. The measurement of function points is based on identifying and counting the functions that the system has to perform. Function Points are computed in terms of

- No of external inputs
- No of external outputs
- No of external inquiries
- No of external files
- No of internal files

Each function identified in the system is then classified to three levels of complexity: simple, average and complex. According to the complexity and the function type, a weight is assigned to each function and all the weights are summed up to give the unadjusted function point count. The final function point count is then obtained by multiplying the unadjusted count by an adjustment that expresses the influence of 14 general system characteristics.

3.1 Object oriented metrics

3.1.1. MOOD metrics suite

f. Method Hiding Factor (MHF) and Attribute Hiding Factor (AHF)

These metrics are basically measures of encapsulation [6, 16, 17]. In the encapsulation, MHF and AHF act as measures of "the use of the information hiding concept". MHF is defined formally [17, 18] as:

$$\frac{\sum_{i=1}^{TC_1} \sum_{m=} M(Ci) (1-V(M}{\sum_{i=1}^{TC} (M_d(C_i)_{mi}))} \quad (3)$$

Where M_d (C_i) is no of methods in class, TC is total No of classes and,

$$V(M_{mi}) = \frac{\sum_{i=1}^{TC} (\text{is visible } (M_{mi}, C_j))}{TC-1} \quad (4)$$

For all classes, C_1, C_2, ..., C_n, a method counts is 0 if it can be used by another class, and 1 if it cannot. The total for the system is divided by the total number of methods, to give the percentage of hidden methods in the system. AHF was defined in similar way, but using attributes rather than methods. The definitions of MHF and AHF cause discontinuities for systems with only one class. Data encapsulation, Information hiding is very primitive in today's open world. Both MHF and AHF use code visibility to measure information hiding, and thus are validated using all validation criteria. MHF and AHF measure the relative *amount* of information hiding and not the *quality* of the information hiding design decisions.

g. Method Inheritance Factor (MIF) and Attribute Inheritance Factor (AIF) Metrics

MIF and AIF measure the number of inherited methods and attributes respectively as a proportion of the total number of methods/attributes [6, 16]. There is a relationship between the relative amount of inheritance in a system and the number of methods/attributes which have been inherited. It can be defined as:-

$$\frac{\sum_{i=1}^{TC} M_i(Ci)}{\sum_{i=1}^{TC} M_a(C_i)} \quad (5)$$

here, $M_a(C_i) = M_d(C_i) + M_i(C_i)$

$Md(C_i)$ = the number of methods declared in a class,

$Ma(C_i)$ = the number of methods that can be invoked in association with C_i,

$Mi(C_i)$ = the number of methods inherited in C_i.

The total for the system is divided by the total number of methods, including any which have been inherited.

h. Coupling Factor (CF)

This metric measures the coupling between classes, excluding coupling due to inheritance. CF has been defined [16, 17] as:

$$\frac{\sum_{i=1}^{TC} \sum_{j=1}^{TC} (\text{is client } (C_i, C_j))}{TC^2-TC} \quad (6)$$

Here, is client (C_i, C_j)=1 iff C_c => C_s ^ C_c =/C_s otherwise equal to 0

Cc =>Cs represents the relationship between a client class, *Cc*, and a supplier class, *Cs.*
CF is calculated by considering all possible pair wise sets of classes, and asking whether the classes in the pair are related, either by message passing or by semantic association links (reference by one class to an attribute or method of another class). These relationships are considered to be equivalent as far as coupling is concerned [7], [9]. Thus, CF is a direct measure of the size of a relationship between two classes, for all pair wise relationships between classes in a system. There are two possible approaches to validating CF. Firstly, CF can be considered as *direct* measure of interclass coupling. Alternatively, one can consider CF to be an *indirect* measure of the attributes to which it was said to be related [8]. CF will enable us to distinguish between different programs with different levels of coupling. A system with a high level of interclass coupling will have a high CF value.

i. Polymorphism Factor (PF)
It is a measure of polymorphism potential. It can be defined as [6, 16]:

$$\frac{\sum_{i=1}^{TC} M_o(C_i))}{\sum_{i=1}^{TC} [(M_n(C_i) * DC(C_i)]} \quad (7)$$

Here $M_d(C_i) = M_n(C_i) + M_o(Ci)$ and
Mn(*Ci*) = number of new methods,
Mo(*Ci*) = number of overriding methods,
DC(*Ci*) = descendants count

PF is the number of methods that redefine inherited methods, divided by the maximum number of possible distinct polymorphic situations PF is an indirect measure of relative amount of dynamic binding in a system.

3.1.2 CK metric suite*:*
j. Weighted Method per class (WMC)
It can be defined as consider a Class S_1, with methods $M_1 \ldots\ldots M_n$ defined in class. Let $C_1 \ldots. C_n$ be the complexity of the methods then[6, 16].

$$WMC = \sum_{i=1}^{n} C_i \quad (8)$$

WMC relates the complexity of the things as methods are properties of object class and complexity is determined by the cardinality of it sets of properties. The number of methods is a measure of class definition as well as being attributes of a class, as attribute are nothing but properties. If all the method complexity are considered as unity then the WMC=n, the no of methods.

k. Depth of inheritance tree (DIT)
It measures the depth of inheritance of the class. If there are multiple inheritances then DIT will be maximum length from node to root of tree [6, 16]. The deeper a class in the hierarchy, greater the number of methods, it is likely to inherit, making it more complex to predict. Deeper the tree greater the design complexity and so does greater potential to reuse of inherited methods. As, it can be very well predicted from the tree diagram in figure 2.

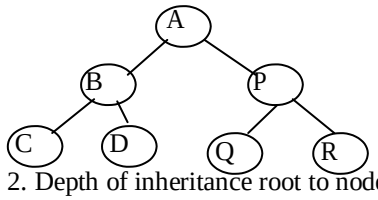


Fig 2. Depth of inheritance root to node

l. Number of Children (NOC)
It is the number of immediate subclasses subordinated to a class in the hierarchy. It measure how many subclasses are going to inherit the methods of parent class [2]. Greater the number of children, greater the reuse and likelihood of improper abstraction of parent class. It will require more testing methods in that class.

m. Coupling between the object classes (CBO)
CBO for a class is the count of number of other classes to which it is coupled. It relates to the notion that as object s cou pled to another if one of them acts on the other. Two classes are coupled when method declared in one class use methods or instance variable defined by other class. The more independent a class is

the easier to reuse it in another application. In order to improve modularity and promote encapsulation, inter object class couples should be kept to a minimum.

n. Response for a class (RFC)

The response set for the class can be expressed as: $RS=\{M\} U_{all\ i} \{R_i\}$

Where R_i is set of methods by method i and {M} is set of all methods in the class.

The response set of a class is a set of methods that can potentially be executed in response to a message received by an object of that class. The cardinality of this ser is a measure of the attributes of objects in the class. It is also a measure of the potential communication between the class and other classes.

o. Lack of cohesion methods (LCOM*)*

LCOM is a count of the number of methods pairs whose similarity is 0 minus the count of methods pairs whose similarity is not zero. The larger the number of similar methods, the more cohesive the class. This uses the notion of degree of similar methods. The degree of similarity for two methods M_1 and M_2 in a class is given by $\sigma() = \{I_1\} \cap \{I_2\}$ where $\{I_1\}$ and $\{I_2\}$ are set of instance variables used by methods M_1 and M_2.

IV. Conclusion

In this paper we have discussed the basic questions about software metrics: why measuring, what to measure, how to measure and when to measure. With the rapid advance of software, their metrics have also developed quickly. Software metrics become the foundation of the software management and essential to the success of software development.We have analyzed the attributes of software, distinguishing attributes of process, product and project. Among all the attributes, complexity is probably the most important one and it comprehends many different aspects of software. The complexity of software will directly affect the eligibility, reliability of the software. We have described various metrics, and most of the metrics defined in the literature, have not been validated using the theory of measurement. Some of them have been validated by showing that they are correlated with other metrics: Function Points have a good correlation with size and effort. With more and more people working hard in this field one can expect to see more thorough and matured software metrics in the near future.

REFERENCES

[1] K.K. Aggarwal and Yogesh Singh, Software Engineering, third edition, 2009 reprinted, new age international publisher.

[2] N.E. Fenton and Shari Lawrencde Pfleeger, Software Metrics: Software metrics A Rigorous and practical approach , second edition. Thomson publication.

[3] S. Morasca, Software Measurement: State of the Art and Related Issues slides from the School of the Italian Group of Informatics Engineering, Rovereto, Italy, September 1995.

[4] Weyuker, Evaluating Software Complexity Measures, IEEE Trans. Software Eng., 14(9), 1988, pp. 1357-1365.

[5] N. Fenton, Software Measurement: a Necessary Scientific Basis, IEEE Trans. Software Eng., 20, 1994, pp. 199-206.

[6] S. Chidamber, C. Kemerer, A Metrics Suite for Object Oriented Design, IEEE Trans. Software Eng., 20/6), 1994, pp. 263-265

[7] Albrecht and J. Gaffney: Software Function, Source Lines of Code, and Development Effort Prediction: A Software Science Validation; in IEEE Trans Software Eng., 9(6), 1983, pp. 639-648

[8] Kumar, A., Kumar, R. and Grover, P. S, "Towards a unified frame work for cohesion measurement in aspect oriented system", Australian conference on software engineering 2008, IEEE, Washington march 2009 pp 57-65.

[9] Thomas J. McCabe, "A Complexity Measure", IEEE transactions on software engineering, IEEE, Washington, Oct 1976, pp 308 320.

[10] N. E. Fenton, "Software Metrics: Successes, Failures & New Directions, " presented at ASM 99: Applications of Software Measurement, S a n J o s e , C A , 1999 .

[11] S. S. Stevens, "On the Theory of Scales of Measurement," Science, vol.103, pp. 677-680. IEEE, "IEEE Std. 1061-1998,

[12] "Standard for a Software Quality Metrics Methodology, revision." Piscataway, NJ,: IEEE Standards Dept., 1998.

[13] L. Briand, S. Morasca, V. Basili, Property-Based Software Engineering Measurement, IEEE Trans. Software Eng. 22(1), 1996, pp. 68-85.

[14] B. Henderson-Sellers, Object Oriented Metrics: Measures of Complexity, Prentice Hall, Upper Saddle River, NJ, 1996

[15] M. Evangelist, Software Complexity Metric Sensitivity to Program Structuring Rules, Journal of Systems and Software, 3(3), 1983, pp. 231-243.

[16] "An Evaluation of the MOOD Set of Object-Oriented Software Metrics", Rachel Harrison, Steve J. Counsell, IEEE transactions on software engineering, vol. 24, no. 6, June 1998.

[17] F. Brito Abreu and W. Melo, "Evaluating the Impact of OO Design on Software Quality," *Proc. Third International Software Metrics Symp.*, Berlin,1996.

[18] "Empirical Analysis of CK Metrics for Object-Oriented Design Complexity: Implications for Software Defects" Ramanath Subramanyam and M.S.Krishnan IEEE transactions on software engineering, vol. 29, no. 4, April 2003

[19] "Applicability of Three Complexity Metrics" D. I. De Silva, N. Kodagoda , H. Perera' The International Conference on Advances in ICT for Emerging Regions - ICTer 2012: 082-088

[20] "The Research on Software Metrics and Software Complexity Metrics" Tu Honglei, Sun Wei, Zhang Yanan, International Forum on Computer Science-Technology and Applications.

Antecedents of Discretionary Risky-Service Behaviour: An Exploratory Study

INTRODUCTION

Given a progressively more competitive environment, it is increasingly claimed that the key factor in the success and indeed survival of service organizations is the effective management of employee-customer relationships (Gro¨nroos, 1994). Empirical results that underline the importance of employee - customer relationships are consistently advanced in the form of studies that highlight the causal linkages between service quality, employee satisfaction, customer satisfaction, customer loyalty, and profitability of the firm (Heskett et al., 1997; Reichheld and Sasser, 1990). The service marketing literature has recognized the importance of developing and maintaining enduring relationships with customers of service businesses (Henning-Thurau et al., 2002). In banks the employees challenge (due to financial crisis and new regulations) is to maintain and enhance the relationships with customer by facilitating and supporting the financial activities of customers and finding ways to be involved in the life (and business) of their customers. In this regard, customer-facing attitudes and behaviours of customer contact employees are forwarded as the most salient factor (Reynolds and Beatty, 1999). Initially quality services (Gurau, 2003) and later on relational strategies (Molina et al., 2007; Dimitriadis, 2010) were considered the determinants for maintaining and enhancing relationships with the customers in banks. The Indian banking sector is no exception to this and the researchers agree that adoption of relational strategies is critical for the survival of Indian banks (Roy and Shekhar, 2010). It was also argued that Asians place more importance on "strong relationships in business" than the West (So and

Speece, 2000). Moreover, as cultural differences influence relational perception and behavior, (Jham and Khan, 2008) researchers aim exploration of unique relational behaviour in an Eastern cultural context such as India.

Researchers argue that service employee' attitudes about their within-role, anti-role and extra-role behaviour determines the nature of their performance in services (Wallace, Chernatony, Buil, 2011). There is significant evidence of service employees anti-role behaviour towards customers like service sabotage ((Blancero & Johnson, 1997; Harris and Ogbonna, 2002) and there is widespread evidence of employees extra-role behaviour (positive) towards organizational outcomes (Organ, 1988) but employee extra-role behaviour towards customer (Discretionary service behaviour) has limited representation in literature (Blancero & Johnson, 1997; Chaoluck et al., 2013). Courage to exercise empowerment by bending the rules and violating company policies to maintain customer satisfaction is forwarded as an important quality in customer oriented employees (Dubinsky, 1994). This combination of positive employee discretionary service behaviour and quality of exercising empowerment by bending the rules and violating company policies to maintain customer satisfaction leads to discretionary risky-service behavior (DRSB) and extends existing service employee discretionary behaviour theory in a new direction forming the foundation of our model. The current research draws upon literature in marketing, organizational citizenship behavior, extra-role behavior, and psychology to offer insights into two broad questions about the DRSB phenomenon: (1) What are the antecedents to DRSB? (2) Are there certain identifiable employee traits or individual factor that explains this kind of service initiation?

The purpose of this article is to present a description of DRS behaviour through exploration. We begin with the broad influences and dimensions of contact employee behavior. We then narrow our focus to pro-organization; pro-customer employee behaviour to develop a specific model of DRS behaviour and a series of propositions regarding its

antecedents. The proposed framework emerges from the data and then to deepen the framework support from the theory and propositions from empirical data were added. Specifically, we base our propositions on the factors affecting contact employees in relational service environments where maintaining customer relationships are the norm. Relational service jobs with long-term customer encounters over time such as relationship banking, and hospitality etc. typically are more amenable to exercising empowerment by service employees. Additionally the emphasis on employee behavior in relational service environments is based on customer influences rather than management influences. By restricting our discussion to relational service encounters and that too in a collectivist culture (India) where personal stance is preferred over professional stance as compared to individualistic cultures that support professional stance in employee-customer relationships we set boundary conditions for our model. Our model of DRS behaviour is not intended to be broadly generalizable to all service environments e.g. a brief exploration of DRS behaviour in foreign banks in India revealed that employees in foreign banks are trained to take a 'professional stance' rather than a 'personal stance' in all relational interactions with the customers and employee behaviour is managed through stringent audit processes and therefore, they lack the means to exhibit DRS behaviour. Drawing from extant literature and our own research, we propose antecedents of DRS behaviour supported by a series of propositions. We conclude with suggestions for future research and implications that follow form DRSB framework.

LITERATURE REVIEW

Researchers' and practicing managers' interest in study of employee behaviours is not new. Employee being an active resource - their attitudes, behaviours, and qualities contribute significantly towards the organizational outcomes. Prior researchers focussed only on task related performances for evaluation of employee contribution to the organization. Organ's

(1988), study argued for other virtues existent in employees for superior organizational outcome rather than focus on individual productivity only. The early studies capturing employee behaviour focused on Sportsmanship (tolerating less than ideal circumstances without complaining), Civic virtue (performing tasks not required to be performed by employees for the benefit of the organization), Conscientiousness (working beyond requirements - for example, working after hours), Altruism (helping co-workers in their tasks), and Courtesy (treating fellow employees with respect) (Organ, 1988). The above mentioned behaviours were collectively termed as organizational citizenship behaviour (OCB) and have a wide spread presence in HRM, OB and Psychology literature (Wayne and Green, 1993; Haworth and Levy, 2001; Podsakoff and Mackenzie, 1997; Smith et al., 1983) as it got acceptance from researchers in various streams.

OCB periphery being employee-supervisor interaction and relationship it was argued that the employees actions performed may be tactic aimed at impressing managers responsible for their promotions and rewards (BoliNo, 1999). Additional studies contributed to the advancement and conceptualization of the construct and proposed extra role behaviour (ERB) (Van et al. 1995), and discretionary service behaviour (DSB) (Blancero & Johnson, 1997; Kelley, 1993) constructs. 'Discretion' was found to be a common component in all these constructs. Discretionary behaviour is defined as behaviour which is outside the explicit job description and employee role expectations (Reychav and Sharkie, 2010; Mosavi et al., 2013). Another important component along with 'discretion' was 'intent' and literature reported the existence of both positive and negative intent of employees while exhibiting discretionary behaviour. Literature focussed on negative behaviour surfaced the existence and understanding of various employee behaviours like sweethearting (Tarnowski 2008; Brady et al., 2011), feigning commercial friendships (Rosenbaum, 2009), sabotage (Wallace, Chernatony, and Buil 2011; Analoui 1995), etc. termed as anti-role behaviour. Where on one

hand employee positive behaviour positively contributes to the organisation as a whole (Ferrin et al., 2007) rather than supporting individuals vested interests, employees' anti-role behaviour is unconventional behaviour employed at work with the most subtle form being the misuse of facilities for one's own end, and the most notorious being the destructive action referred to as sabotage (Analoui, 1995). The above mentioned literature developed our understanding that customer contact employee enactment is multifaceted and comprised of customer orientation, extra-role behaviour, and anti-role behaviour (Analoui, 1995; Harris and Ogbonna, 2002), or sabotage (Wallace, Chernatony, and Buil 2011).

Employee behaviour positive as well as negative significantly relates to employee satisfaction. Literature aimed at understanding the 'why' of employee negative behaviour proposed dissatisfaction with job, working conditions, and relationship with boss as significant determinants (Harris and Ogbonna, 2002. Research studies conducted in service management literature extends the employee behaviour studies to service organizations and identifies two sets of relationships a) employee-organization (internal) , b) employee-customer (external) ; and two demeanour a) positive , b) negative. Based on these relationship directions and intent four categories of intentional (accidental behaviours excluded) employee behaviour were proposed (Blancero & Johnson, 1997):

1. Pro-organization Pro-customer: This is an ideal and most sought after employee behaviour. Such employee behaviours are directed at organizations benefit as well as the customers benefit. Employees' work satisfaction and employees' customer orientation are significant drivers of such behaviours.
2. Anti-organization Pro-customer: Service employees not satisfied with job or their relationship with their managers tend to 'give away' to the customers. This leads to inflated customer satisfaction scores but causes damages to the organization in terms of revenue losses.

3. Pro-organization Anti-customer: In this kind of employee behaviour employees are not satisfied with their jobs (though the dissatisfaction may be borderline) and in guise of adhering to procedures and processes ignore the needs of the customers. This behaviour may also be exhibited by employees who are process oriented and not people oriented. It may also be employee's response to a problem customer. With customer gaining centre stage in all service organizations and customer satisfaction scores being used to evaluate employee's performance anti-customer behaviour may be subtle and difficult to identify.
4. Anti-organization Anti-customer: Strongly negative behaviour where employee is completely dissatisfied with the organization and takes revenge from customer in forms of slowing down the work, not giving complete information, rude behaviour etc. This form of behaviour may be exhibited by employees due to unfair or inequitable demands by either customers or managers.

Our focus was Pro-organization Pro-customer employee behaviour but we considered all above mentioned behaviours to understand 'what drives' and 'what withholds' the employee behaviour of our study. The following construct were found to be of interest due to their relationship with employee discretionary behaviour:

Trust: Trust based relationships in an important determinant in success of service organizations. Although the literature weighs towards trust in employee-organization relationship (Dirks and Ferrin, 2001; Ferrin et al., 2007; Kerkhof at al., 2003; McEvily et al., 2003; Reychav and Sharkie, 2010), trust as a key dimension for evaluating employee-customer relationships (De Wulf et al., 2001) is also identified.

Management Support: Psychological support (Dirks and Ferrin, 2001) and organizational supportiveness (Rhoades et al., 2001) are forwarded as strong antecedents of employee perceptions, which in turn were positively correlated with employee extra-role behavior

(Wang, 2010). Employee perception of psychological support they receive influences the amount of reciprocation by the employees back to the organization (Dirks and Ferrin, 2001; Rhoades et al., 2001). It applies to both employee-organization relationship and employee-customer relationship.

Job Satisfaction: Many research studies have analysed the link between Service quality, employee satisfaction, and customer satisfaction (Heskett et al., 1997; Reichheld and Sasser, 1990). Satisfaction is identified as a key dimension for evaluating employee-customer relationship (De Wulf et al., 2001) and both internal customer (employee) satisfaction and external customer satisfaction are important. It has been suggested that the employee satisfaction is the most crucial determinant of employee discretionary behaviour (Kelley, 1993).

Rewards: Early OCB research highlighted discretionary behaviour as unrewarded behaviour (Wayne and Green 1993) but later studies explored employee discretionary behavior as a tool to impress management in the expectation of rewards (BoliNo, 1999). Haworth and Levy (2001) advanced the concept and the researchers claimed that employees will enact and sustain discretionary behaviour only when they believe that their managers will fairly reward such behaviours. Not only rewards (Reychav and Sharkie, 2010) but reward indicators like promotional opportunity, employment continuity, and pay (Rhoades et al., 2001) are also positively correlated with employee extra-role behavior as they express the organization's dependability and supportiveness (Rhoades et al., 2001).

Relationship: Service employees establish and nurture relationships with customers through communication, customization, preferential treatment, and personal relationships (Claycomb and Martin 2002). Employee-customer relationships that relates significantly to relational outcomes are social bonding (Bendapudi and Berry 1997), commercial friendship (Price and Arnould 1999; Gremler and Gwinner, 2008), and rapport (Gremler and Gwinner , 2008).

Relationship oriented organizations proactively involve in creating, developing and maintaining engaged, committed, interactive, and profitable exchanges with high value customers over time (Camarero, 2007). Literature suggests that customer facing discretions (like discretionary service behaviour) are influenced by employee-customer relationship (Camarero, 2007) and organization facing discretions (like OCB) are influenced by the quality of the relationship between employee and employee's immediate supervisor (Wong et al., 2003).

Customer Orientation: Customer-oriented service employees' view customer relationships from a long-term prospective and concentrate on what is in the best interest of the customer and that augments service relationship (Gremler and Gwinner 2008). Customer orientation is proposed as a vital trait in customer-employee relationship models (Liliana and Lester, 2000). Research stream that advocates employees' customer orientation suggest expertise, and attitude as requisite qualities which customer contact employee must possess (Dubinsky, 1994; Henning-Thurau, 2004). Employees' expertise refers to technical knowledge and skills that a service employee must possess in order to fulfil the customer's needs during the interaction process (Henning-Thurau, 2004). It is suggested that for organizational success customers' needs must be addressed proficiently, expeditiously, and acceptably (Dubinsky, 1994). Attitude refers to genuine empathy, concern, respect, courtesy, sociability and friendliness for the customers (Camarero, 2007; Dubinsky, 1994). Service employees possessing this quality are able to develop social bond, commercial friendship, personal friendship, and rapport with their customers (Claycomb and Martin, 2002; Price and Eric, 1999).

Empowerment: Service quality in organizations is associated with employee empowerment. Dubinsky (1994), proposed empowerment as an important quality possessed by customer oriented service employee. It was defined as audacity to exercise empowerment with

responsibility and authority and go beyond the defined roles (Bowen and Lawler, 1992; Dubinsky, 1994). Breaking the rules or violating company policy to maintain customer satisfaction (Dubinsky, 1994) comprises empowerment and therefore is known to have both positive and negative consequences (Harris and Ogbonna, 2002). The said behaviour is also termed as preferential treatment (Claycomb and Martin, 2002; De Wulf et al., 2001) by some researchers.

Risk-propensity: Risk taking is proposed as an antecedent attitude influencing negative employee behaviour. The research proposed a significant correlation between greater tendency for risk taking and anti-role behaviour (Harris and Ogbonna, 2002). Risk taking is a dimension required for exercising empowerment, exhibiting entrepreneurial abilities and therefore has a positive element. In service management literature risk taking is suggested as a vital trait in customer-employee relationship models (Liliana and Lester, 2000).

Career Orientation: (Harris and Ogbonna, 2002) highlighted relationship between career orientation and service sabotage behaviour. Employees found to express a wish to pursue a career in their current organization appeared more committed to customer service and abstained from negative behaviours. Opportunities of career advancement and fairness perception influences discretionary behaviour ((Blancero & Johnson, 1997).

Preliminary understanding and knowledge of the above mentioned literature facilitated in theme identification from our empirical data. On identifying a theme we further reviewed the literature to find support. This back and forth examination to match support from literature and exploration from our own data translated into the suggested propositions and the exploratory framework.

METHODOLOGY

A two-stage explorative study was conducted in the Indian banking context. Stage 1, was a qualitative research to explore - What discretions do relationship managers in banks

exercise to strengthen relationship with their customers?, and their intent for engaging into such behaviours. The source of qualitative data included personal interviews. Qualitative research design with purposive sampling was adopted as adjudged that relationship managers can provide the best required information.

Responses were obtained from service employees (relationship managers) of two public and three private leading banks in India. Due to the sensitive nature of the issue under investigation, personal references and a snowball sampling design was adopted. As the use of confederates may provide more valid data than if researchers attempted to directly access respondents involved in discretionary behaviour. Data collection procedure was based on previous research designs for the study of deviant behaviours (Liao, Joshi, and Chuang, 2004).

35 relationship managers (confederates) were contacted and 25 mentioned tasks related to what DRS behaviour contains. Grounded theory approach was used and coding categories were derived directly and inductively from the raw data (Glaser & Strauss, 1967 *cf* Suter, 2012) and the process of constant comparison was used. The key components of the behaviour were operationalized based on the behaviours mentioned and support from theory (Blancero & Johnson, 1997; Dubinsky, 1994). The operationalization were

a) discretionary behaviour

b) positive intent (pro-organization and pro-customer)

b) customer service behaviour

c) risk-taking behaviour (breaking the rules or violating company policy to maintain customer satisfaction)

Behaviours not encompassing all four key components did not qualify and were dropped e.g. certain behaviours like reversing the charges levied incorporated the first three components namely discretionary behaviour, positive intent, and customer service behaviour

but did not include the risk-taking behaviour. Analysing the discretionary behaviour containing all four components the definition of DRSB was developed.

Stage 2 was conducted two months away from stage 1(after the DRSB definition was developed) with an objective to identify DRSB antecedents. We began contacting respondents from the list of the 25 respondents who contributed to the DRSB definition for personal interview without a predetermined number in mind. An interview conducted with 21 respondents out of the 25 as after the 19st interview the data saturation point reached as we were not getting any new information. All the respondents were currently employed and had an experience of 2 or more years as relationship managers.

The interview required respondents to answer three questions. The first question provided respondents with a clear definition of DRSB and asked them if they had ever exhibited DRSB. Then, respondents were asked to describe their discretionary risky-service incident in which they had participated during their tenure with the bank. Thereafter, they were asked to list and elaborate on the factors that influenced their decision to either engage in or abstain from DRSB.

Simultaneous data analysis and data collection were done using analytic induction. Potential categories and themes began to surface as early as from first participant's data. The subsequent data either compared with the previous categories and themes or presented new categories. We began with coding the responses. Thereafter, codes were combined to form a single concept. A critical procedure of back-and-forth comparison within categories, and between categories conducted. Categories were related to each other in some analytical way. The next level abstraction was performed by interpreting all the categories and themes developing a representation of an exploratory framework (Figure 1) (Suter, 2012).

To extend support to the exploratory framework we reviewed prior research on organizational citizenship behavior (OCB), discretionary extra-role behavior (DERB), and discretionary

service behaviour (DSB) to determine whether the emergent factors were shared by those studies. Although the organizational citizenship literature and discretionary extra-role behavior literature does not address discretionary risky-service per se, there may be common factors that influence employees discretionary behavior.

Reliability and Validity: The researchers used inductive reasoning, whereby through careful examination and constant comparison themes and categories emerged from the data (Zhang and Wildemuth, 2009). The two authors worked on the qualitative data independently and the dependability of the qualitative findings enhanced due to agreement between the two coders on the themes as the same coding/observation occurred more than once using the same human instrument (Lincoln & Guba, 1985 *cf* Suter, 2012). The principle of triangulation applied by collecting data from two sources (private banks and public banks) and the agreement (convergence) between the data collected from private banks and public banks to support the conclusion enhanced credibility of the qualitative. Furthermore, to enhance the validity of the study's conclusions researchers asked five participants to review the framework and the conclusions and they judged it to be accurate, and relevant (Suter, 2012).

UNDERSTANDING OF DATA

Figure 1 presents an exploratory framework for the discussion that follows. The framework comprises two sets of factors: (a) DRS behavior construct, and (b) the antecedent factors that facilitate or impede DRS behavior. What follows is a discussion of this model and the forwarding of supporting propositions based on theory and field interviews.

Explicating DRS behavior

Exploring unique employee discretionary behaviour in public and private Indian banks we found that managers exhibit what we term as Discretionary risky-service behaviour (DRSB) to their customers. DRSB is concerned with service offered by an employee that is discretionary as it is work beyond what could reasonably be expected from an employee's job

role and risky in a sense that it encompasses bending the rules or violating the procedures for customer satisfaction. Discretionary service behaviour (bending rules for customer satisfaction) can also be an illicit behaviour that costs firms billions of dollars annually in lost revenues like service sweethearting (Brady , Voorhees , and Brusco 2011) but what we have explored is a service which is a positive behaviour and positively contributes to the organisation (Ferrin , Bligh, Kohles 2007).

This behaviour, which we term DRSB, we studied and found it in the banking industry (public and private banks) where staff members may advance the transaction on the promise from their customers that requisite amount would be deposited soon or honouring a customer's cheque in spite of the requisite amount available in the customer's account and transferring funds from their personal account and likewise. Analysing the discretionary tasks mentioned by respondents during the field interview the researchers developed the following definition:

DRSB can be defined as a bold and an unusual service offered by employees to customers where employees discount/ ignore/ do-not-follow some of their organization's regulation or procedure to serve or help a customer that moment and defer the procedure closure for a later time. Employees doing so are aware that they are taking personal risk but yet they do it with the intent to help the customer and serve the organization.

DRSB belongs to this perspective that employees are working for the customers with the firm to provide excellent service. The question arises, is it different from similar constructs like extra role behavior? Extra role behavior is defined as "behavior which benefits the organization and/or is intended to benefit the organization, which is deliberate /discretionary and which goes beyond existing role expectations" (Van Dyne, Cummings, and Parks, 1995). Unlike discretionary extra role behavior, DRSB is a bold and an unusual task

performed by employees involving customer service behavior only, not execution of job content or other non-customer focused employee role behavior.

One major distinction of DRSB is – unlike extra role behaviour where intent is the benefit of organization, the intent of DRSB is focused towards benefit to both organization as well as customer. Second distinction is bending the rules or violating the procedures (exercising empowerment and taking risk) rather than adhering to procedures and rules (professional stance) and the behavior itself is exhibited during service encounters with customers. It is this employee intent (to help the customer and build relationship), not the actual behavior or outcomes of the behavior, that distinctly defines the nature of DRS. Yet another distinction is unlike organizational citizenship behavior (OCB) research which suggests that employees will either engage in OCB or withhold OCB, depending upon their cognitive perceptions (Robinson & Morrison, 1995), DRSB display depends upon employee's affective perception.

Thus, we contend that employees can choose to exhibit varying degrees of positively intended DRSB or choose not to exhibit DRSB. In other words, employees may exhibit DRSB with the intent of helping their customers by taking personal risks or they may not be taking any risk but yet exhibiting discretionary service behavior (DSB).

DRSB Antecedents

In banks financial transactions and related services are generally carried out step by step and the preceding step in the service process must be complete before the succeeding step to be carried out. Employees who carry out a step in the financial transaction process without the completion of the previous step are doing it at a risk (personal as well as organizational). A likely risk, if the customer defaults. Yet some service employees exercise this and the question arises what are the antecedents or drivers of DRSB?

With respect to employees, little is known about the contextual factors and individual factors that may stimulate DRSB. An exploratory study and insights from various literatures namely service management (Bitner, 1990; Kelley, 1993), organization behavior (Dirks and Ferrin, 2001; Haworth and Levy, 2001; Organ, 1988), and psychology (Eisenberger et al., 1986; Rhoades, Eisenberger, and Armeli, 2001) informs our understanding of DRSB antecedents. Further analysis shows that specific employee traits, such as need for empowerment (Bowen and Lawlers, 1992), risk taking (Liliana and Lester, 2000), and customer orientation (Bitner, Booms, & Mohr, 1994; Liliana and Lester, 2000) accentuates DRSB frequency.

The antecedents to DRSB are those factors that facilitate or obstruct employees who deliberately positively and riskily influence service encounters and standards. The 12 DRSB antecedents that emerged from the data (field interview) and analysis of extant literature were organized into themes. Antecedents with prior support in a particular literature were organized under themes named for the literature from which they emerged. For antecedents unique to DRSB, we attempted to organize them under existing themes; in cases where this was not possible, we created a new theme. This process resulted in the recognition of three existing themes (job-related, reward antecedents, and trait antecedents) and the formation of a new theme that we term relationship antecedents.

Relationship antecedents

Perceptions of trust in organisation are forwarded as significant antecedent for employee extra-role behavior and it is argued that in absence of trust in organization employees are likely to withheld skill and knowledge (Dirks and Ferrin, 2001). Reychav, and Sharkie (2010), proposed perceptions of trust in management (Dirks and Ferrin, 2001; Amabile, 2005), psychological support (Dirks and Ferrin, 2001), management values and rewards as strong antecedents of employee perceptions, which in turn were positively

correlated with employee extra-role behavior (Reychav and Sharkie, 2010). McEvily, Perrone, and Zaheer (2003) developed the notion of trust as an organisational principle by specifying that trust influences organisational outcomes. This finding is also reflected in the field interview where researchers note that in employee-customer relationship trust in the customer is a significant antecedent for DRSB. A central theme of the relationship marketing literature is that firms benefit from developing strong and long-term relationships (Reynolds and Beatty 1999) with customers. The development of long-term relationships between customers and employees result into higher customer satisfaction scores (Reynolds and Beatty 1999), and increased customer loyalty (more business, word of mouth etc.). If employee perceptions of trust are favourable, then the prospects for DRSB for the benefit of the customer are likely to be present. For example one employee stated, *"I'll take such a risk in delivering service to a customer who I have complete confidence in (who has never defaulted on his promise) based on my past good relationship with him."* Another employee stated- *"I offer such a service to customers who give more business to me."* Thus DRSB is generally initiated for a high value customer. As one employee expressed *"I do this to retain a high value customer. A high value customer is important for my bank."* Yet another employee said, *"I do it for customers I have a social relationship with."* Thus we propose that relationship factor trust, value, and social bond directly influences DRSB occurrence:

Proposition 1: DRSB ensues when (a) customer is trustworthy; (b) customer is a high value customer; and (c) there is social bonding

Reward antecedents

In research on customer-employee interactions generally the common perspective is that employees are working for the firm and with customers to provide excellent service (Berry 1995). Various intrinsic and extrinsic rewards serve as motivations for providing such exemplary service. In the early stages of OCB research, the emphasis was on discretionary,

unrewarded behaviour, for which employees do not receive training to perform (Wayne and Green 1993) but later on the concept was advanced by Haworth and Levy (2001) and the researchers claimed that employees will enact and sustain OCBs only when they believe that their managers will fairly reward such behaviours. Moreover employees evaluate their contribution (importance) to the organisation through organisational reward indicators like promotional opportunity, employment continuity, and pay (Rhoades, Eisenberger, and Armeli 2001). Therefore establishing rewards as strong antecedents of employee perceptions, which in turn positively correlate with employee extra-role behaviour (Reychav and Sharkie, 2010). During field interview too it was found that material gain is a reason of DRSB occurrence but in a subtle way. As one employee stated, *"I believe such service leads to customer satisfaction and customer satisfaction is rewarded at my bank."* Another reward that employees perceive is the obligation customer feels and reciprocates by giving more business/funds or withholding funds for a specified time on request to help employee achieve his targets. As expressed by one employee, *"To a customer I offer DRS, I can anytime make a request to transfer/withhold funds to meet my targets and that would be honoured."* Thus, we propose that reward factors directly impact DRSB:

Proposition 2: DRSB ensues when (a) employee perceives reciprocity; and (b) there is likelihood of financial gain/reward.

Employee traits antecedents

Employee traits influence employee behaviour. Certain personality traits are also associated with DRSB. In particular, findings indicate that the need for empowerment, customer orientation, career orientation, and risk propensity are key factors in DRSB. The first trait factor, need for empowerment, is often discussed in HRM literature. Consider the statement-

"My customers interact with me and I only have to take these decisions and actions to serve them." We suggest this is need for empowerment and is particularly relevant to DRSB. The second employee trait that may influence DRSB is being customer centric. Consider Organ's (1988) five categories of organizational citizenship behavior (OCB) Conscientiousness, Altruism, Civic virtue, Sportsmanship, and Courtesy. Courtesy according to him means that employee treat others (fellow employees) with respect. We expand the meaning of courtesy to treating others (fellow employees and customer) with respect and not only showing respect to customers but respecting their needs. Employees with such traits are more customer centric as they have a customer-mind-set (Kennedy, Lassk, and Goolsby 2002).Consider the following statement from a relationship manager: *"At times of urgency if the bank does not supports a client with such favours then how would he perceive it 'my bank'."* Another employee said *"If you are not fulfilling a customer's need you are not customer focussed."* A similar expression echoing the sentiment was *"If I do not give him such service we'll lose him because there is no differentiation in the product we can only differentiate on service."*

The third trait antecedent that we uncovered is risk-taking tendency, which refers to employees' propensity of taking risk and being extremely service oriented while delivering service. For example an employee remarked *"Of course its risk, but you can do this for strong and long term relations."* The fourth trait antecedent discovered is career orientation. Consider the following statement from an employee, *"I believe this would help me in growing and advancing in my job."* Based on insights from the exploratory study and substantiation from several literatures, we propose:

Proposition 3: DRSB is greater when employees are (a) higher in need for empowerment; (b) highly customer oriented; (c) highly career oriented and (d) higher in risk taking

Job related Antecedents

The quality of the relationship between an employee and the employee's immediate supervisor, characterised by mutual trust, respect and obligation was found to be related to OCB (e.g. Wong, Ngo, and Wong 2003). Organizational support is an important factor in discretionary behavior as well as job satisfaction. Psychological support and trust in manager/management has a strong relationship with perceptions of trust in the organisation and employee extra role behavior (Kerkhof, Winder, and Klandermans 2003; Eisenberger et al., 1986). Field interview also found evidence that if employees believe that they are being supported by their leaders in carrying out their duties this affects their extra-role behavior (Amabile, 2005). The statement echoing the sentiment is, *"I go to that extent because my branch manager supports me completely."* The following comment made by an employee exemplifies the effect of appraisal process on DRS: *"I would probably not do this if such behavior is not factored in the appraisal process."* Another employee echoing the same sentiment said, *"I would not do so if the appraisal is not transparent and fair."* Thus we propose:

Proposition 4: DRSB occurs when (a) supervisor (bank manager) supports the employee and shares the risk/responsibility (b) employee is satisfied with the job (c) DRSB would not occur if employees are dissatisfied with the appraisal process.

EXPLORATORY MODEL OF DISCRETIONARY RISKY-SERVICE BEHAVIOR (DRSB)

An exploratory model of DRS behaviour is now proposed based on the findings and the above four propositions (Figure 1). Service employees exhibit extra-role/discretionary behaviour. Complex services and contexts of close customer-employee interactions such as banks allow employees to exercise more customer-facing discretions. Employee's customer orientation, their risk propensity, career orientation, need for empowerment, and their trust in their customer affects discretionary service they offer to the customer.

Insert fig 1 here

Despite the considerable progress that has been made in understanding extra-role/discretionary employee behaviour in service firms and its determinants, research scholars appear to be seeking a deeper understanding of it. A theory of employee-customer relationship grounded in extra-role/discretionary behaviour holds tremendous promise in accomplishing these ends.

Our research surfaces the existence of distinctive employee behaviour in banks - DRSB and our exploratory model provides first-ever insights into antecedents of DRSB and, in turn, offers guidance to managers who may be struggling to advance employee-customer relationship. These findings are important because DRSB phenomenon forces managers to examine its impact on relational outcomes and support /encourage them in building strong bonds with customers.

Our exploratory model suggest that there are four measurable employee traits namely risk propensity, need for empowerment, career orientation, and customer orientation that enhances DRSB frequency. Taken together our results indicate that it may be possible to encourage DRSB in a service organization. We discuss the implications of these findings next.

MANAGERIAL IMPLICATIONS

In addition to contributing to theory development in service employee behavior, the propositions in this paper have practical implications for managing relational service encounters. Employee traits like empowerment, risk-taking, career orientation, and customer orientation are positively related to DRSB therefore, service firms should depute contact personnel with these traits for high value customers. Also managers are advised to see how employee rewards are structured, whether these are in line for taking risk and improving

relationships with customers or not. Another important implication is for setting up systems for monitoring employee-customer relationship strength periodically.

DISCUSSION AND CONTRIBUTION

This study has endeavoured to elucidate the organizational phenomenon of customer oriented discretionary risky-service through supplying a definition of DRSB designed to act as groundwork for future studies. Furthermore, analysis of field interviews and extant theory led to the highlighting of a series of individual, organizational, and relationship factors that appear to act as antecedents to DRSB. The determinants of the DRSB as indicated by the respondents in this study are consistent with reward expectation (Eisenberger et al. 1986) and trust in management (Dirks and Ferrin, 2001) dimension of employee discretionary behaviour or extra-role behavior appearing in the international literature.

Another important determinant relationship (Dess and Shaw, 2001) that affect employees' willingness to engage in discretionary extra-role behaviour has a mention in international literature and pertains to employer-employee relationship but relationship dimension of our study is an individual determinant and pertains to employee-customer relationship.

The main contribution of our study is relationship antecedent, which may be taken as a starting point for relationship advancement research. Another contribution is a new perspective to the trust construct. As DRSB involves employee trusting the customer and taking personal risk in serving him/her, traditional relationship paradigms where trust in relationship only means customer's trust on the employee/organization do not apply and hence offers a new perspective to employee-customer relationship research.

The framework and propositional inventory proposed is exploratory in nature but grounded in both existing theory and insights gained through 35 field interviews. The aim of the study is not to provide a definitive description, and model of DRS behavior but rather to supply the

first tentative steps toward subsequent testing. In this regard, the aim of theory construction has been achieved and the next stage of theory testing is left for future studies. Clearly, the developed framework and propositions require further empirical work before more descriptive insights into DRSB can be provided.

LIMITATIONS AND DIRECTIONS FOR FUTURE RESEARCH

This paper can be improved upon in several ways. First, we still need to explore whether or not DRSB is exhibited by employees in more regulatory banking environments (like the USA and the UK). Further the above discussion is clearly limited by lack of empirical data. While the exploratory data suggest the general antecedents of DRSB, the specific propositions have not been tested. This is clearly the next, necessary step.

Additionally, the findings of this study should be interpreted in light of the following limitations. First, the objective of the study was to explore new/unique employee discretionary service behaviour and investigate its antecedents so, there is limitation on the generalizability of the results. Though exploratory study is the best method when the purpose of research is to gain familiarity with a phenomenon or acquire new insight into it. As our study was conducted among bank employees (relationship managers) employed in a particular country. Therefore, generalizations beyond the specific context of this research must be guarded. Replications among bank employees (relationship managers) in other countries would provide further insights and would broaden the understanding for further generalizations. Also, the sample features (i.e. small sample, and use of confederates) warrant caution before generalizing the results beyond the population studied.

Although the present model focuses on the antecedents of employee's DRSB, future research might include the study of both antecedents and consequences of this service behaviour. It would be interesting to study DRSB from customer perspective as it would help in addressing – "how value emerges for customers and how through a sense-making process

customers construct their experience of value of a service provider's participation in their activities and tasks" (Heinonen et al. , 2010).

Acknowledgments

The authors gratefully acknowledge and thank Dr. Kristina Heinonen (Guest Editor) and the two anonymous reviewers for their valuable insight and helpful comments on earlier version of the article.

Biographical Details (if applicable):

Shilpa Sharma Bhaskar is a Doctoral Candidate in Delhi School of Management at Delhi Technological University, New Delhi, India and Assistant Professor at Apeejay School of Management, New Delhi, India. Her teaching interests include Marketing Management, Services Marketing, Consumer Behaviour, Brand management, and Professional Selling. Her research interests outline Loyalty Programs, Employee satisfaction- Customer satisfaction link, Employer Branding, and Consumer Behavior. Shilpa Sharma Bhaskar is the corresponding author and can be contacted at: shilpasharmabhaskar@gmail.com

Dr. Shikha N Khera is Assistant Professor of OB and HRM at Delhi School of Management at the Delhi Technological University, New Delhi, India. Her teaching interest includes Organizational Behaviour, Cross cultural HRM, and Training and Development. Her research interest outlines Knowledge Management and Human Resources, Employee extra-role behaviour, and Customer Relationship Management.

References

Amabile, T. (2005), "Creativity: it's all about routine", *The Australian Financial Review*, 12 July, p. 59 (Harvard Management Update).

Analoui, F. (1995), "Workplace sabotage: its styles, motives and management", *Journal of Management Development*, Vol. 14, No. 7, pp. 48-65.

Bendapudi, Neeli and Leonard L. Berry (1997), "Customers' Motivations for Maintaining Relationships with Service Providers," *Journal of Retailing*, Vol. 73, No.1, pp. 15-37.

Berry, Leonard L. (1995), *On Great Service*. New York, NY: The Free Press.

Bitner, M. J., Booms, B. H., & Tetreault, M. S. (1990), "The service encounter: Diagnosing favorable and unfavorable incidents", *Journal of Marketing*, Vol. 54, pp. 71-84.

Bitner, M. J., Booms, B. H., & Mohr, L. A. (1994), "Critical service encounters: The employee's viewpoint", *Journal of Marketing*, Vol. 58, pp. 95-106.

Blancero, D. & Johnson, S. A. (1997), "Customer service employees and discretionary service behavior: A psychological contract model", (CAHRS Working Paper #97-07). Ithaca, NY: Cornell University, School of Industrial and Labor Relations, Center for Advanced Human Resource Studies. http://digitalcommons.ilr.cornell.edu/cahrswp/149

BoliNo, M.C. (1999), "Citizenship and impression management: good soldiers or good actors?", Academy of Management Review, Vol. 24, No. 1, pp. 82-98.

Bowen, D. E., & Lawler, E. E. (1992), "The empowerment of service workers: What, why, how, and when", *Sloan Management Review*, Spring: 31-39.

Brady, Michael, K.; Voorhees, Clay M., and Brusco, Michael J. (2011), "Service Sweethearting: Its Antecedents and Customer Consequences", *Journal of Marketing* Article Postprint.

Camarero, C. (2007), "Relationship orientation or service quality? What is the trigger of performance in financial and insurance services?", *The International Journal of Bank Marketing*, Vol. 25, No. 6, pp. 406-426.

Chaoluck, P., Medlin, C. J., and Conduit, J. (2013), "The Distinctions between In-Role and Extra-Role Behaviours in Service Quality: The Perceptions of Frontline Service Employees and Customers", Australia and New Zealand Marketing Academy Conference, Auckland, New Zealand. Retrieved January 22, 2014 from http://anzmac.org/conference/2013/papers/anzmac2013-228.pdf

Claycomb, C. and Martin, C.L. (2002), "Building customer relationships: an inventory of service providers' objectives and practices", *Journal of Services Marketing*, Vol. 16, No. 7, pp. 615-35.

De Wulf, K., Odekerken-Schro¨der, G. and Iacobucci, D. (2001), "Investments in consumer relationships: a cross-country and cross-industry exploration", *Journal of Marketing*, Vol. 65, pp. 33-50.

Dess, Gregory G. and Shaw, Jason D. (2001), "Voluntary Turnover, Social Capital, and Organizational Performance", *The Academy of Management Review*, Vol. 26, No. 3, pp. 446-456.

Dimitriadis, S. (2010), "Testing perceived relational benefits as satisfaction and behavioural outcomes drivers", *International Journal of Bank Marketing*, Vol. 28, No. 4, pp. 207-13.

Dirks, K. and Ferrin, D. (2001), "The role of trust in organisational settings", *Organisation Science*, Vol. 12, No. 4, pp. 450-67.

Dubinsky, A.J. (1994), "What marketers can learn from the tin man?", *Journal of Services Marketing*, Vol. 8, No. 2, pp. 36-45.

Eisenberger, R., Huntington, R., Hutchison, S. and Sowa, D. (1986), "Perceived organisational support", *Journal of Applied Psychology*, Vol. 71, pp. 500-7.

Ferrin, D., Bligh, M. and Kohles, J. (2007), "Can I trust you to trust me? A theory of trust, monitoring and cooperation in interpersonal and intergroup relations", *Group and Organisation Management*, Vol. 32, pp. 465-99.

Glaser, B. G., & Strauss, A. L. (1967), *The discovery of grounded theory: Strategies for qualitative research,* New York, NY: Aldine.

Gremler Dwayne D., Gwinner Kevin P., (2008), "Rapport-Building Behaviors Used by Retail Employees", *Journal of Retailing*, Vol. 84, No. 3, pp. 308-324.

Gro¨nroos, C. (1994), "From marketing mix to relationship marketing: towards a paradigm shift in Marketing", *Management Decision*, Vol. 32, No. 2, pp. 4-20.

Gurau, C. (2003), "Tailoring e-service quality through CRM", *Managing Service Quality*, Vol. 13, No. 6, pp. 520-31.

Harris, Lloyd C., and Ogbonna, Emmanuel (2002), "Exploring service sabotage: The antecedents, types and consequences of frontline, deviant, anti service behaviors", *Journal of Service Research*, Vol. 4, No. 3 , 163-183.

Haworth, C.L. and Levy, P.E. (2001), "The importance of instrumentality beliefs in the prediction of organisational citizenship behaviours", *Journal of Vocational Behaviour*, Vol. 59, No. 1, pp. 64-75.

Heinonen, K., Strandvik, T., Mickelsson, K-J., Edvardsson, B., Sundstro¨m, E., and Andersson, P. (2010), "A customer-dominant logic of service", *Journal of Service Management*, Vol. 21, No. 4, pp. 531-548.

Henning-Thurau, T. (2004), "Customer orientation of service employees: its impact on customer satisfaction, commitment, and retention", *International Journal of Service Industry Management*, Vol. 15, No. 5, pp. 460-78.

Hennig-Thurau, Thorsten ; Gwinner, Kevin P ; Gremler, Dwayne D. (2002), "Understanding relationship marketing outcomes: An integration of relational benefits and relationship quality", *Journal of Service Research*, Vol. 4, No. 3, pp. 230-247.

Heskett, J., Sasser, E. and Schlessinger, L. (1997), *The Service Profit Chain*, Free Press, New York, NY.

Jham, V. and Khan, K.M. (2008), "Determinants of performance in retail banking: perspectives of customer satisfaction and relationship marketing", *Singapore Management Review*, Vol. 30, No. 2, pp. 35-45.

Kelley, S. W. (1993), "Discretion and the service employee", *Journal of Retailing*, Vol. 69, pp. 104-125.

Kennedy, Karen Norman; Lassk, Felicia G; Goolsby, Jerry R. (2002), "Customer mind-set of employees throughout the organization", *Academy of Marketing Science Journal*, Vol. 30, No.2, pp. 159.

Kerkhof, P., Winder, A. and Klandermans, B. (2003), "Instrumental and relational determinants of trust in management among members of Works Councils", *Personnel Review*, Vol. 32, No. 5, pp. 623-37.

Liao, Hui; Joshi, Aparna and Chuang, Aichia (2004), "Sticking out Like a Sore Thumb: Employee Dissimilarity and Deviance at Work," *Personnel Psychology*, Vol. 57, No. 4, pp. 969-1000.

Liliana L. Bove, Lester W. Johnson, (2000) "A customer-service worker relationship model", *International Journal of Service Industry Management*, Vol. 11, No. 5, pp. 491 – 511.

McEvily, B., Perrone, V. and Zaheer, A. (2003), "Trust as an organizing principle", *Organisation Science*, Vol. 14, No. 1, pp. 91-103.

Lincoln, Y. S., & Guba, E. G. (1985), *Naturalistic observation,* Thousand Oaks, CA: Sage.

Molina, A., Martı´n-Consuegra, D. and Esteban, A., (2007), “Relational benefits and customer satisfaction in retail banking”, *International Journal of Bank Marketing*, Vol. 25, No. 4, pp. 253-71.

Mosavi, S. A , Abedi, M. and Ghaedi, M. (2013), “Reviewing the relationship between perception of trust in organization with employee extra-role behaviour”, *African Journal of Business Management*, Vol. 7, No. 35, pp. 3620-3629.

Organ, D.W. (1988), *Organizational Citizenship Behavior: The Good Soldier Syndrome*. Lexington Books, Lexington, MA.

Podsakoff, P.M. and Mackenzie, S.B. (1997), "The impact of organizational citizenship behavior in organizational performance: review and suggestion for future research", *Human Performance*, Vol. 10, pp. 133-51.

Price, Linda L. and Eric J. Arnould (1999), “Commercial Friendships: Service Provider-Client Relationships in Context,” *Journal of Marketing*, Vol. 63, No. 4, pp. 38-56.

Reichheld, F.E. and Sasser, W.E. Jr (1990), “Zero defections: quality comes to service”, *Harvard Business Review*, Vol. 68, No. 9, pp. 105-11.

Reychav, I, Sharkie, R. (2010),"Trust: an antecedent to employee extra-role behaviour", *Journal of Intellectual Capital*, Vol. 11, No. 2, pp. 227 – 247.

Reynolds, Kristy E. and Sharon E. Beatty (1999), “Customer Benefits and Company Consequences of Customer-Salesperson Relationships in Retailing,” *Journal of Retailing*, 75, No. 1, pp. 11-32.

Rhoades, L.; Eisenberger, R.; Armeli, S. (2001), “Affective commitment to the organization: The contribution of perceived organizational support”, *Journal of Applied Psychology*, Vol. 85, No. 5, pp. 825-836.

Robinson, S. L., & Morrison, E. W. (1995), "Psychological contracts and OCB: The effect of unfulfilled obligations on civic virtue behaviour", *Journal of Organizational Behavior*, Vol. 16, pp. 289-298.

Rosenbaum, Mark S. (2009), "Exploring commercial friendships from employees' perspectives", *The Journal of Services Marketing*, Vol. 23, No.1, pp. 57-66.

Roy, S.K. and Shekhar, V. (2010), "Dimensional hierarchy of trustworthiness of financial service providers", *International Journal of Bank Marketing*, Vol. 28, No. 1, pp. 47-64.

Smith, C.A., Organ, D.W. and Near, J.P., (1983), "Organizational citizenship behavior: its nature and antecedents", *Journal of Applied Psychology*, Vol. 68, No. 4, pp. 653-63.

So, S.L.M., and Speece, M.W., (2000), "Perceptions of relationship marketing among account managers of commercial banks in a Chinese environment", *International Journal of Bank Marketing*, Vol. 18, No. 7, pp. 315-27.

Suter, W. Newton (2012), Qualitative Data, Analysis, and Design, *Introduction to Educational Research: A Critical Thinking Approach*, Second Edition. SAGE Publications, Inc., pp. 342-386.

Tarnowski, Joseph (2008), "Good Night, Sweethearting", *Progressive Grocer*, Vol. 87, No. 6, pp. 126.

Van Dyne, L., Cummings, L. L., & Parks, J. M. (1995), "Extra-role behaviors: In pursuit of construct and definitional clarity (a bridge over muddied waters)", In L. L. Cummings & B. M. Staw (Eds.), *Research in organizational behavior*, 17: 215-285. Greenwich, CT: JAI Press.

Wallace, E., Chernatony, L. D., Buil, I., (2011), "Within-role, extra-role and anti-role behaviours in retail banking", *International Journal of Bank Marketing*, Vol. 29, No. 6, pp. 470-488.

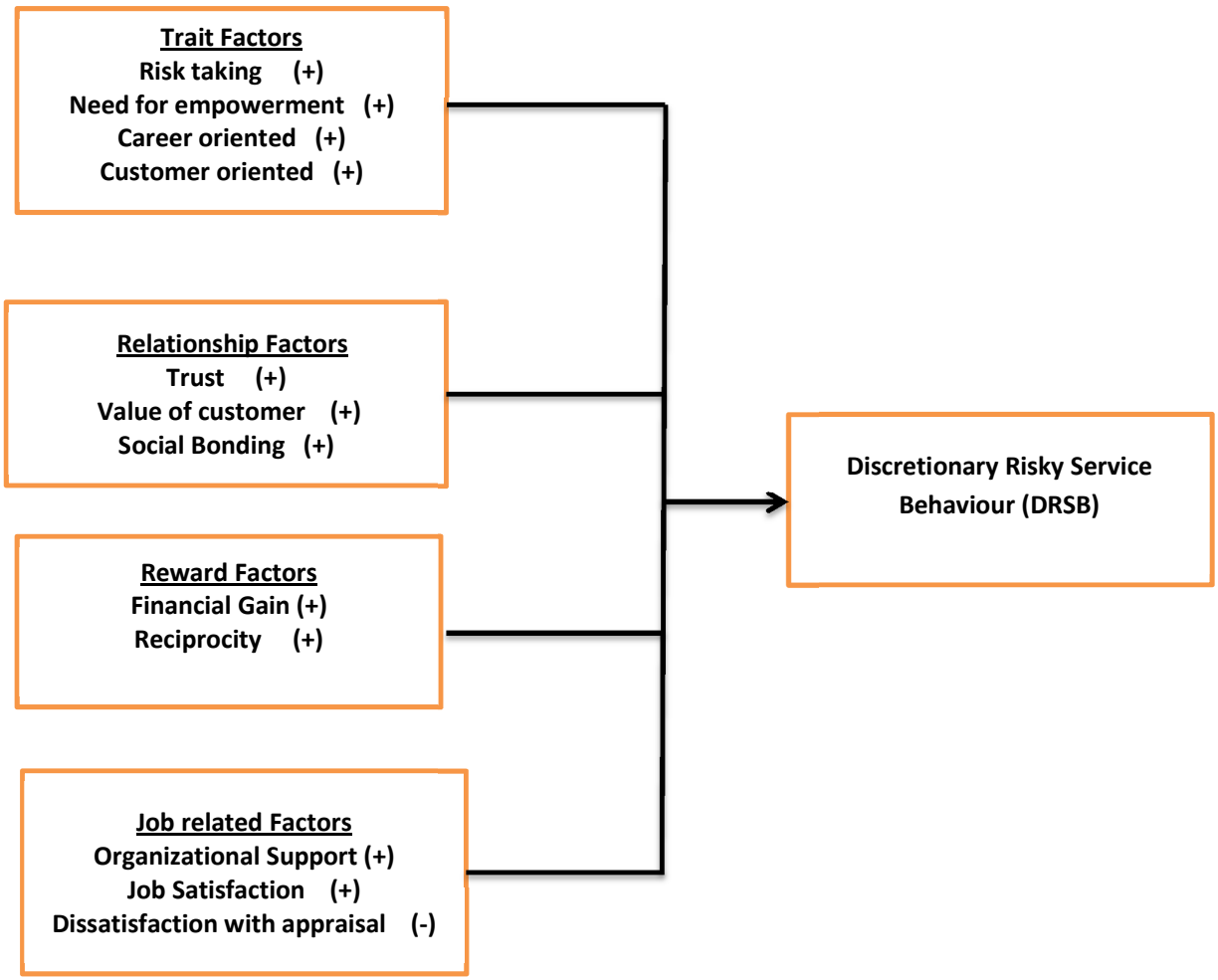
Wang, M-L. (2010), "Does organisational support promote citizenship? The moderating role of market-focused HRM", *The Service Industries Journal*, Vol. 30, No. 7, pp. 1077-95.

Wayne, S.J. and Green, S.A. (1993), "The effects of leader-member exchange on employee citizenship and impression management behaviour", *Human Relations*, Vol. 46, No. 12, pp. 1431-40.

Wong, Y.T., Ngo, H.Y. and Wong, C.S. (2003), "Antecedents and outcomes of employees' trust in Chinese joint ventures", *Asia Pacific Journal of Management*, Vol. 20, No. 4, pp. 481-99.

Zhang, Y. and Wildemuth, B. M. ,(2009), "Qualitative Analysis of Content" in B. Wildemuth (Ed.) *Applications of Social Research Methods to Questions in Information and Library*. Retrieved April 18, 2012 from https://www.ischool.utexas.edu/~yanz/Content_analysis.pdf

Figure 1: Exploratory Framework of Antecedents of Discretionary Risky-Service Behavior (DRSB)



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

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 3, Issue 6, June 2014

Data Mining Techniques for Intrusion Detection: A Review

Abhaya[1], Kaushal Kumar[2], Ranjeeta Jha[3], Sumaiya Afroz[4]

M.Tech (IS), Department of Computer Scince & Engineering, Birla Institute of Technology, Mesra (Ranchi), India[1,3,4]

M.Tech (SE), Department of Software Engineering, Delhi Technological University, Delhi, India[2]

Abstract: With the dramatically development of internet, Security of network traffic is becoming a major issue of computer network system. Attacks on the network are increasing day-by-day. The most publicized attack on network traffic is considered as Intrusion. Intrusion detection system has been used for ascertaining intrusion and to preserve the security goals of information from attacks. Data mining techniques are used to monitor and analyze large amount of network data & classify these network data into anomalous and normal data. Since data comes from various sources, network traffic is large. Data mining techniques such as classification and clustering are applied to build Intrusion detection system. An effective Intrusion detection system requires high detection rate, low false alarm rate as well as high accuracy. This paper presents the review on IDS and different Data mining techniques applied on IDS for the effective detection of pattern for both malicious and normal activities in network, which helps to develop secure information system.

Keywords: Intrusion Detection System; Anomaly Detection; Misuse Detection; Data mining; Clustering; Classifications

I. INTRODUCTION

With the speedy escalation of Internet, there is enhancement in the lifestyle of people but at the cost of threats, which are created by either individuals or any organization. They are used to break the security of network.

Security means degree of protection given to the network or system. The main goals of security are confidentiality, Integrity and availability of data [1]. Attacks on network can be referred as Intrusion. Intrusion means any set of malicious activities that attempt to compromise the security goals of the information.

In early days, only conventional approaches were used for network such as encryption, firewalls, virtual private network etc but they were not enough to secure network completely. It is difficult to depend completely on static defense techniques. This increases the need for dynamic technique, which can be monitors system and identify illegal activities. Thus to enhance the network security dynamic approach is introduced and known as Intrusion Detection System. Intrusion detection system collects online information from the network after that monitors and analyzes these information and partitions it into normal & malicious activities, provide the result to system administrator [2].

IDS is the area, where Data mining is used extensively, this is due to limited scalability, adaptability and validity. In IDS data is collected from various sources like network log data, host data etc. Since the network traffic is large, the analysis of data is too hard. This give rise to the need of using IDS along with different Data mining techniques for intrusion detection . Lee & Salvatore J. Stolfo, Columbia University were first to apply Data mining techniques in the IDS [3]. Data mining techniques such as classification and clustering easily extract the information from large dataset.

The remaining part of the paper is structured in this way. Section I introduction, Section II review the related work on IDS using Data mining techniques, Section III explanation of IDS. In Section IV, Data mining and its techniques which are used in IDS are described and finally Section V brings us to the conclusion.

II. LITERATURE REVIEW

Xiang M.Y. Chang et.al.(2004) [4], designed a multiple-level tree classifier for Intrusion detection system and increase the detection rate. Classifier is more efficient in case of known attacks but for unknown vulnerabilities it gives low detection rate. **Peddabachigiri S. et.al.**(2007)[5], proposed a model of intrusion detection system combining decision tree and support vector machine (DTSVM) classification techniques and produces high detection rate. **Mrutyunjaya panda et. al.**(2008)[6], compares different data mining techniques for intrusion detection system and found that accuracy & performance of Naïve bayes classifier for all classes is better than the accuracy obtained in the case of different Decision tree algorithm but Decision tree is robust in detecting unknown intrusions in comparison to Naïve bayes classification algorithm. **M.Govindarajan et.al.**(2009)[7], proposed new K-nearest neighbour classifier applied on Intrusion detection system and evaluate performance in term of Run time and Error rate on normal and malicious dataset. This new classifier is more accurate than existing K-nearest neighbour classifier. **Mohammadreza Ektela et.al.**(2010)[8], used Support Vector Machine and classification tree Data mining technique for intrusion

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

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 3, Issue 6, June 2014

detection in network. They compared C4.5 and Support Vector Machine by experimental result and found that C4.5 algorithm has better performance in term of detection rate and false alarm rate than SVM, but for U2R attack SVM performs better. **Song Naiping et.al.**(2010)[9], studied on Intrusion detection based on Data mining. Here, types of IDS means Misuse detection and Anomaly detection are described by the author along with different Data mining techniques which are used to build IDS. **T. Velmurugan et.al.**(2010)[10], compute the complexity between k-means and k-medoids clustering algorithm for uniform and normal distribution of data points and concluded that average time taken by k-Means algorithm is more in both the cases.

P. Amudha et.al.(2011)[11], observed that Random forest gives better detection rate, accuracy and false alarm rate for Probe and DOS attack & Naive Bayes Tree gives better performance in case of U2R and R2L attack. Also the execution time of Naive Bayes Tree is more as compared to other classifier. **Deepthy k Denatious et.al.**(2012)[1], describe different data mining techniques applied for detecting intrusions. Also describe the classification of Intrusion detection system and its working. For large amount of network traffics, clustering is more suitable than classification in the domain of intrusion detection because enormous amount of data needed to collect to use classification.

R. China Appala Naidu et.al.(2012)[12], used three Data mining techniques SVM, Ripper rule and C5.0 tree for Intrusion detection and also compared the efficiency. By experimental result, C5.0 decision tree is efficient than other. All the three Data mining technique gives higher than 96% detection rate. **Roshan Chitrakar et.al.**(2012)[13], proposed a hybrid approach to intrusion detection by using k-Medoids clustering with Naïve Bayes classification and observed that it gives better performance than K-Means clustering technique followed by Naïve Bayes classification but also time complexity increases when increase the number of data points.

Roshan Chitrakar et.al.(2012)[14], proposed a hybrid approach of combining k-Medoids clustering with Support Vector Machine classification technique and produced better performance compared to k-Medoids with Naïve Bayes classification. The approach shows improvement in both Accuracy and Detection Rate while reducing False Alarm Rate as compared to the k- Medoids clustering approach followed by Naïve bayes classification technique. **Sumaiya Thaseen et.al.**(2013)[15], analyzed different tree based classification techniques for IDS. Experimental results show that Random tree model reduces false alarm rate and has highest degree of accuracy.

III. INTRUSION DETECTION SYSTEM

The concept of IDS was proposed by Denning(1987), to identify, detect and trace the intrusion[11]. An IDS is a combination of software and hardware which are used for detecting intrusion[1]. It gathers and analyzes the network traffic & detect the malicious patterns and finally alert to the proper authority. The main function of IDS includes:[16]

- Monitoring and analyzing the information gathered from both user and system activities.
- Analyzing configurations of system and evaluating the file integrity and system integrity.
- For static records, it finds out the abnormal pattern.
- To recognize abnormal pattern, it use static records and alert to system administrator.

A. Classification of IDS

According to techniques used for intrusion detection based on whether attack's patterns are known or unknown, IDS classified into two category [2][17]:

(1) Misuse detection
(2) Anomaly detection

Misuse detection: It is Signature based IDS where detection of intrusion is based on the behaviors of known attacks like antivirus software. Antivirus software compares the data with known code of virus. In Misuse detection, pattern of known malicious activity is stored in the dataset and identify suspicious data by comparing new instances with the stored pattern of attacks.

Anomaly detection: [16][18][1] It is different from Misuse detection. Here baseline of normal data in network data in network eg load on network traffic, protocol and packet size etc is defined by system administrator and according to this baseline, Anomaly detector monitors new instances. The new instances are compared with the baseline, if there is any deviation from baseline, data is notified as intrusion. For this reason, it is also called behavior based Intrusion detection system.

TABLE 1
COMPARISON BETWEEN MISUSE DETECTION AND ANOMALY DETECTION

Signature – Based(Misuse Detection)	**Behaviour–Based(Anomaly Detection)**
Advantages	**Advantages**
-Higher Detection rate, Accuracy for known behaviors. -Simplest and effective method. -Low False alarm rate.	-can examine unknown and more complicated intrusions. - Rate of Missing report is low. -Detect new and unforeseen vulnerabilities.
Disadvantages	**Disadvantages**
- It can detect only known attacks. - Needs a regular update of the rules which are used. - Often no differentiation between an attack attempt and a successful attack. - Rate of Missing report is high.	- Needs to be trained and tuned model carefully, otherwise it tends to false – positives -low detection rate and high false alarm rate. - It can't identify new attacks because intrusion detection depends upon latest model.

B. Working of Intrusion Detection System [1]

Author presents 4-steps for working of IDS.

1) *Data Acquisition:* Data is collected from various sources by using particular software.
2) *Feature Selection:* Huge amount of data is collected from network traffic. So dataset for IDS becomes

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

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 3, Issue 6, June 2014

large. For working on large dataset generate feature vectors, which contains only necessary data.

3) *Analysis:* In this step, Collected data is analyzed to determine whether data is suspicious or not. Here, various Data mining techniques are used for Intrusion detection.
4) *Action:* IDS alarms the administrator about attack which has been detected.

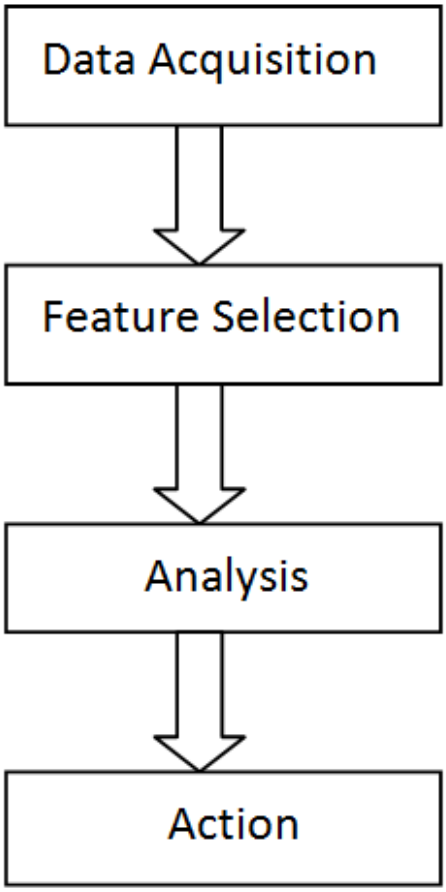


Fig. 1. Working of Intrusion Detection System

C. Performance Measurement of IDS [17] [19]

There are some primary factors which are used during performance measurement of Intrusion detection system.
True positive (TP): The total number of normal data which are detected as a normal data during intrusion detection process.

True negative (TN): In Intrusion detection, number of detected abnormal data which are actually abnormal data in dataset.

False positive (FP): Or false alarm, total number of detected normal data but they are actual attack.

False negative (FN): Number of detected abnormal instances but in real they are normal data.

Performance of IDS is measured in terms of detection rate, accuracy and false alarm rate.

Detection Rate (DR) = (TP/TP+FN) x 100% *False Alarm Rate (FAR)* =FP/Number of Attacks *Accuracy* = (TP+TN/TP + TN + FP + FN)x 100%

IV. DATA MINING BASED INTRUSION DETECTION SYSTEM

Data mining is the activity of extracting relevant information from a large amount of data.[20]

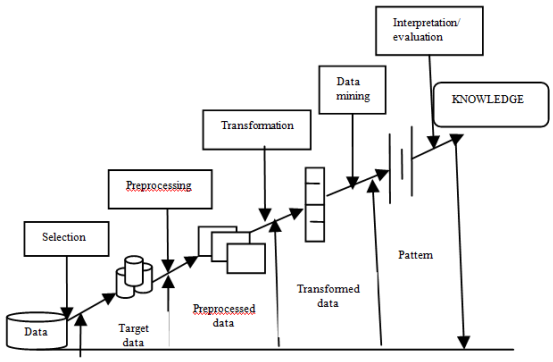


Fig. 2. Data Mining

Network traffic is massive and information comes from different sources, so the dataset for IDS becomes large. Hence the analysis of data is very shard in case of large dataset. Data mining techniques are applied on IDS because it can extract the hidden information and deals with large dataset. Presently Data mining techniques plays a vital role in IDS. By using Data mining techniques, IDS helps to detect abnormal and normal patterns.

This section describes different Data mining techniques such as clustering and classification, which are used in IDS to obtain information about vulnerability by monitoring network data.[1]

A. Classification [1]

Classification is the task of taking each and every instances of dataset under consideration and assigning it to a particular class normal and abnormal means known structure is used for new instances. It can be effective for both misuse detection and anomaly detection, but more frequently used for misuse detection. Classification categorized the datasets into predetermined sets. It is less efficient in intrusion detection as compared to clustering. Different classification techniques such as decision tree, naive bayes classifier, K-nearest neighbour classifier, Support vector machine etc are used in IDS.

1) Decision Tree [21]

Decision tree is a recursive and tree like structure for expressing classification rules. It uses divide and conquer method for splitting according to attribute values. Classification of the data proceeds from root node to leaf node, where each node represents the attribute and its value & each leaf node represent class label of data. Tree based classifier have highest performance in case of large dataset. Different decision tree algorithms are described below[6]

ID3 algorithm

It is famous decision tree algorithm developed by Quinlan. ID3 algorithm basically attribute based algorithm that constructs decision tree according to training dataset. The attribute which has highest information gain is used as a root of the tree.

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

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 3, Issue 6, June 2014

J48 algorithm
It is based on ID3 algorithm and developed by Ross Quinlan. In WEKA, C4.5 decision tree algorithm is known as J48 algorithm. It construct decision tree using information gain, attribute which have highest information gain is selected to make decision. The main disadvantage of this algorithm is that it takes more CPU time and memory in execution. Another different tree based classifier [15]:

AD Tree
Alternating decision tree is used for classification. AD Tree have prediction node as both leaf node and root node.

NB Tree
NB Tree algorithm uses both decision tree and naive bayes classifier. Root node uses decision tree classifier and leaf nodes uses naive bayes classifier.

Random Forest [22]
Random Forest is first introduced by Lepetit et.al. and it is ensemble classification technique which consists of two or more decision trees. In Random Forest, every tree is prepared by randomly select the data from dataset. By using Random Forest improve the accuracy and prediction power because it is less sensitive to outlier data. It can easily deals with high dimensional data.

2) K-Nearest Neighbor [21]
It is one of the simplest classification technique. It calculates the distance between different data points on the input vectors and assigns the unlabeled data point to its nearest neighbor class. K is an important parameter. If k=1, then the object is assigned to the class of its nearest neighbor. When value of K is large, then it takes large time for prediction and influence the accuracy by reduces the effect of noise.

3) Naive Bayes classifier [13]
Naive Bayes classifier is probabilistic classifier. It predicts the class according to membership probability. To derive conditional probability, it analyzes the relation between independent and dependent variable.
Bayes Theorem:

$$P(H/X)=P(X/H).P(H)/P(X)$$

Where, X is the data record and H is hypothesis which represents data X and belongs to class C. P(H) is the prior probability, P(H/X) is the posterior probability of H conditioned on X and P(X/H) is the posterior probability of X conditioned on H.

Construction of Naive Bayes is easy without any complicated iterative parameter. It may be applied to large number of data points but time complexity increases.

4) Support Vector Machine [8]
Support Vector Machine is supervised learning method used for prediction and classification. It separate data points into two classes +1 and -1 using hyperplane because it is binary classification classifier. +1 represents normal data and -1 for suspicious data.
Hyperplane can be expressed as: W. X +b=0

Where $W=\{w_1,w_2,.....,w_n\}$ are weight vector for n attributes $A=\{A_1,A_2,.....,A_n\}$, $X=\{x_1,x_2,.....,x_n\}$ are attribute values and b is a scalar. The main goal of SVM is to find a linear optimal hyper plane so that the margin of separation between the two classes is maximized. The SVM uses a portion of the data to train the system.

B. Clustering [1]
Since the network data is too huge, labelling of each and every instances or data points in classification is expensive and time consuming. Clustering is the technique of labelling data and assign into groups of similar objects without using known structure of data points. Members of same cluster are similar and instances of different clusters are different from each other. Clustering technique can be classified into four groups: Hierarchical algorithm, Partitioning algorithm, Grid based algorithm and Density based algorithm. Some clustering algorithms are explained here.

1) K-Means Clustering algorithm [23][13]
K-Means clustering algorithm is simplest and widely used clustering technique proposed by James Macqueen. In this algorithm, number of clusters K is specified by user means classifies instances into predefined number of cluster. The first step of K-Means clustering is to choose k instances as a center of clusters. Next assign each instances of dataset to nearest cluster. For instance assignment, measure the distance between centroid and each instances using Euclidean distance and according to minimum distance assign each and every data points into cluster. K –Means algorithm takes less execution time, when it applied on small dataset. When the data point increases to maximum then it takes maximum execution time. It is fast iterative algorithm but it is sensitive to outlier and noise.

2) K-Medoids clustering algorithm [13]
K-Medoids is clustering by partitioning algorithm as like as K-means algorithm. The most centrally situated instance in a cluster is considered as centroid in place of taking mean value of the objects in K-Means clustering. This centrally located object is called reference point and medoid. It minimizes the distance between centroid and data spoints means minimize the squared error. K-Medoids algorithm performs better than K-Means algorithm when the number of data points increases to maximum. It is robust in presence of noise and outlier because medoid is less influenced by outliers, but processing is more expensive.

V. CONCLUSION

On the basis of detection rate, accuracy, execution time and false alarm rate, the paper has analyzed different classification and clustering data mining techniques for intrusion detection. According to given necessary parameter, execution time of Support vector machine is less and produces high accuracy with smaller dataset, while construction of Naive Bayes classifier is easy. Also decision tree has high detection rate in case of large dataset. In clustering techniques, execution time of K-Means clustering algorithm is less in case of small dataset, but when number of data point increases, K-Medoids performs better.

REFERENCES

[1] Deepthy K Denatious & Anita John, "Survey on Data Mining Techniques to Enhance Intrusion Detection", International Conference on Computer Communication and Informatics (ICCCI - 2012), Jan. 10 – 12, 2012, Coimbatore, INDIA

[2] Rung-Ching Chen, Kai-Fan Cheng and Chia-Fen Hsieh, "Using Rough Set And Support Vector Machine For Network Intrusion Detection", International Journal of Network Security & Its Applications (IJNSA), Vol 1, No 1, April 2009

[3] Deepak Upadhyaya and Shubha Jain, "Hybrid Approach for Network Intrusion Detection System Using K-Medoid Clustering and Naïve Bayes Classification", IJCSI International Journal of Computer Science Issues, Vol. 10, Issue 3, No 1, pp 231-236, May 2013

[4] Xiang, M.Y. Chong and H. L. Zhu, "Design of Multiple-level Tree classifiers for intrusion detection system", IEEE conference on Cybernetics and Intelligent system, 2004

[5] Peddabachigiri S., A. Abraham., C. Grosan and J. Thomas, "Modeling of Intrusion Detection System Using Hybrid intelligent systems" , Journals of network computer application, 2007

[6] Mrutyunjaya Panda and Manas Ranjan Patra, "A Comparative Study Of Data Mining Algorithms For Network Intrusion Detection", First International Conference on Emerging Trends in Engineering and Technology, pp 504-507, IEEE, 2008

[7] M.Govindarajan and Rlvl.Chandrasekaran, "Intrusion Detection Using k-Nearest Neighbor" pp 13-20, ICAC, IEEE, 2009

[8] Mohammadreza Ektefa, Sara Memar, Fatimah Sidi and Lilly Suriani Affendey, "Intrusion Detection Using Data Mining Techniques", pp 200-203, IEEE, 2010

[9] Song Naiping and Zhou Genyuan, "A study on Intrusion Detection Based on Data Mining", International Conference of Information Science and Management Engineering , Pp 135- 138, IEEE,2010

[10] T. Velmurugan and T. Santhanam, "Computational Complexity between K-Means and K-Medoids Clustering Algorithms for Normal and Uniform Distributions of Data Points", Journal of Computer Science 6 (3): 363-368, 2010

[11] P Amudha and H Abdul Rauf, "Performance Analysis of Data Mining Approaches in Intrusion Detection", IEEE, 2011

[12] R.China Appala Naidu and P.S.Avadhani, "A Comparison of Data Mining Techniques for Intrusion Detection", International Conference on Advanced Communication Control and Computing Technologies (ICACCCT), pp-41-44, IEEE, 2012

[13] Roshan Chitrakar and Huang Chuanhe, "Anomaly based Intrusion Detection using Hybrid Learning Approach of combining k-Medoids Clustering and Naïve Bayes Classification", IEEE,2012

[14] Roshan Chitrakar and Huang Chuanhe, "Anomaly Detection using Support Vector Machine Classification with k-Medoids Clustering", IEEE, 2012

[15] Sumaiya Thaseen and Ch. Aswani Kumar, "An Analysis of Supervised Tree Based Classifiers for Intrusion Detection System", International Conference on Pattern Recognition, Informatics and Mobile Engineering (PRIME), IEEE, February 21-22 2013

[16] David Ndumiyana, Richard Gotora and Hilton Chikwiriro, "Data Mining Techniques in Intrusion Detection: Tightening Network Security", International Journal of Engineering Research & Technology (IJERT), Vol. 2 Issue 5, May – 2013

[17] Muhammad K. Asif, Talha A. Khan,Talha A. Taj, Umar Naeem and Sufyan Yakoob, " Network Intrusion Detection and its Strategic Importance", Business Engineering and Industrial Applications Colloquium(BEIAC), IEEE, 2013

[18] Kapil Wankhade, Sadia Patka and Ravindra Thools, "An Efficient Approach for Intrusion Detection Using Data Mining Methods", IEEE 2013

[19] Fatin Norsyafawati Mohd Sabri, Norita Md Norwawi and Kamaruzzaman Seman, "Hybrid of Rough Set Theory and Artificial Immune Recognition System as a Solution to Decrease False Alarm Rate in Intrusion Detection System", IEEE 2011

[20] Vaishali B Kosamkar and Sangita S Chaudhari, "Data Mining Algorithms for Intrusion Detection System: An Overview", International Conference in Recent Trends in Information Technology and Computer Science (ICRTITCS), 2012

[21] Hind Tribak , Blanca L. Delgado-Marquez, P.Rojas, O.Valenzuela, H. Pomares and I. Rojas, " Statistical Analysis of Different Artificial Intelligent Techniques applied to Intrusion Detection System", IEEE, 2012

[22] S. Revathi and A. Malathi, "Data Preprocessing for Intrusion Detection System using Swarm Intelligence Techniques", International Journal of Computer Applications , Volume 75– No.6, August 2013

[23] Iwan Syarif, Adam Pruge Bennett and Gary Wills, "Unsupervised clustering approach for network anomaly detection", IEEE.

2014 IEEE International Conference on
Autonomous Robot Systems and Competitions (ICARSC)
May 14-15, Espinho, Portugal

Evaluation Of Signal Smoothing Algorithms for Stability of a Quadrotor MAV

Asheesh Ranjan, Pranav Jetley
Department of Computer Engineering
Delhi Technological University
Delhi, India

***Abstract*—The control of mobile robots involves the use of wireless signals. In this paper, we compare and evaluate the real time performance of multiple smoothing algorithms for a pulse width modulated digital signal on an Arduino Uno microcontroller board. The input signal is generated in a real world setting via a manually controlled RC transmitter-receiver setup. The signal smoothing is performed on the input signal and performance is evaluated in terms of standard deviation and signal to noise ratio. We also focus on the needs of a critical real-time system, a remotely piloted UAV, and perform a secondary analysis in terms of mean run time and signal accuracy to find the most suitable algorithm.**

I. Introduction

There has been a spurt of research and developmental activities in the field of unmanned aerial vehicles (UAVs) in the last decade. They are used to perform various types of missions ranging from fire fighting, surveillance and reconnaissance, locating survivors in natural disasters, to offensive military application. They offer the advantage of being un-piloted and can therefore be used for more dangerous missions, where the risks for piloted aircraft are high. Recent advancements in the field have focused on the development of smaller UAVs, which are called micro aerial vehicles or MAVs [1]. These are usually smaller in size which make them applicable for use in areas where larger UAVs cannot be used, such as being able to fly indoors. One type of MAV is a quadrotor, which contains four motors mounted at the ends of two beams, usually arranged in a plus or X configuration. The simple design and lack of any moving parts makes the quadrotor a popular choice for various missions.

For remote operation, MAVs are operated by a radio controlled transmitter receiver system. The control signals are usually transmitted across a fixed frequency range. As they travel through the channel, these signals are subject to interference by noise, which can corrupt the control signal and compromise the functionality of the MAV. A variation in the control signal to the MAV will cause abrupt changes in motor RPM, leading to varying thrust values. This variation in the thrust generated by the propellers can significantly change the orientation of the craft, in turn compromising flight stability. Thus, there is a need to filter out the noisy part of the signal to ensure that the MAV responds reliably to different control inputs.

This paper addresses the removal of noise in a control signal for a MAV using various time domain filters. Our test bed is a quadrotor MAV, which is controlled by an onboard Arduino Uno microcontroller board receiving control inputs from a 2.4 Ghz Avionic RCB7X transmitter receiver system. We will be implementing smoothing algorithms on the pulse width modulated (PWM) digital signals sent by the receiver to the Arduino Uno. The processed output will be sent to the four motors.

This paper is organized as follows; section II gives a background about different types of noise and their generation. It will also give a backdrop about different signal smoothing algorithms used for filtering noise from a signal, and the methods applied to compare the performance of different algorithms. Section III describes the hardware platform that has been used for conducting the experiment. Section IV describes the experiment in detail. The methods used for gathering data, factors behind the use of different algorithms, the statistical and application specific measures used to compare them are described. Section V focuses on the results and evaluation of our experiment. Finally, conclusion and future work are presented in section VI.

II. Background

A. Types of Noise and Noise Generation

The error introduced into a signal during transmission from the sender to the receiver is called noise. The presence of noise in a signal impairs the signal to noise ratio, which is a measure of signal quality, and causes fluctuation in the signal value. This in turn, makes the signal inaccurate and unreliable and may even lead to a loss of data.

There are various sources of noise in our experimental setup. Noise can either be generated at the transmitter side circuit, in the transmission channel, or at the receiver side circuit. The types of noise being generated can be categorized depending on the cause. Random fluctuations in the current flowing in the circuits of the transmitter and receiver may be a source of shot noise. Flicker noise may occur in electronic devices and can show up as irregularities in the conducting wires causing fluctuation in the voltage and current values. Another type of noise is transient noise, which occurs during signal transmission and is caused by interference in the transmission channel. These are short pulses followed by decaying low frequency oscillations. Another source of noise is the thermal agitation of electrons inside a conductor, called thermal noise. Lastly, a major source of noise in our experiment is quantization noise. The analog signal sent by the transmitter is converted to a (PWM) signal before being sent

978-1-4799-4254-1/14/$31.00 ©2014 IEEE

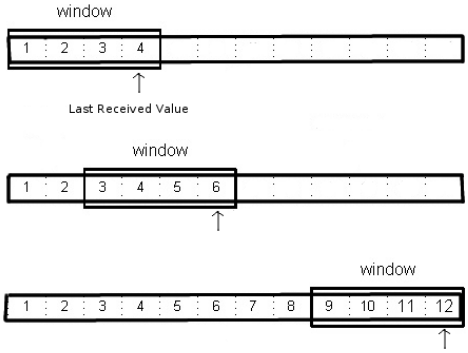


Fig. 1. Input channel with moving window

to the motors. The rounding off of errors that occur during this process of signal conversion is known as quantization error, which presents itself as noise in the original signal.

B. Signal Smoothing and Related Algorithms

Signal smoothing is a process that attempts to capture essential information while leaving out the noise in the signal, by interpolating the raw signal to estimate the original signal. As our system updates control signals in real time, we have considered filters in the time domain for signal smoothing [2] i.e. the simple moving average (SMA), cumulative moving average (CMA), exponential moving average (EMA), the Savitzky-Golay (SG) filter, and the Ramer-Douglas-Peucker (RDP) algorithm. These time domain filters usually have a finite window width and the window moves along the data set, as shown in the figure below. For our application, the most recently received signal value is the first value in the moving window. The data points in the smoothing window are assigned different weights and the actual value is interpolated. The weighting function is the primary factor differentiating the algorithms.

C. Comparison Methodology

The comparison methodology applied for the various algorithms is two fold: a statistical comparison and an application specific comparison. In the statistical comparison we compute the standard deviation (SD) of the data sets obtained for different algorithms. The SD is an indicator of the variation of a data set from its mean value. A lower SD in the control signal will result in more stable flight. We will also be comparing the ratios of the signal to noise ratio (SNR) of the smooth and noisy signal. The SNR ratio will provide us with a measure of the noise present in the signal.

For a signal consisting of a constant signal part a varying noisy part, the SNR is given by

$$SNR = \frac{S^2}{\sigma_n^2} \tag{1}$$

where S is the constant signal value and n^2 is the variance of the noise. Thus, the SNR ratio of the smooth and noise signal is given by

Fig. 2. Quadrotor MAV Test Platform

$$SNR\,Ratio = \frac{\frac{S^2}{\sigma_{n1}^2}}{\frac{S^2}{\sigma_{n2}^2}} \tag{2}$$

where σ_{n1}^2 is the variance of the noise of the original signal and σ_{n2}^2 is the variance of the noise of the noisy signal. As the input signal is the same in both the data sets, the useful portion of the signal denoted by S can be assumed to be equal. Thus, the above equation simplifies to

$$SNR\,Ratio = \frac{\sigma_{n2}^2}{\sigma_{n1}^2} \tag{3}$$

While the previous analyses provide an indication of the signal parameters in terms of statistics, these parameters alone do not provide sufficient information regarding real world performance of the algorithms. One of the constraints is that the processing power on the Arduino Uno is constrained [3], with an Atmega 328 running at clock frequency of 16 MHz. To run the motors stably, the Arduino Uno needs to update the control signal at a frequency of 80 Mhz. This means that the run time of each algorithm needs to be within reasonable limits. Therefore, we will compute the mean run time (MRT) for each algorithm and the ones having the least run times are the ones which are most suitable for a control system application.

As the smoothing algorithm attempts to interpolate the control signal from the received signal, we will also analyse the accuracy of the smoothed output by calculating the difference of the integrals (DOI) of the raw and smooth signals. This term will indicate the extent of signal deviation from the original signal after smoothing. In the real world, this translates to the system reacting as expected to the control signal.

III. HARDWARE PLATFORM

The MAV used in the experiment consists of two 45cm x 2.5cm beams arranged in a X configuration. The beams are made of balsa wood and have the electronics unit is housed at the center of the beams. Four 8x4 APC E propellers are mounted on Turnigy D2830-11 1000kV brushless outrunner motors, which are powered by a Turnigy 3000 mAh 3S1P lithium polymer battery and are

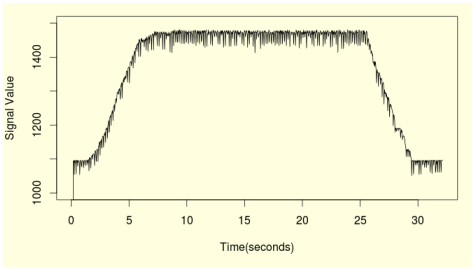


Fig. 3. Sample input *(motor activates at a signal value 1150)*

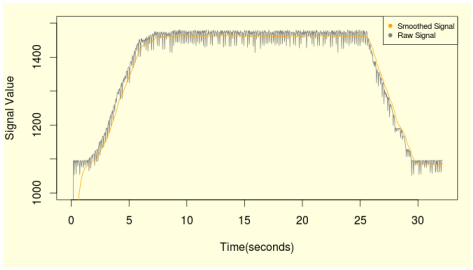


Fig. 4. Plot representing input signal and smoothed output for EMA 0.10

connected via a Hobbyking UBEC 30A electronic speed control. The electronic speed controllers are controlled by the Arduino microcontroller board. Our experiment makes use of the Avionic RCB7X 2.4Ghz 7-channel transmitter receiver system, which incorporates frequency hopping spread spectrum (FHSS) technology. FHSS transmits radio signals by switching between different frequency channels, using a particular sequence known to both the transmitter and receiver. This makes it highly resistant to narrow-band interference. The receiver converts the control signal into a digital PWM signal and sends it to the Arduino board.

The Arduino board being used here is an Arduino Uno Revision 3. This board forms the main controller of our quadrotor platform. The board is programmed using the official Arduino IDE, which comes packaged with platform specific libraries and configuration files. The Arduino Uno board has limited processing and memory resources. It consists of Atmega 328 microcontroller operating at 16 MHz, 1KB EEPROM, 2KB SRAM, and 32KB flash memory. It provides 14 digital and 6 analog general-purpose input-output pins, out of which 6 support PWM. We will be connecting the PWM pins to the receiver as the receiver generates PWM signals. It also has a Atmega 16U2 microcontroller which allows the board to communicate with a computer via a serial connection as well as allows the main microcontroller to be programmed.

IV. Experiment

We have considered a situation where the quadrotor starts from rest and accelerates to a throttle level of approximately 50%, which is sufficient for the quadrotor to hover. The quadrotor stays in the hover position for 15 seconds, and finally decelerates to idle throttle in the last 5 seconds. The throttle is controlled manually to closely resemble real world operation. The algorithms are implemented in C++. The Arduino board processes the digital PWM signals being sent by the receiver using the implemented smoothing algorithms. We monitor the system in real-time using a serial connection and record the input signal value, both raw and smoothed, and the times at which the smoothing function is called and returns. This process is repeated multiple times per algorithm to ensure a significant sized data set.

We have considered filters in the time domain for signal smoothing i.e. the simple moving average (SMA), cumulative moving average (CMA), exponential moving average (EMA), the Savitzky-Golay filter, and the Ramer-Douglas-Peucker (RDP) algorithm. The SMA is the mean of the set of data points distributed uniformly on either side of a central value. The computed mean is then set as the control value in place of the received value at that time instant. The number of data points, also known as the window size can be varied.

The CMA takes the mean of all the data that has arrived till that particular instant. It is given by the formula:

$$C_{k+1} = \frac{x_{k+1} + xC_k}{k+1} \tag{4}$$

$$C_k = \frac{\sum_{i=1}^{k} x_i}{k} \tag{5}$$

where C_k is the cumulative average of k data points and x_k is the value of k^{th} data point. C_0 is taken as zero. Similar to SMA, we again compute the CMA and set it as the control value, in place of the received value at that time instant.

The EMA is a type of weighted moving average with an exponential weighting function [4]. The weighting function uses a continuously decreasing exponential function. A constant factor, alpha, represents the degree of decrease in successive weights, and lies between 0 and 1. A smaller value of alpha takes more of the previous readings into account for calculating the current signal value. The EMA for a series S can be calculated as:

$$S_1 = C_1 \quad \text{for } t = 1$$
$$S_t = \alpha Y_t + (1-\alpha) S_{t-1} \quad \text{for } t > 1 \tag{6}$$

where Y_t is the data value at a time instant t. S_t is the value of the EMA at any time instant t.

The SG filter is a digital filter, which works on the method of fitting least squares, of a polynomial of a given order, to data points in a moving window. The polynomial is evaluated at the center of the moving window, and that value is the filtered value. The window is shifted over the entire data set and the central value of the window is calculated in each iteration of the algorithm. The degree of the polynomial used for least squares fitting and the moving window size can be varied and tested to obtain the most appropriate combination of parameters. The convolution coefficients for smoothing the signal can be obtained from the original paper of Savitzky and Golay [5].

The final algorithm implemented was RDP algorithm

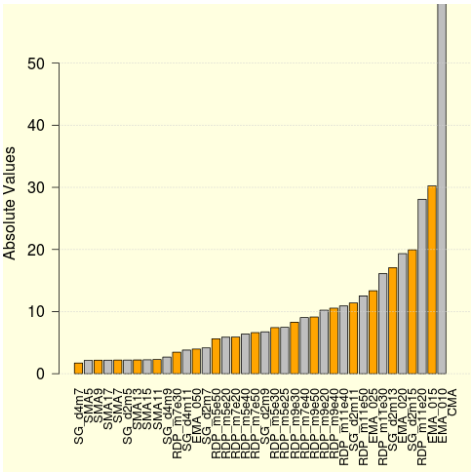


Fig. 5. SNR Ratios

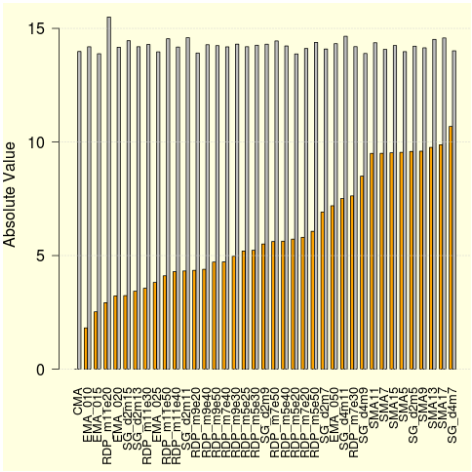


Fig. 6. Standard deviation for input and smoothed signal *(orange represents the smoothed signal and grey represents the original signal)*

[6]. Given a set of curves, it attempts to reduce the number of points required to best approximate the curve. The algorithm uses the perpendicular distance of points from the curve to estimate which points need to be ignored. The ones at a distance greater than a fixed value (ϵ) are considered significant and remain part of the final smoothed output. It can be applied for smoothing noisy signals by filtering the data set to include fewer points.

Now that the algorithms were implemented, we executed the program on the Arduino board and observed the performance in real time. An open serial connection between the Arduino board and the workstation was used to collect data for a period of 30 seconds using the input signal as described above. Each record of the data set collected consisted of four parameters: time of call of smoothing function, raw value of input signal, time of return of smoothing function, and the smoothed signal value. To compare the performance of the different algorithms, we first performed a statistical comparison. The firs part of this analysis involved computing the SD of the raw and smooth value of the data set. This was done using the SD function present in the stats package of R. Furthermore, the SNR ratio was also computed for the raw and smooth signal. The SNR values were computed using the formula given in section II. We used the variance function from the stats package in R for this analysis.

The next phase of analysis involved an application specific evaluation of the different algorithms on the quadrotor platform. For this, the MRT for each algorithm was computed. The run time was estimated by calculating the difference of the time of call of smoothing function and the time of return of smoothing function. The run time was then averaged over the entire data set. The second part of this analysis consisted of calculating the DOI of the smooth and noisy signal. To compute the integrals, we implemented an integral function.

V. Results and Evaluation

As mentioned above, the data was collected in real-time from the MAV using a serial connection. The Arduino was refreshing control signal to the motors at a variable frequency depending on the algorithm being used. The data obtained was then preprocessed to remove any instances of invalid characters. An example of the input signal versus the smoothed output of EMA for alpha 0.10 is shown in figure 8.

A. Statistical Evaluation

First, we calculated the SD of each data set in the time interval when the signal was expected to be stable i.e. from 10 to 25 seconds. It is expected that the SD of the smooth data set will be lower than that of the raw data set and that lower SDs will result in a more stable control signal. This will result in a uniform thrust being generated leading to a more stable flight.

The SDs showed a visibly improved output signal being generated in all cases. As can be seen in figure 6, certain algorithms performed markedly better than the rest, with CMA resulting in a SD of 0. EMA with alpha 0.10 and 0.15 performed well with low SDs. Thus, we observed that a higher dependence on the previous values leads to a smoother signal. The SG filter was most efficient for a quadratic fitting and a larger window size. For the quartic fitting, a larger window size gave a lower SD. All variations of window size and epsilon for RDP resulted in higher SDs than EMA, but were approximately close to each other. The worst performing algorithms were variations of SMA, resulting in consistently large SDs.

Next, we computed the ratio of SNRs for the smoothed and raw signal, which can be seen in figure 5. It is expected a higher SNR ratio of the control signal will indicate a signal with lower fraction of noise present in it. The variation in SNR was found to be significant ranging from 1.7 to 6. The CMA case in particular had a zero variance in the smoothed signal resulting in a extremely large SNR ratio. The highest SNR ratio was observed for EMA with alpha 0.10 followed by EMA with alpha 0.15. A large majority of variations of RDP were found to have SNR ratios distributed in a small range from 5.6 to 10.9. However, a window size of 11 and

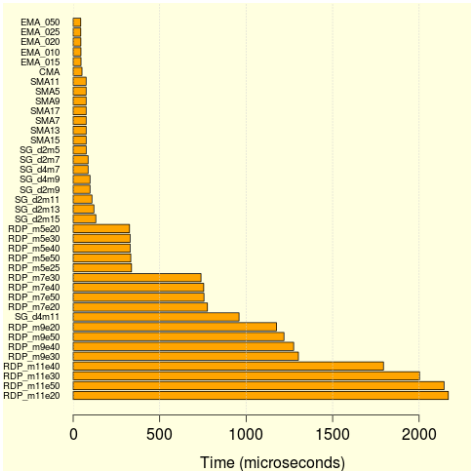


Fig. 7. Mean run time

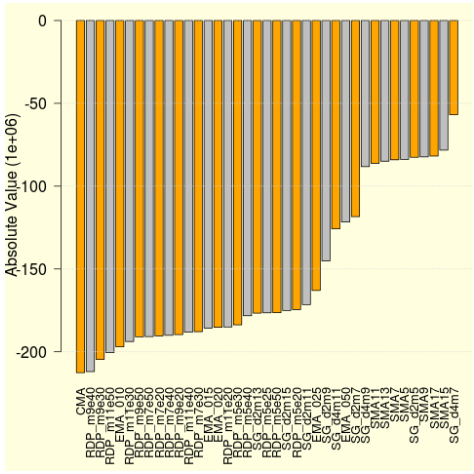


Fig. 8. Difference of integrals

an epsilon value of 20 gave the highest SNR ratio of 28. All variations of SMA resulted in very low SNR ratios. SG filter showed variable SNR ratios depending on the window size and degree of fitting. However, the highest SNR ratio was observed for a quadratic fitting and a window size of 15.

B. Application Oriented Evaluation

Our second phase of evaluation consisted of comparisons based on the requirements for our quadrotor MAV. The primary concerns as mentioned above were the accuracy of the signal being generated after smoothing as well as the time taken for the execution of the algorithm.

Our expectation from the DOI measure was that a value close to zero will indicate a more accurate output. The accuracy of the signal will translate to a more responsive MAV. On observing figure 8, we see that CMA has the lowest DOI, thus indicating that it was the least accurate. This makes it unfeasible for use in control systems. RDP for large window sizes also resulted in significantly low values. The EMA with alpha 0.10 performed poorly whereas EMA with alpha 0.15 performed reasonably well with a DOI of -185. Different variations of SG filter had a varying performance ranging from a DOI of -176 for quadratic fitting with window size 13 to -57 for a quartic fitting with window size 7. The SMA performed extremely well on this metric having the least DOI with no major difference across different window sizes.

Next, we compared the mean run time as the severely constrained requirement of the Arduino processor required smoothing to be performed efficiently. This will allow the Arduino Uno to update the control signal at a higher frequency. The mean run time for EMA was the lowest followed by CMA. SMA took approximately double the time required for EMA. SG filter performed reasonably well with MRT varying from 75 to 130 for the smallest degree and smallest window size. As the window size and the degree of fitting increased, there was an increase in the MRT. The worst performing algorithm was RDP as it involved multiple recursions, which made it unfeasible for use in our application.

Finally, we intended to combine the above four measures into an aggregate measure based on which we would estimate the best performing algorithm for our specific application. To do this, we normalized each data set on a zero-one scale with higher values indicating a better performance. A weighting function that would assign a priority to each measure was decided. The function assigned the following weights: 4-MRT, 3-DOI, 2-SNR ratio, and 1-SD. We assigned MRT the maximum priority as it would effect the responsiveness of the entire system to any input being provided. We calculated the aggregate measure or score by adding the four weighted values. The results of this analysis were as shown in table I.

TABLE I. FINAL SCORES OF ALGORITHMS

Rank	Algorithm	Score
1	EMA_010	9.707
2	EMA_015	8.371
3	EMA_020	7.916
4	SG_d2m15	7.577
5	SG_d2m13	7.509

VI. CONCLUSION AND FUTURE WORK

In this paper, multiple signal smoothing algorithms were implemented and tested on a quadrotor MAV platform. The control input was processed in real time using various signal processing algorithms and their performance evaluated. The evaluation was done on the basis of SD and SNR, followed by DOI and MRT. Exponential moving average filters were found to be the best performing algorithms with alpha parameters of 0.10 and 0.15. EMA was found to perform well on all metrics and came across as the best choice for application in a MAV.

In the future, we will explore other filters and experiment with different scenarios and larger data sets. Efficient signal smoothing on a low power platform such as Arduino Uno has the potential to be applied in many areas where control systems may be involved. Furthermore, these results are applicable to other areas utilising digital signal processing such as audio, SONAR and RADAR systems. Specifically, for MAVs, we

will incorporate an inertial measurement unit and evaluate the performance of the smoothing algorithms for varying flight patterns. Indoor flying is one of the core application areas of MAVs and stable flight with minimal vibrations is critical in closed and constrained environments.

REFERENCES

[1] C. Galiski and R. bikowskib *"Some problems of micro air vehicles development,"* Bulletin of the Polish Academy of Sciences, Technical Sciences, Vol. 55, No. 1, 2007

[2] Seng Hansun, *"A new approach of moving average method in time series analysis,"* New Media Studies (CoNMedia), 2013 Conference on , vol., no., pp.1,4, 27-28 Nov. 2013

[3] Koswatta, R.; Karmakar, N.C., *"Moving average filtering technique for signal processing in digital section of UWB chipless RFID reader,"* Microwave Conference Proceedings (APMC), 2010 Asia-Pacific , vol., no., pp.1304,1307, 7-10 Dec. 2010

[4] Wilson, A., *"Event Triggered Analog Data Acquisition Using the Exponential Moving Average,"* Sensors Journal, IEEE , vol.PP, no.99, pp.1,1

[5] Abraham. Savitzky and M. J. E. Golay; *Smoothing and Differentiation of Data by Simplified Least Squares Procedures.* Analytical Chemistry 1964 36 (8), 1627-1639

[6] Douglas, D. H. and Peucker, T. K. (2011) *Algorithms for the Reduction of the Number of Points Required to Represent a Digitized Line or its Caricature*, in Classics in Cartography: Reflections on Influential Articles from Cartographica (ed M. Dodge), John Wiley & Sons, Ltd, Chichester, UK.

[7] Chen, H.C.; Chen, S.W., *"A moving average based filtering system with its application to real-time QRS detection,"* Computers in Cardiology, 2003 , vol., no., pp.585,588, 21-24 Sept. 2003

[8] Ziluan Liu; Yuehui Jin; Yidong Cui; Qiyao Wang, *"Design and implementation of a line simplification algorithm for network measurement system,"* Broadband Network and Multimedia Technology (IC-BNMT), 2011 4th IEEE International Conference on , vol., no., pp.412,416, 28-30 Oct. 2011

Article

Field Emission with Ultra-Low Turn-On Voltage from Metal Decorated Carbon Nanotubes

Srividya Sridhar, Chandrasekhar Tiwary, Soumya Vinod, Jose Jaime Taha-Tijerina, Srividvatha Sridhar, Kaushik Kalaga, Benjamin Sirota, Amelia H. C. Hart, Sehmus Ozden, Ravindra Kumar Sinha, Harsh ,, Robert Vajtai, Wonbong Choi, Krisztián Kordás, and Pulickel M Ajayan

ACS Nano, **Just Accepted Manuscript** • DOI: 10.1021/nn500921s • Publication Date (Web): 15 Jul 2014

Field Emission with Ultra-Low Turn-On Voltage from Metal Decorated Carbon Nanotubes

Srividya Sridhar, [1] *Chandrasekhar Tiwary,* [2,6] *Soumya Vinod,* [2] *Jose Jaime Taha-Tijerina,* [2] *Srividvatha Sridhar,* [3] *Kaushik Kalaga,* [2] *Benjamin Sirota,* [5] *Amelia H. C. Hart,* [2] *Sehmus Ozden,* [2] *Ravindra Kumar Sinha,* [1] *Harsh,* [4] *Robert Vajtai,* [2] *Wongbong Choi,* [5] *Krisztián Kordás* [7] *and Pulickel M. Ajayan*[2*]

[1]Delhi Technological University (Formerly Delhi College of Engineering), Department of Applied Physics, Bawana Road, Delhi 110042, India.
[2]Department of Materials Science and Nano Engineering, Rice University, Houston, Texas 77005, United States.
[3]Department of Chemical and Biomolecular Engineering, Rice University, Houston, Texas 77005, United States.
[4]Department of Physics, Jamia Millia Islamia, New Delhi 110025, India.
[5]Department of Material Science and Engineering, University of North Texas, Denton, Texas 76203, United States.
[6]Materials Engineering, Indian Institute of Science, Bangalore, Karnataka-560012, India.
[7]Microelectronics and Materials Physics Laboratories, Department of Electrical Engineering, University of Oulu, P.O. Box 4500, FI-90014 Oulu, Finland.

*Corresponding Author: Dr. P. M. Ajayan; Email Address: ajayan@rice.edu

ABSTRACT

A simple and scalable method of decorating 3D-carbon nanotube (CNT) forest with metal particles has been developed. The results observed in aluminum (Al) decorated CNTs and copper (Cu) decorated CNTs on silicon (Si) and Inconel are compared with undecorated samples. A significant improvement in the field emission characteristics of the cold cathode was observed with ultra-low turn on voltage (E_{to}~0.1 V/μm) due to decoration of CNTs with metal nanoparticles. Contact resistance between the CNTs and the substrate has also been reduced to a large extent because of which we were able to get stable emission for longer duration without any current degradation, thereby promising a possibility of their use in vacuum microelectronic devices.

Key words: Field emission, metal decoration, screening effect, work function, edge effect

Due to the high aspect ratio, high mechanical strength and high thermal conductivity, CNTs are considered to be an ideal electron emitting sources in field emission displays. [1–8] Aligned and well-separated CNT cathode morphology is important for many potential applications, where a high electric-field is needed such as in field emission devices. The growth in these cases is usually achieved by the means of chemical vapor deposition (CVD). It is important to mention, however, the highly dense growth of CNTs by thermal CVD method may compromise the field emission properties, due to a field-screening effect caused by the proximity of neighboring tubes. Also some critical issues such as adhesion of CNTs with the substrate, reliability, and stability have to be addressed while using CNTs as electron emitters in field emission applications. To improve the field emission performance from CNT emitters, different methods such as doping of CNTs with nitrogen, [9,10] surface coatings with low work function materials such as LaB_6 or Ha for low turn on field [11,12] and decorations with ZnO and Ru [13–15] have been investigated. Previous studies have also shown that tuning the structure of the CNTs such as the radius, [16] geometry, [17,18] structural change by oxygen plasma treatment [19] and improving their density [20] could improve the field emission property. Also by decorating the CNT walls with organic functional groups [21–23] or by inorganic semiconductor [24–31] the field emission properties of CNTs could be enhanced. Chi *et al.* [32] reported a threshold field of 0.9 V/μm at 1 mA/cm^2 which correlated to a turn on voltage of 0.6 V/μm at 10 μA/cm^2 by growing CNTs on a mesh electrode. Also, Zuo *et al.* [33] demonstrated low turn on field of 0.63 V/μm at 10 μA/cm^2 by decorating the surface of the CNT with titanium nanoparticles. Pandey *et al.* [34] arrived at a threshold field of 0.8 V/μm in strontium titanate coated CNTs. On the other hand, Liu *et al.* [15] decorated CNTs with Ru nanoparticle and the turn on field reduced from 2.5 to 1.3 V/μm after Ru decoration. In the recent report of Zannin *et al.* [35] hybrid diamond-like carbon

and CNT composite structures showed threshold field of 2 V/μm. Thus, to the best of our knowledge, such low turn on field of 0.1 V/μm achieved in this work has not been reported so far.

Electrophoretic deposition (EPD) technique is a common method used for decorating the CNTs with the desired metal particle to improve the field emission properties of the cold cathode. Fan *et al.* [36] used CNT/Ni composite and Chen *et al.* [37] fabricated CNT/Cu composite to enhance the field emission properties of the CNTs. But during the fabrication, the CNTs in the composites would be easily covered by the consequently deposited nanoparticles in the chemical process resulting in the decrease of effective field emitters in the CNT cathode. Another important issue of CNTs getting bundled together is also common in this process. Once they are bundled, without the use of aggressive treatment like ultra-sonication or chemical functionalization, it is extremely difficult to achieve a stable dispersion or homogeneous mixture. These measures are known to damage the walls of the CNTs and degrade their electrical properties to a larger extent. Moreover, this process cannot alter the distance between the CNTs. Although it is important to lower the turn on and threshold voltage for better field emission results, reducing the screening effect is crucial for CNTs to be used in field emission application. It is well known that if the separation between the tubes, *"d"*, is much less than the tube height, *"h"*, in an array, the electrostatic shielding between the CNTs can drastically affect the field emission performance of the cold cathode. Numerical simulations show that if the distance between the CNTs is equal to or more than twice the height of the CNTs, then the screening effect can be considerably reduced which in turn could improve the field emission property of the cathode. [38–40] Electrostatic shielding becomes a significant hurdle for future device applications because of the exponential dependence of the emission current on the electric field

in accordance with the Fowler-Nordheim theory. Work has also been done on thermally evaporated nanoparticles on nanotubes. [41–43] The results in these work indicate that the morphology of the metal particle or the film on the CNTs is dependent on the intrinsic properties of the metal themselves than on the process parameters. On the other hand, the work of Muratore *et al*. [44] reported the effect of particle growth temperature and time on the decorated metal nanoparticles' density and morphology. However, the sputtered metal nanoparticles were attached to the surface of the CNTs without penetrating in between the CNTs to change the inter tubular spacing of those.

The current paper proposes a simple and scalable procedure for decorating the vertically aligned CNTs forest with metal nanoparticles for improved field emission properties. This process enabled the CNTs to bundle towards a metal particle forming a pattern due to which the CNT bundles are separated by almost 3 times the height of the bundle reducing the screening effect considerably. Further, by the deposition of metal nanoparticles on the surface of the nanotubes, emitting centers are obtained that ensure highly conductive paths for the electrons from the nanotubes towards the vacuum helping to by-pass the amorphous carbon impurities, known as one of the major hurdles in CNT based emitters.

RESULTS AND DISCUSSION

Figure 1a shows the scanning electron microscope (SEM) images of the multi-walled carbon nanotubes (MWCNTs) grown on Si with Al decoration on them. Before the metal decoration, the CNTs were vertically aligned *via* self-supporting mechanism as described in our previous work, [45] due to extremely high density of CNTs. Interestingly, after metal evaporation, the morphology of the CNTs changed completely. Groups of CNTs became linked together by

their tips while attaching to the Al particles forming microscopic patterns in the nanotube forest. Accordingly, the CNTs are no longer individual strands for field emission; instead they are individual bundles. Choi. *et al.* [20] also observed similar morphology by H_2 plasma treatment. The surface morphology of vertically aligned MWCNTs changed from flat surface to sharp conical stacks of CNTs due to the post plasma treatment. They claimed that there was improvement in the field emission properties of the CNTs due to the formation of such stacks.

Figure S1a and 1b show the distribution histogram of the number of CNTs with their height and the number of CNT bundles with the inter-bundle distance respectively. From the histogram, it is very clear that most of the CNT bundles are of around 2 μm in their height with inter bundle distances varying mostly between 3 and 5 μm. It is to be noted that bundled CNTs have the advantage of realizing the ideal ratio of inter tube distance to the height for achieving maximum field emission by reducing charge screening effects in adjacent CNTs in the forest. [46] According to Suh *et al.* [47] the field emission would be optimum when the tube height is similar to the inter-tube distance, while the results of Ren *et al*. [48] suggest 3-fold inter-tube distance in reference to CNT height to achieve maximum emitter efficiency. In our case, since most of the CNT bundles are separated by almost 3 times the height of the bundle and the bundle are emitting as a whole, the screening effect is reduced considerably and the emitter efficiency has increased substantially. Figure 1b corresponds to the SEM images of Al decorated CNTs grown on Inconel. SEM images clearly show that the CNTs on Inconel are also no longer a uniform film after the metal deposition, however the surface texture of the forests are not the same as for the other CNTs grown on the other substrate material. The nanotubes in this case seems to have formed bundles which has in turn is expected to reduce the screening effect. A schematic of the morphology of CNTs grown on Si and Inconel substrate is shown in Figure 1c. Although the

mechanisms responsible for the differing surface textures after evaporating the metals on the different films are not clear, a number of different reasons associated with the wetting behavior of the CNT forests with the molten metal (and/or with the condensing metal from the vapor phase) might explain our observations. Figure 2a shows the transmission electron microscope images (TEM) and high resolution transmission electron microscope images (HRTEM) of CNTs grown on Si substrate with Al decoration on them. The low magnification image resembles a similar morphology as seen in SEM. The TEM observation of the CNTs revealed the formation of MWCNTs consisting of 4-6 graphite layers. It is clear that individual CNTs form bundles with adjacent nanotubes and share a metal tip. Selected area diffraction (SAD) of the particle confirms these particles to be FCC plane (111) of Al. The high magnification image at the point of joint in a bundle confirms presence of metal particles and several CNTs are found to be tangled together. The high magnification image also reveals the nanoparticles embedded in the amorphous carbon layer are in reasonably close contact with the nanotubes, while the other side of the particles is unraveled. The increase in wall diameter due to bundle formation is further verified with Raman measurements. The Al deposition on Inconel shows very similar morphology as shown in Figure 2b. The large area SAD shows the presence of Al planes. The high magnification image shows an increase in diameter with a narrow distribution because of CNTs joining together to form a bundle. The size of the metal nanoparticles was found to be 5-10 nm in diameter.

Raman spectra of the pristine CVD grown CNTs with and without metal deposition are shown in Figure 3a. Raman spectroscopy is the inelastic scattering of light usually associated with absorption or emission of phonons and is rich in information about the structure and chemical bonding of CNTs. Defects and sp^3 hybridized carbon atom give rise to D- band (1340

cm-1) and height of this band is inversely related to the quality of the nanotube (*i.e.* presence of disorder in the graphitic material). The G- band (~1580 cm-1) is associated with the sp^2 hybridized carbon atom in the nanotube wall and is a good measure of the graphitization of the sample. The G' or 2D peak (around 2600 cm-1) arises due to two phonon second order scattering process and indicates long range order in a sample. The ratio of the intensities of the defect or the disorder induced D-band to the symmetry allowed graphitic band or the G-band I_d/I_g characterizes the defect density or degree of disorder in sp^2 hybridized carbon material. [49] The ratio I_{2d}/I_g can be used for identifying the number of walls in the given MWCNT. [50] It is also to be noted that increase in the number of layer leads to a significant decrease in the peak intensity of the 2D peak and this peak becomes hardly distinguishable if the number of layer exceeds 5. [51]

The intensity ratios I_d/I_g and I_{2d}/I_g are given in Table.1. From the results, it is clear that the intensity ratio I_d/I_g of CNTs grown on Si substrate without any metal decoration and with Al decoration was 0.78 and these values are very close to the values reported by Athipalli *et al.* [52] For the CNTS grown on Si substrate with Cu decoration on them this ratio became 0.51. The lower the value of this ratio, lesser is the defect, which in turn means lesser amorphous carbon, and greater is the degree of graphitic crystallinity. Also, the 2D peak for Al decorated CNTs grown on Inconel was hardly distinguishable and hence the ratio I_{2d}/I_g became zero. As predicted by the work of Ferrari [51] the number of walls must have exceeded five and hence this peak became hardly distinguishable. This was also consistent with TEM, where it is clear that the CNTs join together thereby increasing the number of walls. From the above predictions, it is clear that the CNTs grown on Si with Cu decoration on them have higher graphitic crystallinity

with lesser defect. The occurrence of sharp 2D peak on this sample indicates a good long range order in those samples.

Table 1: Values of I_d/I_g and I_{2d}/I_g ratio for different types of CNTs

Type of CNTs	I_d/I_g	I_{2d}/I_g
Si Grown	0.78	0.71
Inconel Grown	0.65	0.41
Si grown with Al decoration	0.78	0.80
Inconel grown with Al decoration	0.67	0.00
Si grown with Cu decoration	0.51	0.58
Inconel grown with Cu decoration	0.66	0.77

In order to characterize the phase deposited on the CNT forest, the samples are analyzed by X-ray diffraction. Figure 3b shows the representative image of the Al deposition and Cu deposition on Si. In the diffraction pattern we observe reflections at $2\Theta=38°$ and 44° which can be assigned to the (111) and (200) planes of FCC phase of Al. The reflection at $2\Theta=43°$ corresponds to the FCC plane (111) of Cu. These diffraction patterns confirm the presence of metallic particles of Al and Cu and absence of their crystalline oxides. The broadened reflections from the metals clearly reveals the small crystallite size of the metallic particles (~5 nm) supporting observations done with TEM.

The field emission characteristics of the CNTs grown on Si with and without any metal decorations are shown in Figure 4a. The turn on field E_{to}, which is the field required to obtain a current density of 1 $\mu A/cm^2$ and the threshold field E_{th}, which is the field required to obtain a current density of 100 $\mu A/cm^2$ are summarized in Table 2. The data clearly reveals that Si grown CNTs with Al decoration on them gave the lowest E_{to} and E_{th}, 0.13 V/μm and 0.14 V/μm

respectively. One of the explanation for our low E_{to} and E_{th} for Si grown CNTs with Al decoration on them is due to the lower work function of Al as predicted by Lee *et al*. [53] Similar field emission characteristics were observed for CNTs grown on Inconel with and without any metal decorations on them. (Figure 4b) The J-E plot was repeatedly measured and it showed good reproducibility. No current saturation was observed over E_{th}.

It is well known that during field emission, the electrons have to cross two barriers. [54] Barrier 1 is the barrier between the substrate and the CNT and the barrier 2 is the barrier between the CNT and the vacuum. Since the Inconel and Al coated Si substrates on which CNTs are directly grown have good electrical conductivity, barrier 1 is expected to be reduced substantially enhancing the overall field emission properties of our structures. The measured low contact resistance values for the CNT-substrate interfaces are in agreement with our assumptions as seen in Figure S3. Barrier 2 on the other hand has also improved. According to Tanaka *et al.* [55] the presence of amorphous carbon can increase the work function and the E_{th} to a larger extent. One of the explanations for the low E_{th} values for metal decorated CNTs can be due to the reduced amount of amorphous carbon. [45] Although the mechanism that would reduce the amorphous carbon content obtained after Cu decoration is not clear, a plausible explanation might be a partial dissolution of carbon by Cu at the process temperatures applied in the course of evaporation. Another, probably more important, effect that helps electron emission is the promoted electron passage from the nanotubes towards the vacuum through the metal nanoparticles, which may reasonably explain why we observe significant reduction of E_{th} for both metals. The metal nanoparticles form highly conductive electrical paths for the electrons through the amorphous carbon layer (Figure 2a) covering the nanotubes thus ensuring emission centers with clean surface towards the vacuum.

Table 2: Values of E_{to} and E_{th} for different type of substrates with and without metal decoration.

Type of CNTs	E_{to}	E_{th}
Si Grown	0.78	2.6
Inconel Grown	1.2	2
Si grown with Al decoration	0.13	0.14
Inconel grown with Al decoration	0.14	0.18
Si grown with Cu decoration	0.19	0.22
Inconel grown with Cu decoration	0.53	0.58

Furthermore, according to Fuji *et al.* [56] the electric field of the bundle is significantly higher at the edge than at the center when compared to the electric field of the flat film which is constant all over the emitter surface. In our samples, due to the metal decoration, the emitter surface is no longer a flat film; instead they are transformed into individual bundles since some of the Al particles went in between the CNTs, thereby increasing the number of edges. The CNTs at the periphery of the bundle formed due to metal decoration acted as a major emission sites. Thus, the excellent field emission property of our emitters can also be attributed to the edge effect.

In addition to the improved field emission I–V characteristics, better emission uniformity is observed from the emission pattern of Al decorated CNTs grown on both Si and Inconel, as shown in the inset of the Figure 4. This is because of the increase in the number of available emitters and reduced screening effect. Figure 4c shows the time trace of current density at the fixed field of 0.15 V/μm for Si grown CNTs and 0.25 V/μm for Inconel grown CNTs, both

decorated with Al nanoparticles. It was found that there was stable emission for more than an hour due to metal decoration on the CNTs.

CONCLUSIONS

In conclusion, vertically aligned CNTs grown on Al coated Si and Inconel substrates were decorated with Al and Cu particles using a simple and scalable process. The synthesized hybrid structures showed enhanced field emission properties with ultra low turn on and threshold voltages of 0.13 V/μm and 0.14 V/μm respectively measured for the Al decorated CNTs grown on Al coated Si. Contact resistance also got reduced substantially in the metal decorated structures which resulted in stable emission for a longer duration without any current degradation. The excellent field emission properties of the metal decorated CNTs can be attributed to the edge effect, reduced screening effect, lower contact resistance, which may pave the road for future devices that require substantially lower bias than the currently existing ones.

METHODS

First the CNTs were grown on Al coated silicon (Si) and Inconel as described elsewhere. [45] The grown CNTs were kept inside the CVD chamber with aluminum (Al) sheet of thickness 150 μm over them and the temperature of the system was gradually increased to 700°C for 10 minutes and cooled gradually. Since the melting point of Al is 660°C, it got melted and the CNTs were decorated with Al particles. As the sheet was kept over the array of CNTs, some of the Al particles penetrate in between the CNTs in the array. This process of annealing was done in the presence of Argon (Ar) at a low flow rate. The schematic procedure for the fabrication of metal

decorated CNTs is shown in Figure 1c. The same procedure was repeated for the deposition of Cu particles using Cu sheet to obtain Cu decorated CNTs.

The surface morphologies were characterized using scanning electron microscopy (SEM, FEI Quanta 400 ESEM FEG) and high resolution transmission electron microscopy (HRTEM JEOL 2100 F TEM). The content of Al and Cu particles is identified using X-ray diffraction (Rigaku D/Max Ultima II Powder XRD with a Cu Kα source) and Raman spectroscopy (Renishaw in *Via* Raman Microscope). The area of the sample used was 1 cm^2. After the deposition with metal particles; the sample was transferred to a vacuum chamber with vacuum better than 2 x 10^{-6} Torr, for field emission measurement. The silicon (Si) and Inconel substrates with CNTs were used as the cathode and indium tin oxide (ITO) coated glass plate as the anode. The cathode and anode mounting stands were machine ground to ensure that they are perfectly parallel. The distance between the cathode and the anode (100 μm) was adjusted using micrometer screw gauge arrangement and a suitable DC voltage (up to 400 V) was supplied using Keithley 2410 high voltage power supply. The electron impedance spectroscopy (EIS) measurements were performed using a two electrode setup with the CNT on the substrate as the working electrode and lithium metal as the counter/reference electrode. In this, 1 M LiPF6 in 1:1 v/v mixture of ethylene carbonate (EC) and dimethyl carbonate (DMC) are used as the electrolyte and glass micro-fiber filter membrane as the separator. The EIS measurements were conducted over 70 kHz to 10 mHz by applying a constant dc bias with sinusoidal signal of 10 mV.

Acknowledgments *The work done at Rice University has been supported by U.S. Department of Defense: U.S. Air Force Office of Scientific Research for the Project MURI: "Synthesis and Characterization of 3-D Carbon Nanotube Solid Networks" Award No. FA9550-12-1-0035.*

C. S. Tiwary would like to thank Indian Institute of Science, Bangalore for their support. J.T.-T. acknowledges the support from CONACYT (213780).

Supporting Information Available *A histogram of distance distribution between the CNT bundles and the height of the bundle is included. This is followed by F-N plot for Al decorated CNTs and Cu decorated CNTs grown on both Si and Inconel substrates. Nyquist plots of EIS spectra collected from electrochemical lithium half cell of raw MWCNTs grown without any metal decoration and with Al decoration on CNTs (grown on Si and Inconel substrates) are shown. Information about SEM, TEM and HRTEM characterization of CNTS grown on Silicon and Inconel with Copper decoration on them is available in this section. Images after field emission on metal decorated CNTs are also available. This material is available free of charge via the Internet at http: //pubs.acs.org*

REFERENCES AND NOTES

(1) Planeix, J. M.; Coustel, N.; Coq, B.; Brotons, V.; Kumbhar, P. S.; Dutartre, R.; Geneste, P.; Bernier, P.; Ajayan, P. M. Application of Carbon Nanotubes as Supports in Heterogeneous Catalysis. *J. Am. Chem. Soc.* **1994**, *116*, 7935–7936.

(2) Fu, Y.; Zhang, L.; Chen, G. Preparation of a Carbon Nanotube-Copper Nanoparticle Hybrid by Chemical Reduction for Use in the Electrochemical Sensing of Carbohydrates. *Carbon* **2012**, *50*, 2563–2570.

(3) Liu, Z.; Lin, X.; Lee, J. Y.; Zhang, W.; Han, M.; Gan, L. M. Preparation and Characterization of Platinum-Based Electrocatalysts on Multiwalled Carbon Nanotubes for Proton Exchange Membrane Fuel Cells. *Langmuir* **2002**, *18*, 4054–4060.

(4) Sun, Y.; Wang, H. H. High-Performance, Flexible Hydrogen Sensors That Use Carbon Nanotubes Decorated with Palladium Nanoparticles. *Adv. Mater.* **2007**, *19*, 2818–2823.

(5) Li, W.; Liang, C.; Zhou, W.; Qiu, J.; Zhou; Sun, G.; Xin, Q. Preparation and Characterization of Multiwalled Carbon Nanotube-Supported Platinum for Cathode Catalysts of Direct Methanol Fuel Cells. *J. Phys. Chem. B* **2003**, *107*, 6292–6299.

(6) Reddy, A.; Ramaprabhu, S. Hydrogen Storage Properties of Nanocrystalline Pt Dispersed Multi-Walled Carbon Nanotubes. *Int. J. Hydrog. Energy* **2007**, *32*, 3998–4004.
(7) Kong, J.; Chapline, M. G.; Dai, H. Functionalized Carbon Nanotubes for Molecular Hydrogen Sensors. *Adv. Mater.* **2001**, *13*, 1384–1386.
(8) Teo, K. B. K.; Chhowalla, M.; Amaratunga, G. A. J.; Milne, W. I.; Pirio, G.; Legagneux, P.; Wyczisk, F.; Pribat, D.; Hasko, D. G. Field Emission from Dense, Sparse, and Patterned Arrays of Carbon Nanofibers. *Appl. Phys. Lett.* **2002**, *80*, 2011.
(9) Stephan, O.; Ajayan, P. M.; Colliex, C.; Redlich, P.; Lambert, J. M.; Bernier, P.; Lefin, P. Doping Graphitic and Carbon Nanotube Structures with Boron and Nitrogen. *Science* **1994**, *266*, 1683–1685.
(10) Golberg, D.; Bando, Y.; Bourgeois, L.; Kurashima, K.; Sato, T. Large-Scale Synthesis and HRTEM Analysis of Single-Walled B- and N-Doped Carbon Nanotube Bundles. *Carbon* **2000**, *38*, 2017–2027.
(11) Zhang, J.; Yang, C.; Wang, Y.; Feng, T.; Yu, W.; Jiang, J.; Wang, X.; Liu, X. Improvement of the Field Emission of Carbon Nanotubes by Hafnium Coating and Annealing. *Nanotechnology* **2006**, *17*, 257–260.
(12) Wei, W.; Jiang, K.; Wei, Y.; Liu, P.; Liu, K.; Zhang, L.; Li, Q.; Fan, S. LaB[sub 6] Tip-Modified Multiwalled Carbon Nanotube as High Quality Field Emission Electron Source. *Appl. Phys. Lett.* **2006**, *89*, 203112.
(13) Min, Y.-S.; Bae, E. J.; Park, J. B.; Kim, U. J.; Park, W.; Song, J.; Hwang, C. S.; Park, N. ZnO Nanoparticle Growth on Single-Walled Carbon Nanotubes by Atomic Layer Deposition and a Consequent Lifetime Elongation of Nanotube Field Emission. *Appl. Phys. Lett.* **2007**, *90*, 263104.
(14) Green, J. M.; Dong, L.; Gutu, T.; Jiao, J.; Conley, J. F.; Ono, Y. ZnO-Nanoparticle-Coated Carbon Nanotubes Demonstrating Enhanced Electron Field-Emission Properties. *J. Appl. Phys.* **2006**, *99*, 094308.
(15) Liu, C.; Kim, K. S.; Baek, J.; Cho, Y.; Han, S.; Kim, S.-W.; Min, N.-K.; Choi, Y.; Kim, J.-U.; Lee, C. J. Improved Field Emission Properties of Double-Walled Carbon Nanotubes Decorated with Ru Nanoparticles. *Carbon* **2009**, *47*, 1158–1164.
(16) Jonge, N. de; Allioux, M.; Doytcheva, M.; Kaiser, M.; Teo, K. B. K.; Lacerda, R. G.; Milne, W. I. Characterization of the Field Emission Properties of Individual Thin Carbon Nanotubes. *Appl. Phys. Lett.* **2004**, *85*, 1607.
(17) Han, S.; Ihm, J. First-Principles Study of Field Emission of Carbon Nanotubes. *Phys. Rev. B* **2002**, *66*.
(18) Kim, H.-S.; Lee, H.; Han, K.-S.; Kim, J.-H.; Song, M.-S.; Park, M.-S.; Lee, J.-Y.; Kang, J.-K. Hydrogen Storage in Ni Nanoparticle-Dispersed Multiwalled Carbon Nanotubes. *J. Phys. Chem. B* **2005**, *109*, 8983–8986.
(19) Kim, J.-Y.; Jeong, T.; Baik, C.-W.; Park, S. H.; Han, I.; Kim, G.-H.; Yu, S. Field-Emission Performance and Structural Change Mechanism of Multiwalled Carbon Nanotubes by Oxygen Plasma Treatment. *Thin Solid Films* **2013**, *547*, 202–206.
(20) Choi, H.; Ji Shin, Y.; Il Cha, S.; Ho Kang, I.; Bahng, W. Enhanced Field-Emission Capacity by Density Control of a CNT Cathode Using Post-Plasma Treatment. *Solid State Commun.* **2013**, *171*, 50–54.

(21) Xiao, L.; Liu, P.; Liu, L.; Jiang, K.; Feng, X.; Wei, Y.; Qian, L.; Fan, S.; Zhang, T. Barium-Functionalized Multiwalled Carbon Nanotube Yarns as Low-Work-Function Thermionic Cathodes. *Appl. Phys. Lett.* **2008**, *92*, 153108.
(22) Lyth, S. M.; Hatton, R. A.; Silva, S. R. P. Efficient Field Emission from Li-Salt Functionalized Multiwall Carbon Nanotubes on Flexible Substrates. *Appl. Phys. Lett.* **2007**, *90*, 013120.
(23) Jin, F.; Liu, Y.; Day, C. M.; Little, S. A. Enhanced Electron Emission from Functionalized Carbon Nanotubes with a Barium Strontium Oxide Coating Produced by Magnetron Sputtering. *Carbon* **2007**, *45*, 587–593.
(24) Feng, M.; Sun, R.; Zhan, H.; Chen, Y. Decoration of Carbon Nanotubes with CdS Nanoparticles by Polythiophene Interlinking for Optical Limiting Enhancement. *Carbon* **2010**, *48*, 1177–1185.
(25) Wang, W.; Serp, P.; Kalck, P.; Silva, C. G.; Faria, J. L. Preparation and Characterization of Nanostructured MWCNT-TiO2 Composite Materials for Photocatalytic Water Treatment Applications. *Mater. Res. Bull.* **2008**, *43*, 958–967.
(26) Haremza, J. M.; Hahn, M. A.; Krauss, T. D.; Chen, S.; Calcines, J. Attachment of Single CdSe Nanocrystals to Individual Single-Walled Carbon Nanotubes. *Nano Lett.* **2002**, *2*, 1253–1258.
(27) Huang, C.-S.; Yeh, C.-Y.; Chang, Y.-H.; Hsieh, Y.-M.; Ku, C.-Y.; Lai, Q.-T. Field Emission Properties of CNT–ZnO Composite Materials. *Diam. Relat. Mater.* **2009**, *18*, 452–456.
(28) Yang, M.; Liang, T.; Peng, Y.; Chen, Q. Synthesis and Characterization of a Nanocomplex of ZnO Nanoparticles Attached to Carbon Nanotubes. *Acta Phys.-Chim. Sin.* **2007**, *23*, 145–151.
(29) An, G.; Na, N.; Zhang, X.; Miao, Z.; Miao, S.; Ding, K.; Liu, Z. SnO_2 /carbon Nanotube Nanocomposites Synthesized in Supercritical Fluids: Highly Efficient Materials for Use as a Chemical Sensor and as the Anode of a Lithium-Ion Battery. *Nanotechnology* **2007**, *18*, 435707.
(30) Zhai, L.; Wei, Z.; Yang, Z.; Ni, X. Polymerization Initiated by ZnS Nanocrystals Anchored on Carbon Nanotubes. *Mater. Lett.* **2010**, *64*, 531–533.
(31) Cho, N.; Roy Choudhury, K.; Thapa, R. B.; Sahoo, Y.; Ohulchanskyy, T.; Cartwright, A. N.; Lee, K.-S.; Prasad, P. N. Efficient Photodetection at IR Wavelengths by Incorporation of PbSe–Carbon-Nanotube Conjugates in a Polymeric Nanocomposite. *Adv. Mater.* **2007**, *19*, 232–236.
(32) Li, C.; Ding, S.; Lei, W.; Zhang, X.; Wang, B. Enhanced Field Emission from Vertically Aligned Carbon Nanotubes on Metal Mesh Electrode. *Appl. Surf. Sci.* **2013**, *285*, 505–508.
(33) Zuo, Y.; Ren, Y.; Wang, Z.; Han, X.; Xi, L. Enhanced Field Emission and Hysteresis Characteristics of Aligned Carbon Nanotubes with Ti Decoration. *Org. Electron.* **2013**, *14*, 2306–2314.
(34) Pandey, A.; Prasad, A.; Moscatello, J. P.; Engelhard, M.; Wang, C.; Yap, Y. K. Very Stable Electron Field Emission from Strontium Titanate Coated Carbon Nanotube Matrices with Low Emission Thresholds. *ACS Nano* **2013**, *7*, 117–125.

(35) Zanin, H.; May, P. W.; Hamanaka, M. H. M. O.; Corat, E. J. Field Emission from Hybrid Diamond-like Carbon and Carbon Nanotube Composite Structures. *ACS Appl. Mater. Interfaces* **2013**, *5*, 12238–12243.
(36) Fan, Y. C.; Liu, Y. M.; Chen, Y. C.; Sung, Y.; Ger, M. D. Carbon Nanotube Field Emission Cathodes Fabricated with Chemical Displacement Plating. *Appl. Surf. Sci.* **2009**, *255*, 7753–7758.
(37) Chen, Y.; Jiang, H.; Li, D.; Song, H.; Li, Z.; Sun, X.; Miao, G.; Zhao, H. Improved Field Emission Performance of Carbon Nanotube by Introducing Copper Metallic Particles. *Nanoscale Res. Lett.* **2011**, *6*, 537.
(38) Nilsson, L.; Groening, O.; Emmenegger, C.; Kuettel, O.; Schaller, E.; Schlapbach, L.; Kind, H.; Bonard, J.-M.; Kern, K. Scanning Field Emission from Patterned Carbon Nanotube Films. *Appl. Phys. Lett.* **2000**, *76*, 2071.
(39) Bonard, J.-M.; Dean, K.; Coll, B.; Klinke, C. Field Emission of Individual Carbon Nanotubes in the Scanning Electron Microscope. *Phys. Rev. Lett.* **2002**, *89*.
(40) Bonard, J.-M.; Weiss, N.; Kind, H.; Stöckli, T.; Forró, L.; Kern, K.; Châtelain, A. Tuning the Field Emission Properties of Patterned Carbon Nanotube Films. *Adv. Mater.* **2001**, *13*, 184–188.
(41) Bittencourt, C.; Ke, X.; Van Tendeloo, G.; Thiess, S.; Drube, W.; Ghijsen, J.; Ewels, C. P. Study of the Interaction between Copper and Carbon Nanotubes. *Chem. Phys. Lett.* **2012**, *535*, 80–83.
(42) Gingery, D.; Bühlmann, P. Formation of Gold Nanoparticles on Multiwalled Carbon Nanotubes by Thermal Evaporation. *Carbon* **2008**, *46*, 1966–1972.
(43) Charlier, J.-C.; Arnaud, L.; Avilov, I. V.; Delgado, M.; Demoisson, F.; Espinosa, E. H.; Ewels, C. P.; Felten, A.; Guillot, J.; Ionescu, R.; *et al.* Carbon Nanotubes Randomly Decorated with Gold Clusters: From Nano ^{2} Hybrid Atomic Structures to Gas Sensing Prototypes. *Nanotechnology* **2009**, *20*, 375501.
(44) Muratore, C.; Reed, A. N.; Bultman, J. E.; Ganguli, S.; Cola, B. A.; Voevodin, A. A. Nanoparticle Decoration of Carbon Nanotubes by Sputtering. *Carbon* **2013**, *57*, 274–281.
(45) Sridhar, S.; Ge, L.; Tiwary, C. S.; Hart, A. C.; Ozden, S.; Kalaga, K.; Lei, S.; Sridhar, S. V.; Sinha, R. K.; Harsh, H.; *et al.* Enhanced Field Emission Properties from CNT Arrays Synthesized on Inconel Superalloy. *ACS Appl. Mater. Interfaces* **2014**, 140113152340002.
(46) Katayama, M.; Lee, K.-Y.; Honda, S.; Hirao, T.; Oura, K. Ultra-Low-Threshold Field Electron Emission from Pillar Array of Aligned Carbon Nanotube Bundles. *Jpn. J. Appl. Phys.* **2004**, *43*, L774–L776.
(47) Suh, J. S.; Jeong, K. S.; Lee, J. S.; Han, I. Study of the Field-Screening Effect of Highly Ordered Carbon Nanotube Arrays. *Appl. Phys. Lett.* **2002**, *80*, 2392.
(48) Ren, H.; Yang, L.; Zhang, Y. Numerical Calculations on the Field Emission of Carbon Nanotubes. *J. Phys. Conf. Ser.* **2013**, *418*, 012007.
(49) Tuinstra, F. Raman Spectrum of Graphite. *J. Chem. Phys.* **1970**, *53*, 1126.
(50) Saito, R.; Hofmann, M.; Dresselhaus, G.; Jorio, A.; Dresselhaus, M. S. Raman Spectroscopy of Graphene and Carbon Nanotubes. *Adv. Phys.* **2011**, *60*, 413–550.

(51) Ferrari, A. C. Raman Spectroscopy of Graphene and Graphite: Disorder, Electron–phonon Coupling, Doping and Nonadiabatic Effects. *Solid State Commun.* **2007**, *143*, 47–57.
(52) Atthipalli, G.; Wang, H.; Gray, J. L. Catalyst-Assisted Vertical Growth of Carbon Nanotubes on Inconel Coated Commercial Copper Foil Substrates *versus* Sputtered Copper Films. *Appl. Surf. Sci.* **2013**, *273*, 515–519.
(53) Lee, D. H.; Lee, J. A.; Lee, W. J.; Choi, D. S.; Lee, W. J.; Kim, S. O. Facile Fabrication and Field Emission of Metal-Particle-Decorated Vertical N-Doped Carbon Nanotube/Graphene Hybrid Films. *J. Phys. Chem. C* **2010**, *114*, 21184–21189.
(54) Gadzuk, J.; Plummer, E. Field Emission Energy Distribution (FEED). *Rev. Mod. Phys.* **1973**, *45*, 487–548.
(55) Tanaka, H.; Akita, S.; Pan, L.; Nakayama, Y. Barrier Effect on Field Emission from Stand-Alone Carbon Nanotube. *Jpn. J. Appl. Phys.* **2004**, *43*, 864–867.
(56) Fujii, S.; Honda, S.; Machida, H.; Kawai, H.; Ishida, K.; Katayama, M.; Furuta, H.; Hirao, T.; Oura, K. Efficient Field Emission from an Individual Aligned Carbon Nanotube Bundle Enhanced by Edge Effect. *Appl. Phys. Lett.* **2007**, *90*, 153108.
(57) Collins, P. G.; Zettl, A. Unique Characteristics of Cold Cathode Carbon-Nanotube-Matrix Field Emitters. *Phys. Rev. B* **1997**, *55*, 9391–9399.

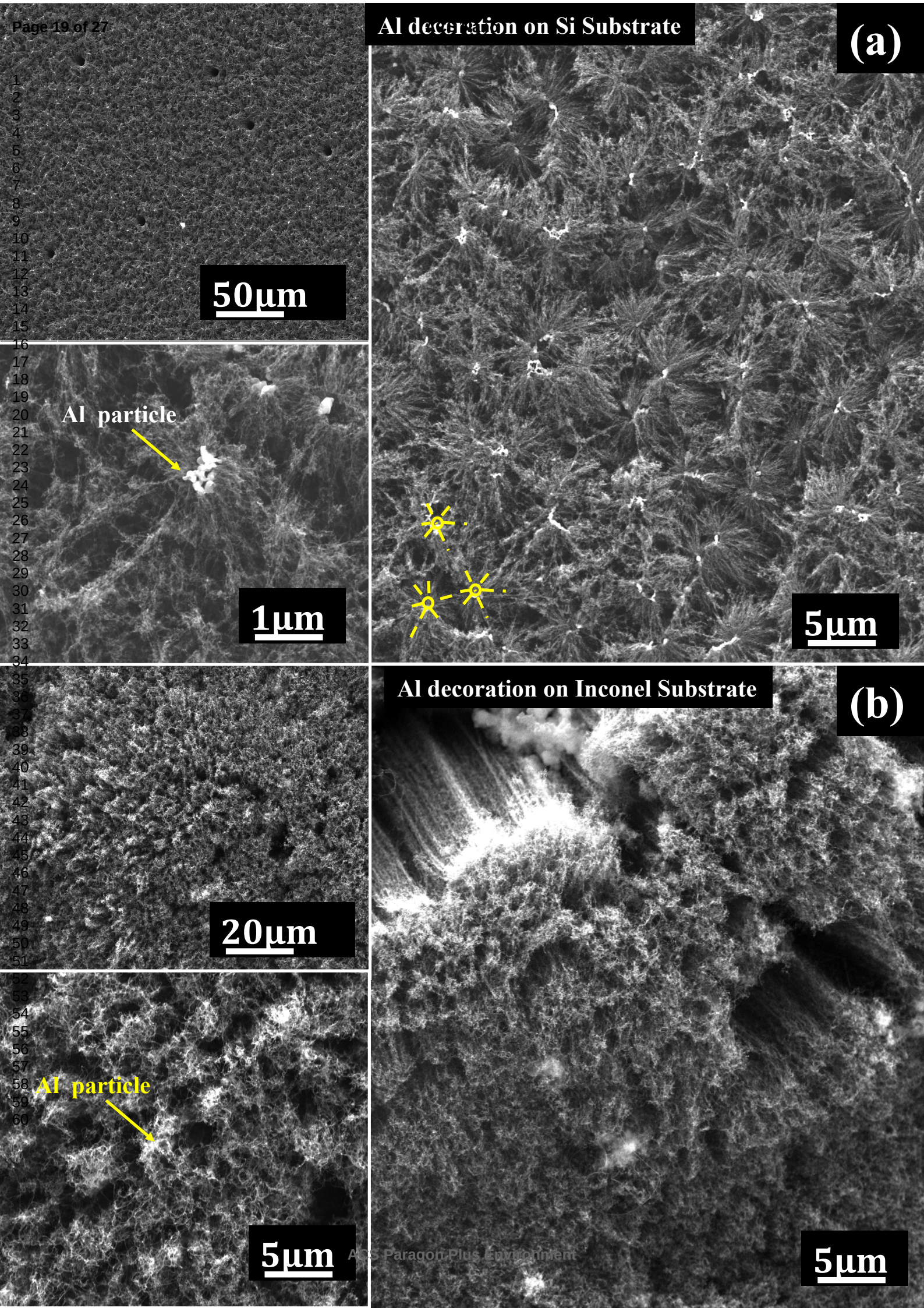
Field Emission with Ultra-Low Turn-On Voltage from Metal Decorated Carbon Nanotubes

Srividya Sridhar, [1] Chandrasekhar Tiwary, [2,6] Soumya Vinod, [2] Jose Jaime Taha-Tijerina, [2] Srividvatha Sridhar, [3] Kaushik Kalaga, [2] Benjamin Sirota, [5] Amelia H. C. Hart, [2] Sehmus Ozden, [2] Ravindra Kumar Sinha, [1] Harsh, [4] Robert Vajtai, [2] Wongbong Choi, [5] Krisztián Kordás [7] and Pulickel M. Ajayan[2]*

[1]Delhi Technological University (Formerly Delhi College of Engineering), Department of Applied Physics, Bawana Road, Delhi 110042, India.
[2]Department of Materials Science and Nano Engineering, Rice University, Houston, Texas 77005, United States.
[3]Department of Chemical and Biomolecular Engineering, Rice University, Houston, Texas 77005, United States.
[4]Department of Physics, Jamia Millia Islamia, New Delhi 110025, India.
[5]Department of Material Science and Engineering, University of North Texas, Denton, Texas 76203, United States.
[6]Materials Engineering, Indian Institute of Science, Bangalore, Karnataka-560012, India.
[7]Microelectronics and Materials Physics Laboratories, Department of Electrical Engineering, University of Oulu, P.O. Box 4500, FI-90014 Oulu, Finland.

*Corresponding Author: Dr. P. M. Ajayan; Email Address: ajayan@rice.edu

Figures



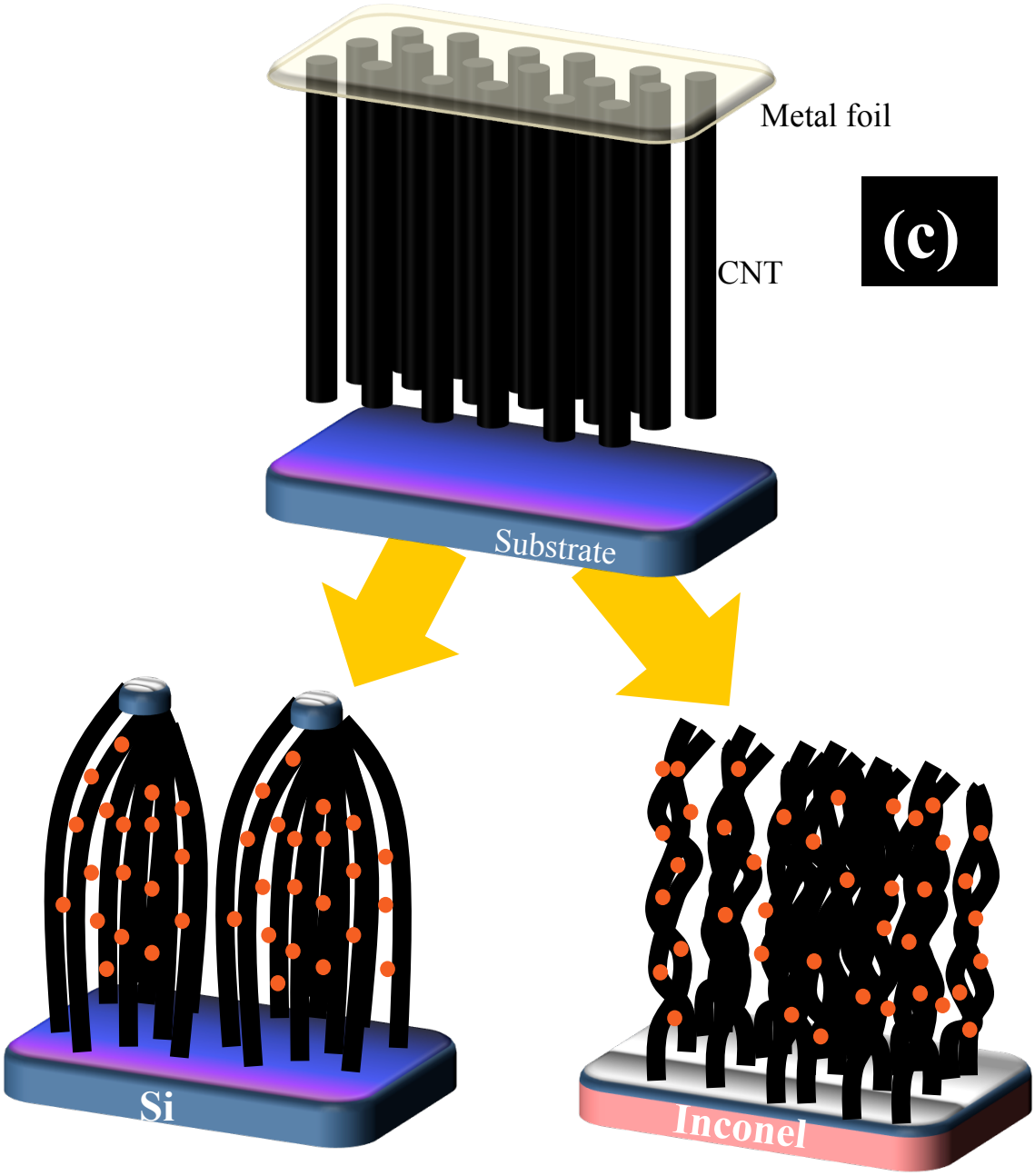
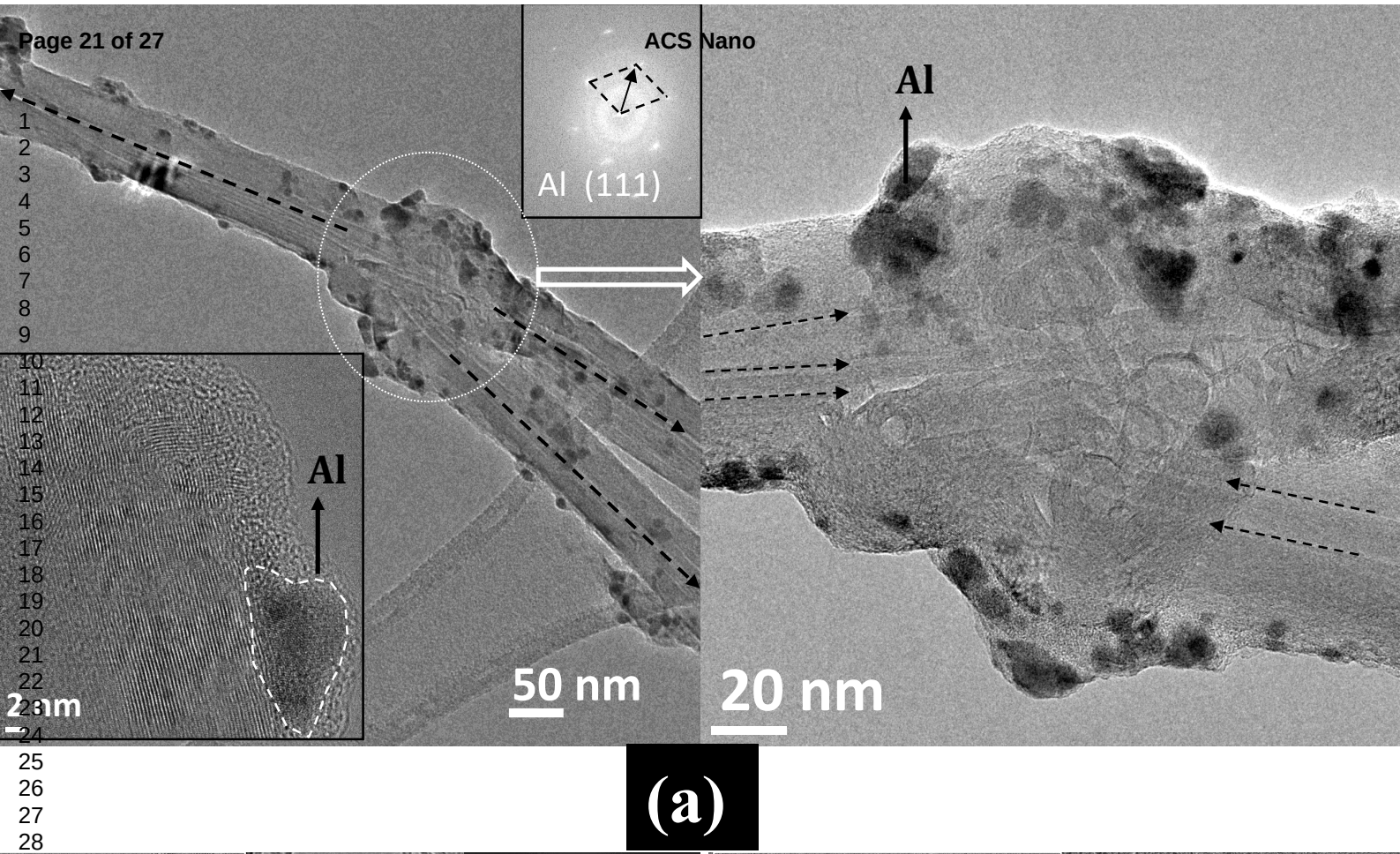


Figure 1

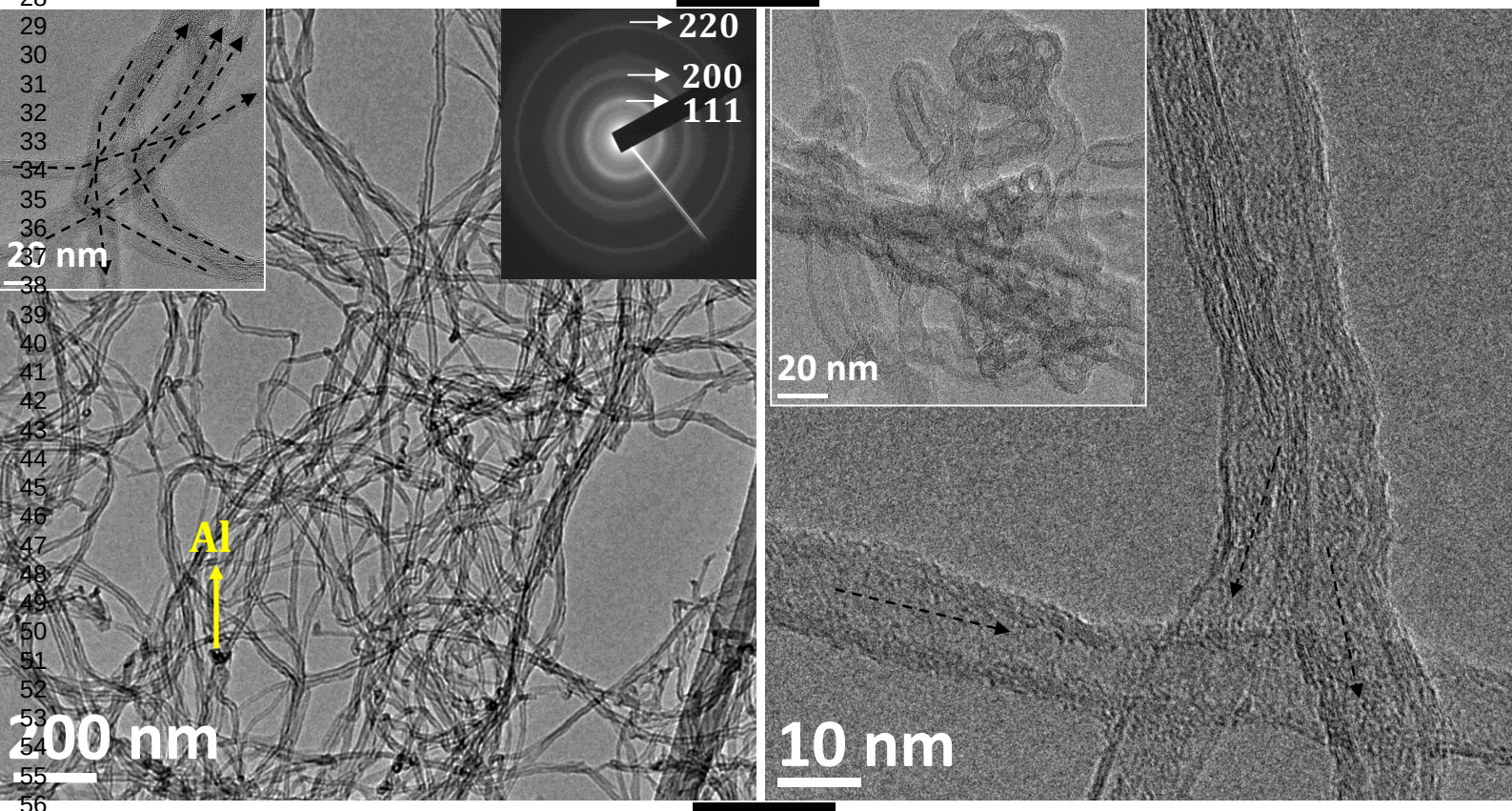
a) SEM image of CNTs grown on Si with Al decoration in different magnification. SEM image shows the way in which the CNTs got attached to Al particle forming a beautiful pattern.

b) SEM image of the CNTs grown on Inconel with Al decoration on them in different magnification.

c) The schematic procedure for the fabrication of metal decorated CNTs



(a)

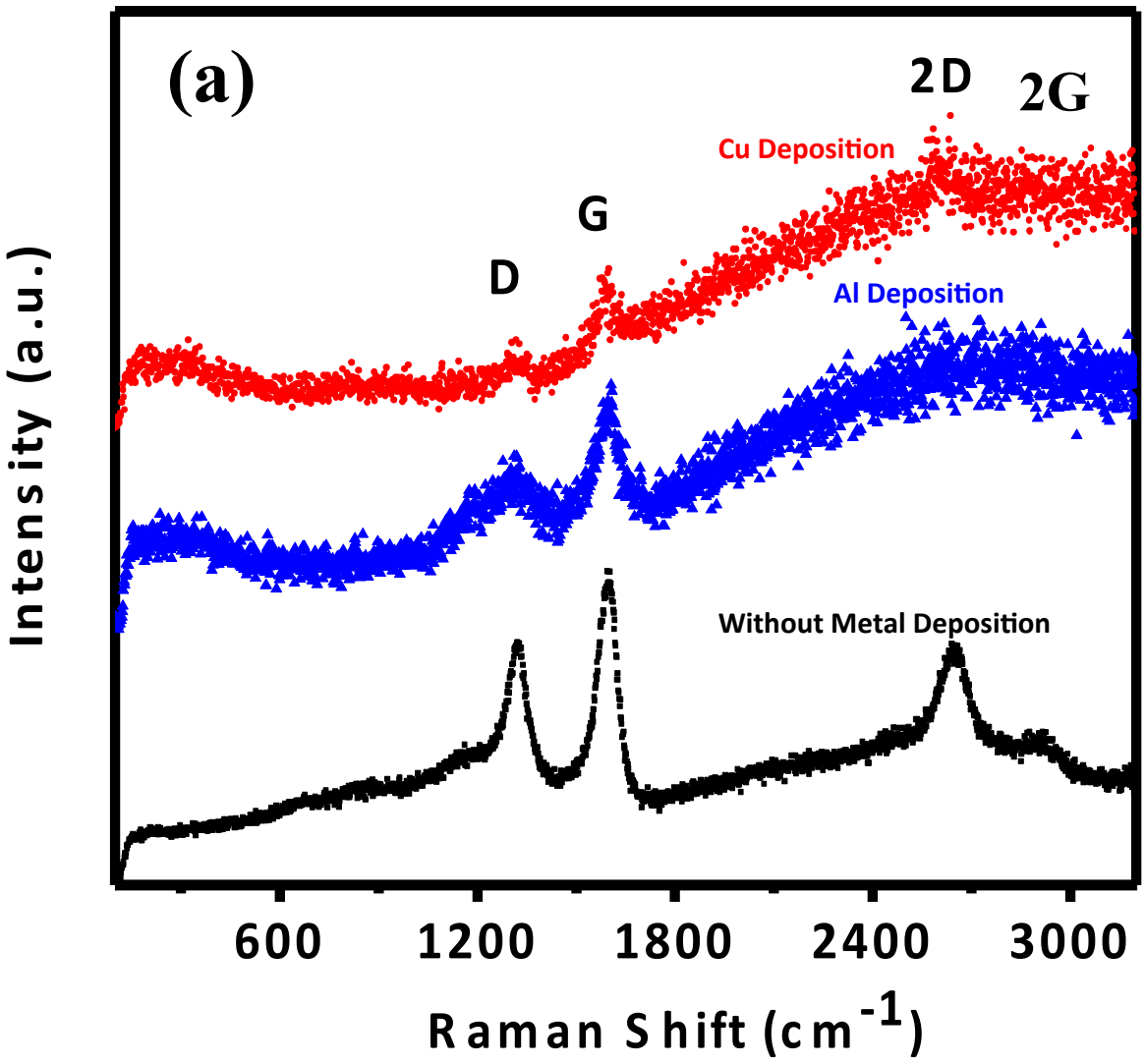


(b)

Figure 2

a) TEM and HRTEM micrograph of the CNTs grown on Si substrate with Al decoration on them. It shows how the CNTs are joined together at the tip and the high magnification confirms the presence of metal particle.

b) TEM and HRTEM micrograph of the CNTs grown on Inconel.


(a)
2D
2G
G
D
Cu Deposition
Al Deposition
Without Metal Deposition
Intensity (a.u.)
600
1200
1800
2400
3000
Raman Shift (cm-1)

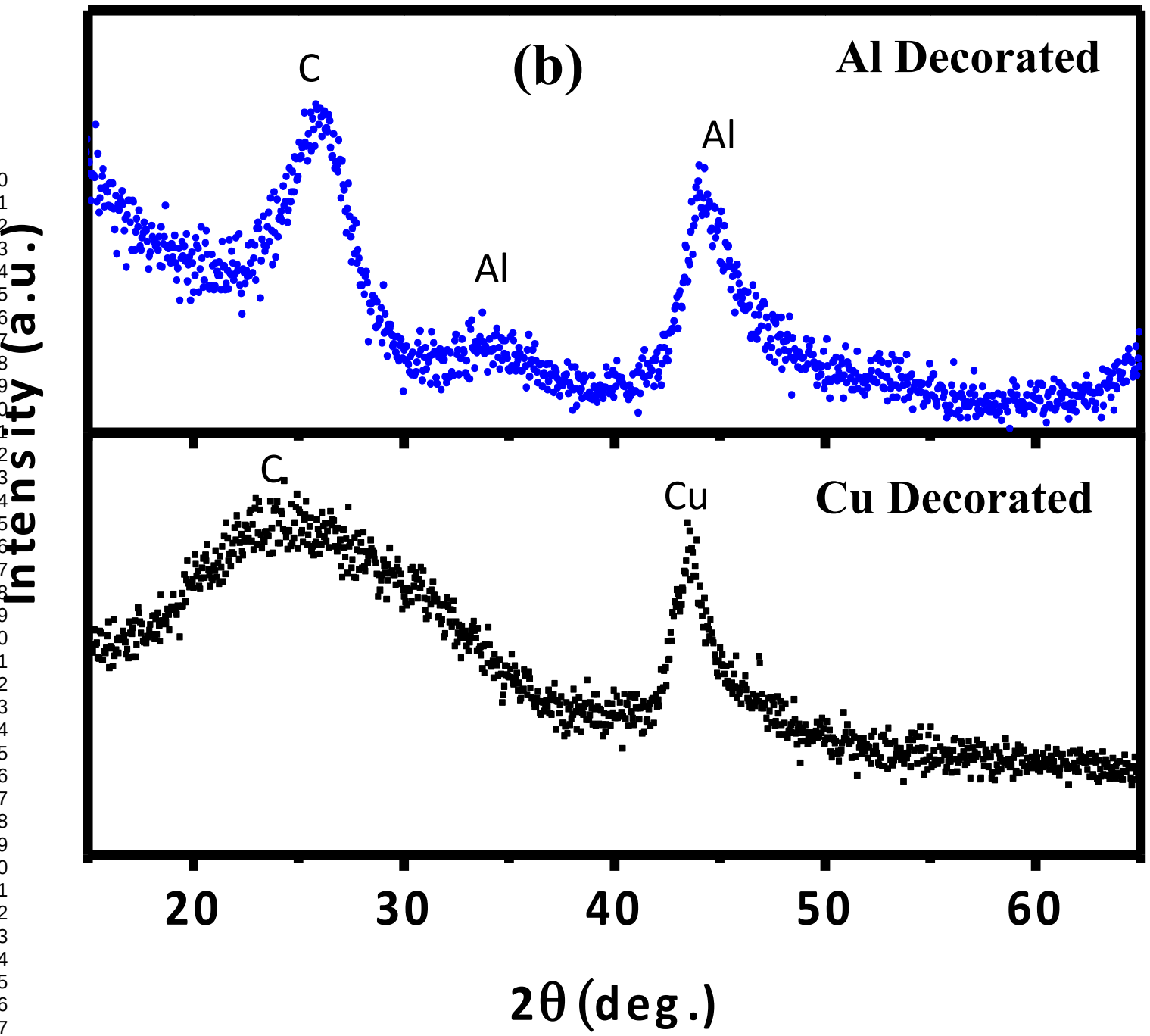
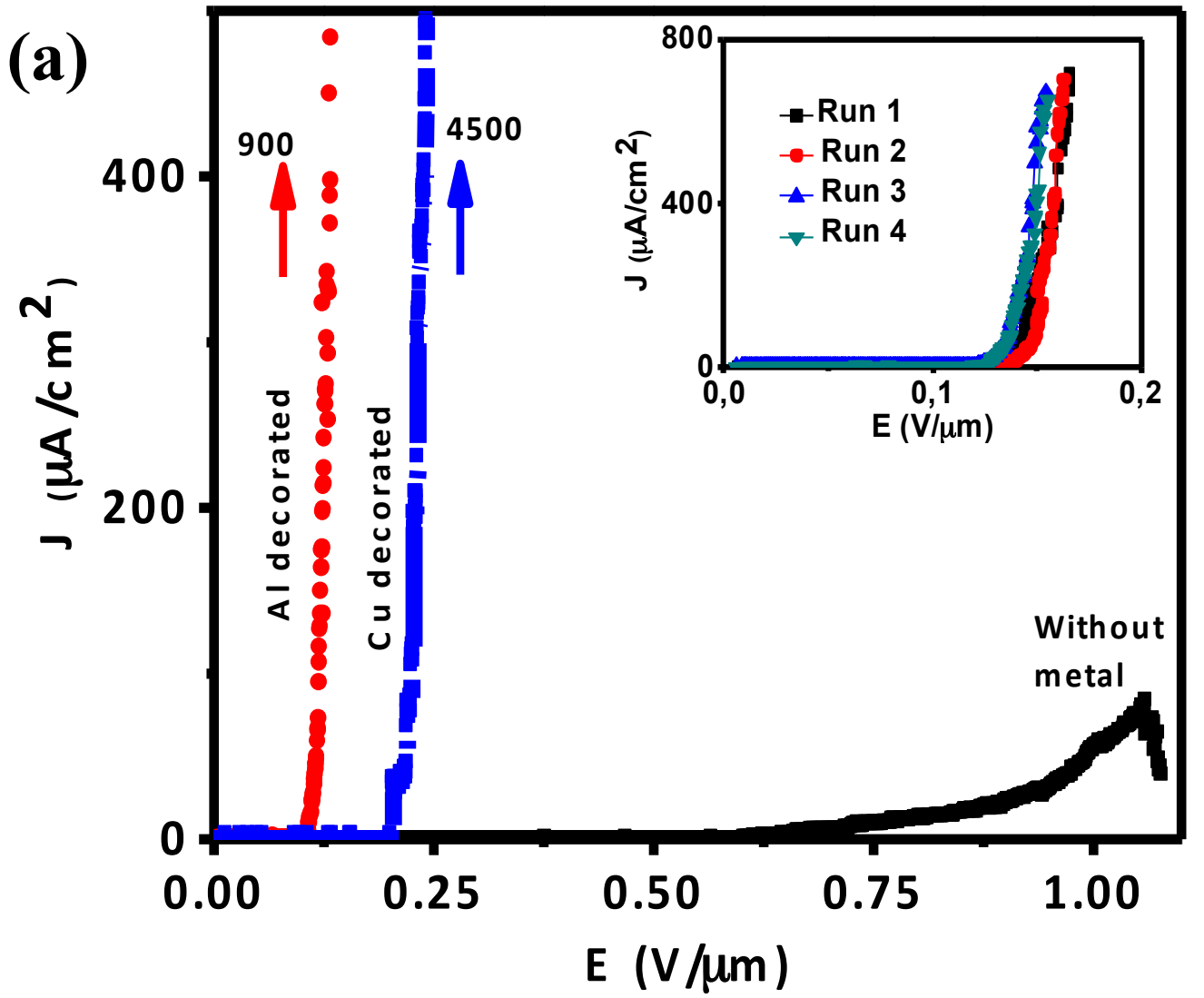


Figure 3

a) Raman spectra showing 3 distinct peaks [dis-order peak D, graphitic peak G and long range order peak G' for Si without metal decoration, with Al decoration and Cu decoration.

b) XRD pattern on Si with Al decoration and XRD pattern on Si with Cu decoration. Al, Cu and C peak clearly prove that there is no oxide formation


(a)
900
4500
Al decorated
Cu decorated
Without metal
J (μA/cm2)
E (V/μm)
Run 1
Run 2
Run 3
Run 4

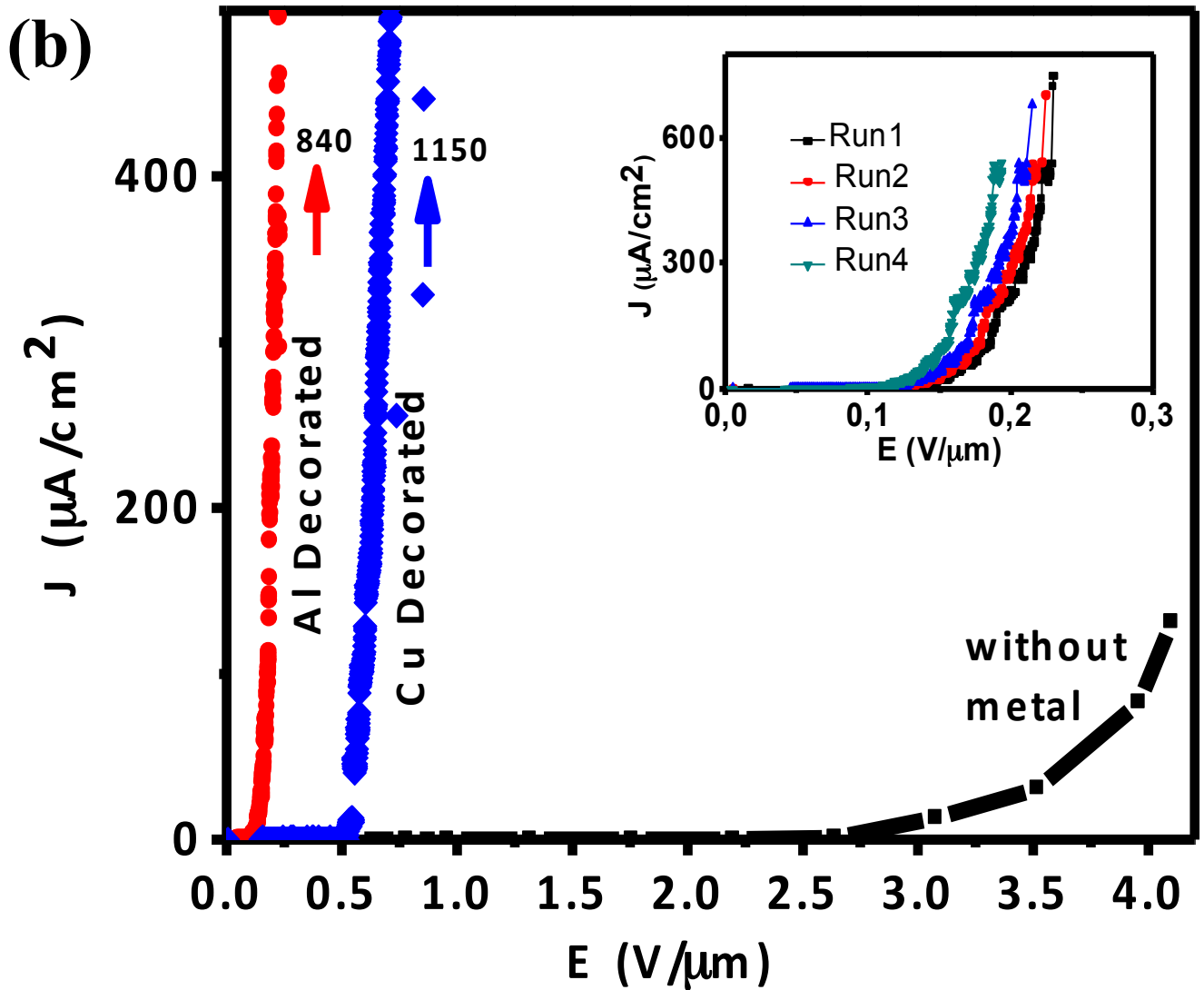

(b)
840
1150
Al Decorated
Cu Decorated
without metal
J (μA/cm2)
E (V/μm)
Run1
Run2
Run3
Run4

Fig.4

(c)

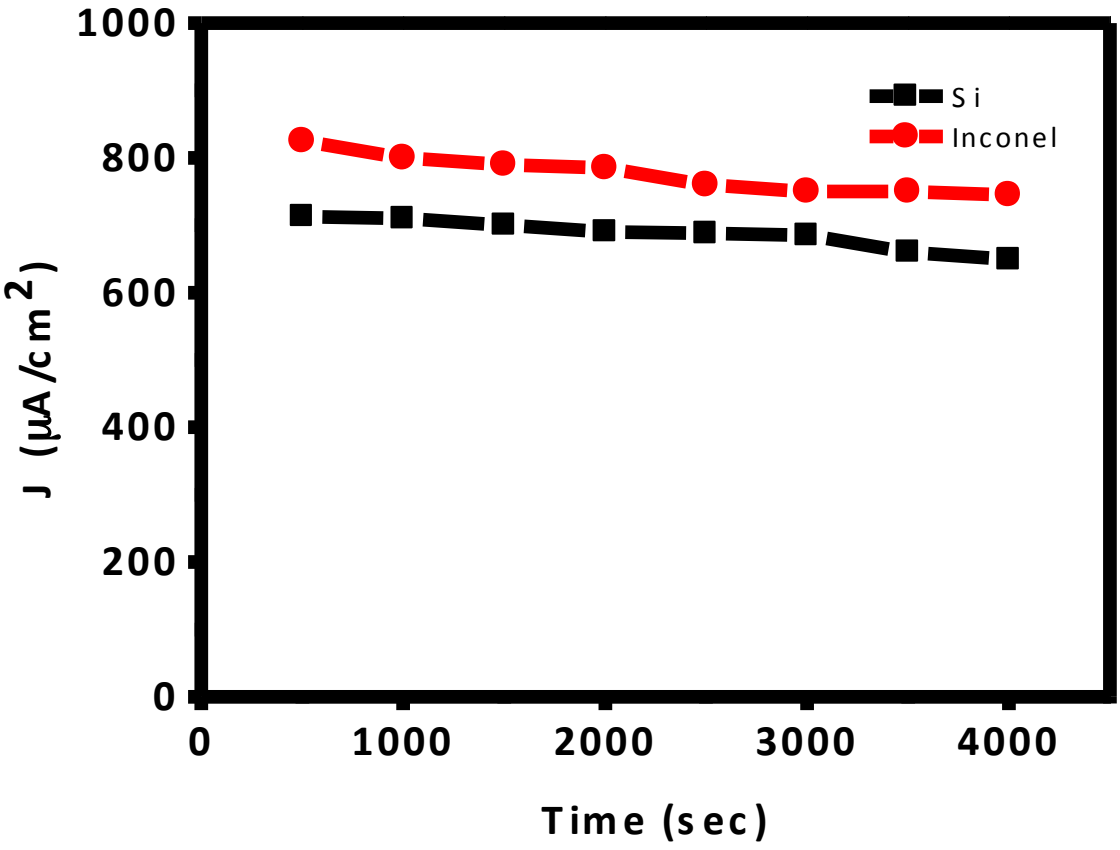


Figure 4

a) Field emission characteristics of the CNTs grown on Si with and without any metal decorations. It is very clear that the turn on field E_{to} and the threshold field E_{th} are significantly reduced after the metal decoration. Inset in the JE plot shows emission stability

b) Field emission characteristics of the CNTs grown on Inconel showing the same phenomena. Inset in the JE plot shows emission stability

c) Time trace of current density at the fixed field of 0.15 V/μm for Si grown CNTs and 0.25 V/μm for Inconel grown CNTs, both decorated with Al nanoparticles.

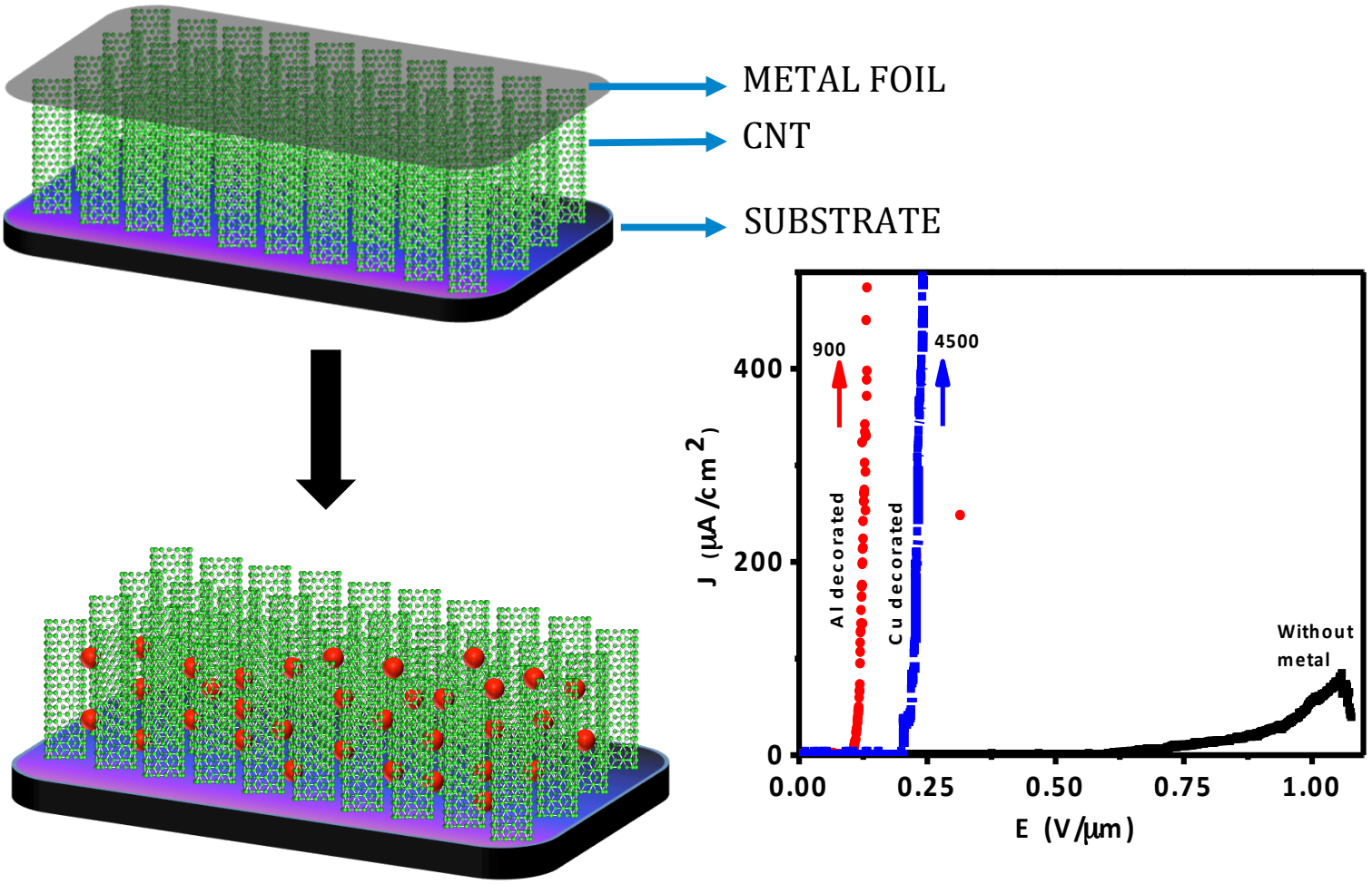
Field Emission with Ultra-Low Turn-On Voltage from Metal Decorated Carbon Nanotubes

Srividya Sridhar, [1] Chandrasekhar Tiwary, [2,6] Soumya Vinod, [2] Jose Jaime Taha-Tijerina, [2] Srividvatha Sridhar, [3] Kaushik Kalaga, [2] Benjamin Sirota, [5] Amelia H. C. Hart, [2] Sehmus Ozden, [2] Ravindra Kumar Sinha, [1] Harsh, [4] Robert Vajtai, [2] Wongbong Choi, [5] Krisztián Kordás [7] and Pulickel M. Ajayan[2]*

[1]Delhi Technological University (Formerly Delhi College of Engineering), Department of Applied Physics, Bawana Road, Delhi 110042, India.
[2]Department of Materials Science and Nano Engineering, Rice University, Houston, Texas 77005, United States.
[3]Department of Chemical and Biomolecular Engineering, Rice University, Houston, Texas 77005, United States.
[4]Department of Physics, Jamia Millia Islamia, New Delhi 110025, India.
[5]Department of Material Science and Engineering, University of North Texas, Denton, Texas 76203, United States.
[6]Materials Engineering, Indian Institute of Science, Bangalore, Karnataka-560012, India.
[7]Microelectronics and Materials Physics Laboratories, Department of Electrical Engineering, University of Oulu, P.O. Box 4500, FI-90014 Oulu, Finland.

*Corresponding Author: Dr. P. M. Ajayan; Email Address: ajayan@rice.edu

Table of Contents (TOC) Graphic


METAL FOIL
CNT
SUBSTRATE
900
4500
Al decorated
Cu decorated
Without metal
J (μA /c m 2)
E (V/μm)
0
200
400
0.00
0.25
0.50
0.75
1.00

BIOINFO Financial Management
ISSN: 2250-091X & E-ISSN: 2250-0928, Volume 3, Issue 1, 2013, pp.-066-069.
Available online at http://www.bioinfopublication.org/jouarchive.php?opt=&jouid=BPJ0000199

GENDER DIFFERENCES IN INVESTMENT

KANSAL P. AND SINGH S.*

Delhi Technological University, Delhi- 110 042, India.
*Corresponding Author: Email- seemasinghdtu@gmail.com

Received: November 14, 2013; Accepted: December 12, 2013

Abstract- There are number of studies which favour the existence of gender bias in investment patterns. These studies concluded that women behave more conservatively than men when comes to their investment decisions and that the women are more risk averse. Some studies also contend that women are less confident about their investment decisions and earn less return in compare to men. This paper aims to study whether these gender differences really exist in India. Using NCAER household survey report, Tthe paper examined the household investment and saving patterns of Indian households. The gender difference has been studied in terms of their risk tolerance *viz* their risk bearing capacity, risk perception *i.e.* how they assume risk within different constraints, time horizon for investment, preference for investment alternatives among the vast number of alternatives available in the Indian capital market etc. We also examined the factors which influence the investment behaviour of women such as age, level of education, their marital status, income, dependency etc. Data has been collected from NCAER household survey report and analysed through chi-square test using SPSS software.

JEL Classification: J16, G11, G23

Keywords- Gender Difference, Risk perception, Investment Behaviour, Women Investment, Time horizon, Preference, Portfolio diversification, Indian capital market

Introduction

After the global financial crisis of 2008 and 2009, the Indian financial sector has now emerged stronger. The investments and savings are increasing in terms of volumes and number of investors. As the number of investors is increasing, the most common discussion based on gender bias again becomes the interest of investors. There are number of studies, which have shown that the financial behaviour of men and women differ significantly. Women hold low risk tolerance i.e. are more risk averse than men and also sometimes earns less returns from their investments. Except this, women invest more conservatively their financial resources and have shows low confidence regarding their financial behaviour. In this paper we are studying the investment behaviour of individuals, residing in the territory of India, dividing them in the strata of male and female with reference to their preference for investment alternative, their risk behaviour, the level of confidence regarding their financial decisions, the factors which influence their behaviour, etc.

Literature Review

There are number of studies available, which have discussed the difference in financial behaviour of male and female investors and the factors which influence this behaviour. Researchers like Sunden, et al [1] Agnew [2], Olsen and Cox [3], Schmidt & Sevak [4] shows that women are more risk averse then men in general and this defines their choice of less risky assets in their portfolios.

Except this, Guiso, et al [5], Bajtelsmit & VenDerhei [6], Hariharan, et al [7], Hartog, et al [8] concluded that males are more risk tolerant than females. Powell & Ansic [9] find that men have significantly higher preference for risk than women. Males prefer "riskier" investment strategies in order to achieve the highest gains, while women select "safer" strategies that allow them avoiding the worst possible losses [10-14]. Graham, et al [15] found that women have less confidence regarding their decision related to financial issues.

Fellner & Maciejovsky [16] reveal a systematic correlation between gender and risk attitudes. Further Fellner & Maciejovsky [17] find that women prefer less volatile investments and exhibit lower market activity, e.g. they submit fewer offers and engage less often in trades.

Women give a lot of priority and importance to the advices given by Financial Advisors (FA) and depend on them for guidance than men. Female investors are more detail oriented; and want to read more and understand financial matters better and they ask more questions than male clients [18].

Jianakoplos & Bernasek [19] test gender differences in investment behavior on a large data set drawn from the Survey of Consumer Finances (CFS) 1989. The analysis reveals that single women are relatively more risk averse than single men or married couples.

Some studies concluded that since women earn less compare to men, they have lower wealth accumulation and hence lower investment and saving rates [20-22].

In India there are very few studies which have presented these gender gaps. Madhusoodanan [22] suggested that risk tolerance serves as an illusion of control and thus overconfidence. Somasundaram [23] concluded that the investors of Coimbatore prefer the bank and chit fund deposits to save. Mutual funds are the least preferred instrument for investment. Rajarajan [24] and Shobhana and Jayalaxmi [25] bought out the fact that there is a strong associ-

ation between the demographic factors and the risk tolerance of the investors. Sharma & Sharma [26] pointed out that the retail investment activity in India is very low. However the young generation of Jammu holds a positive attitude towards the stock market with moderate belief. Rajarajen [27] revealed that the population of Indian investors have increased in the recent years. Similarly, the surveys organised by government bodies such as NCAER, RBI Survey talk about the individual investor's investment and saving behaviour. In the same line, also, Hira and Mugenda [28] state that an advisor needs to understand the factors that underlie a client's financial behaviours before they can effectively advise them, and Deb and Chavali [29] have shown that men and women think and behave differently when it comes to managing money.

However there exist some studies exist which denies the existence of any gender gap. For example Schubert, et al [30] find no influence of gender on financial decisions. Masters & Meier [31] found no difference in the risk taking propensity of male and female entrepreneurs.

Research Methodology

Formulation of Hypotheses: Hypotheses are formulated considering '*no difference'* exists in the investment behaviour among Male and Female investors.

Ho: There will be no difference between male and female households with regard to:

- The preferred investment alternative
- The amount allocated to investments
- The time horizon for investments
- The relative risk aversion
- The risk tolerance
- The perception of risk
- The level of confidence
- The use of windfall income
- The portfolio diversification strategy
- The importance of expert's advice
- The factors influencing the investment in capital market

Sample

The data set used in this study is the 2011 Survey of National Council of Applied Economic Research, titled "How Household Invest: Evidence from NCAER Household Survey" sponsored by Securities Exchange Board o India (SEBI) [32].

The household is the basic unit of analysis in the study. The survey was conducted in 52 major states/ Union Territories of India. The sample size is 38,412.

Chi square test is used for analyze the formulated hypotheses with the help of SPSS 21.0 package.

Data Analysis

Table-1 represents the summarised data which is given in NCAER survey report. However we have summarized the data as per our requirements. We have worked on the following 11 parameters:

Parameter 1: Investment Preference:

As per the survey, assuming a level of risk aversion and information asymmetry, investment preference shows how a rupee of surplus income has been allocated across various investment options by the households. We find that mutual funds constitute the single largest allocation by male (40.75%) and female (41.63%) compared to all other options. Since mutual funds provide returns that are in general greater than market returns and expose investing households to risks that are lower than the market risks, Both male and female households prefer this medium over retail investing. Retail investing is "costlier" in terms of time and information as well as the variability of returns. This explains why a mere 21.38 percent (male) and 19.23 percent (female) of all households prefer to invest in the secondary market. Other choices such as derivatives and bonds are even less preferred by male investors. However it is noticeable that female investors are more interested in derivatives.

Parameter 2: Level of Investment

Report shows that men and women invest almost same amount.

Parameter 3: Time Horizon

This shows the relationship between gender and time horizon for investment. It is interesting to know that women are involve in more speculative behaviour when compared to men.

Parameter 4: Relative Risk Aversion

The relative risk aversion (risk tolerance) of a household has been captured using a risk scale, which can be defined as; The risk scale reflects the proportion of risky assets in an investor's portfolio. The numerator of this ratio is the value of investments in risky assets and the denominator is the total value of financial wealth. The risk scale is bounded between 0 and 1. We divide the risk scale into four categories, viz., less than equal to 0.25, greater than 0.25 but less than equal to 0.50, greater than 0.50 but less than equal to 0.75 and, greater than 0.75. The degrees of the risk tolerance scale are in increasing order from 0 to 1. The study shows that most of the investors are risk averse. Here also, women involve in less risky behaviour.

Parameter 5: Risk Tolerance

The self-perception of households with respect to their willingness to take risk is examined using the following neutral statement:

Which of the following statements is true for you? Willingness to take substantial financial risks/ Willingness to take above-average financial risks, expecting to get above average financial returns/ Willingness to take average financial risks, expecting to get average financial returns/ Not willing to take any financial risks.

The findings are consistent with the findings of Parameter 4 i.e. majority of investors either men or women are low risk taker.

Parameter 6: Confidence Level

NCAER used the following vignette to judge the level of confidence:

"You have saved money for a "world tour" that you were looking forward to for a long time. A month before you plan to leave, you lose your job. You would: [Very low level of confidence=Cancel the trip; Low level of confidence=Take a shorter vacation; High level of Confidence=Go as scheduled, reasoning that you will use that time to prepare for a job search; Very High level of Confidence=Extend your vacation, because this might be your only chance for such a trip]".

The responses are not in the line of previous studies which say that women are less confident. Table-1 shows the less confident behavior of men.

Table 1- Data Analysis

Parameters of Study	Men	Women
Investment Preference		
Mutual Fund	40.75	41.63
Bond Only	15.09	14.79
Debenture Only	8.68	6.74
IPO Only	8.71	4.51
Secondary Market Only	21.38	19.23
Derivatives Only	5.4	13.71
Level of Investment		
Lowest	19.89	21.7
IInd Quintile	19.91	20.69
IIIrd Quintile	20.01	19.94
IVth Quintile	20.47	16.58
Highest	19.71	21.09
Preferred Time Horizon for Investments		
Upto 3 years	33.68	37.98
3 to 5 years	32.23	33.99
More than 5 years	34.08	28.03
*Relative Risk Aversion		
< 0.25	49.22	53.32
0.25 - 0.50	14.06	16.04
0.5 - 0.75	16.91	16.94
> 0.75	19.81	13.69
Risk Tolerance		
No Risk Taker	39.55	36.38
Moderate Risk Taker	20.9	24.09
High Risk Taker	17.76	21.32
Substantially High Risk Taker	21.78	18.21
Confidence Level		
Very Low	32.62	23.99
Low	36.24	44.32
High	24.55	23.99
Very High	7.59	7.69
Use of Windfall Income		
Only Saving	31.39	31.79
invest in risk free securities	23.49	23.48
invest in securities having moderate risk	20.3	23.48
invest in securities having high risk	13.88	11.83
no saving or investment	10.95	9.43
Risk Definition		
Loss	22.13	23.82
Uncertainty of Returns	42.43	42.36
Opportunity for Significant Return	27.32	24.91
Thrill	8.11	8.91
Portfolio Diversification		
60% in Low Risk, 30% in Medium Risk and 10 % in High Risk	27.71	22.45
30% in Low Risk, 40% in Medium Risk and 30 % in High Risk	50.61	52.5
10% in Low Risk, 40% in Medium Risk and 50 % in High Risk	21.68	25.05
Following the Experts		
Truther	25.28	26.74
Liberal	26.65	24.54
Free thinker	27.09	29.3
Contrarian	20.98	19.41
Factors affecting the investment in capital market		
Returns	5.07	3.22
Surety of Returns	13.6	12.22
Liquidity	8.23	8.08
Information about instruments	26.85	22.24
Knowledge	14.27	16.38
Role of Regulator	6.19	7.73
Financial Resources	25.81	30.15

Parameter 7: Use of Windfall Income

Windfall gains are unexpected increase in income. One would expect such unexpected income to be used for relatively risky ventures. The following vignette is used to know the behaviour of households:

"If you unexpectedly receive ' 50,000, what would you do? [Saving Only=Deposit it in a bank account; Invest in risk free securities =Invest in high-quality govt. bonds; Invest in securities having moderate risk=Invest in mutual funds; Invest in securities having high risk=Invest in stocks; No saving or investment=Spend it]".

Responses shows men and women show the same attitude towards the use of windfall income and prefer to save it.

Parameter 8: Definition of Risk

This was the simply asked question the option mentioned in Table-1 under parameter 8. Investors take it as the uncertainty of returns.

Parameter 9: Portfolio Diversification

This parameter discussed the portfolio diversification strategy adopted by me and women. Table-1 show that women consist more risky securities in their portfolio.

Parameter 10: Expert Advice

Basically this parameter shows the herd behaviour of investors. The following vignette is used to know the herd behaviour of households:

"Some experts are predicting that the prices of assets such as gold, jewels, collectibles, and real estate (hard assets) will increase in value; bond prices may fall. However, experts have advised you that government bonds are relatively safe. Most of your investments are currently in high-interest government bonds. What would you do? [Truther=Hold the bonds; Liberal=Sell the bonds, put half the proceeds into the stock market, and the other half into assets such as land; Free Thinker=Sell the bonds and put all the money into buying land and precious metals; Contrarian=Sell the bonds and put all the money into buying assets like land and borrow additional money to buy more assets such as land]".

The study shows that when it comes to follow the advice of experts, both male and female shows the little bit influence of expert's advice.

Parameter 11: Factors Influencing the Investment in the Stock Market

As per the NCAER survey report, women invest less in stock market because of their resource constraints whereas men invest less because they don't know about the investment options.

Interpretation

The above discussed parameters are the basis of our study to know if there is any difference in regarding the financial behaviour of male and female households. As mentioned above the null hypotheses are tested with the help of Pearson Chi- Square Test using SPSS 21.0. The Chi -square test is done at the significant level of 5%, i.e. calculated p-value is compared with 0.05. Table-2 shows that p - value is higher than the level of significant for all the parameters i.e. p> 0.05, hence all the null hypotheses are accepted. This means there is no significance difference among the male and female households with regard to the preferred investment alternative, the amount allocated to investments, the time horizon for investments, the relative risk aversion, the risk tolerance, the perception of risk,

the level of confidence, the use of windfall income, the portfolio diversification strategy, the importance of expert's advice and the factors influencing the investment in capital market.

Table 2- Data Interpretation

Parameters	Chi-square	p-value	df	Critical value	Decision
Investment Preference	5.056	0.409	5	11.07	Accepted
Level of Investment	0.382	0.984	4	9.488	Accepted
Preferred Time Horizon for Investments	0.863	0.649	2	5.991	Accepted
Relative Risk Aversion	1.349	0.718	3	7.815	Accepted
Risk Tolerance	1.021	0.796	3	7.815	Accepted
Confidence Level	2.222	0.528	3	7.815	Accepted
Use of Windfall Income	0.579	0.965	4	9.488	Accepted
Risk Definition	0.218	0.975	3	7.815	Accepted
Portfolio Diversification	0.945	0.623	2	5.991	Accepted
Following the Experts	0.325	0.955	3	7.815	Accepted
Factors affecting the investment in capital market	1.864	0.932	6	12.592	Accepted

Conclusion

As the position of women is changing day by day, now women are also having a strong profile in the society. Now she is playing a role of not only a homemaker, but also dealing with high valued decisions including financial decisions at par with men. The above discussion shows that the decisions taken by the women are not different from the decisions taken by men. This result is however not in the line of previous studies which say that the financial behaviour of men and women have the significant behaviour. The reason may be the recent financial turmoil due to which a significant change has occurred in the investment activities of both men and women.

$$r_i = \frac{RA_i}{TA_i}$$

i = 1,2,3,4, …… N

$r_i \rightarrow 0$ Reflects that household is risk averse

$r_i \rightarrow 1$, Reflects that household tolerate risk

Reflects that household tolerate risk

where *ri* is the proportion of risky assets, *RAi* is asset holdings with some degree of risk, Tai is total asset holdings, *i* is the household.

Conflicts of Interest: None Declared.

References

[1] Sunden A and Surette B. (1998) *American Economic Review*, 88(2), 207-212.

[2] Agnew J. (2005) *Journal of Financial and Quantitative Finance*, 41, 939-962.

[3] Olsen R.A. and Cox C.M. (2001) *Journal of Behavioural Finance,* 2(2(1).

[4] Schmidt L. and Sevak P. (2006) *Feminist Economics*, 12(12), 139-166.

[5] Guiso L., Jaelli T. and Terlizzese D. (1996) *American Economic Review*, 86, (1), 158-172.

[6] Bajtelsmit V.L. and VanDerhei J.A. (1997) *Positioning pensions for the twenty-first century,* University of Pennsylvania Press, Philadelphia.

[7] Hariharan G., Chapman K.S. and Domain D.L. (2000) *Financial Services Review*, 9(2), 159-170.

[8] Hartog J., Ferrer-i Carbonell A. and Jonker N. (2002) *Kyklos*, 55 (1), 3-26.

[9] Powell M. and Ansic D. (1997) *Journal of Economic Psychology,* 18(6), 605-628.

[10] Odean T. (1998) *Journal of Finance*, 53(5), 1775-1798.

[11] Barber B. and Odean T. (2001) *Quarterly Journal of Economics*, 116(2), 261-292.

[12] Shlomo B. and Thaler R.H. (2001) *American Economic Review,* 91(1), 79-98.

[13] Gervais S. and Odean T. (2001) *Review of Financial Studies*, 14(1), 127.

[14] Dorn D. and Huberman G. (2003) *European Finance Association Meetings in Glasgow.*

[15] Graham J., Stendardi E., Meyers J. and Graham M. (2002) *International Journal of Bank Marketing*, 20(1), 17-26.

[16] Fellner G. and Maciejovsky B. (2007) *Journal of Economic Psychology,* 28(3), 338-350.

[17] Worley D. (1998) *Florida Today*, 12C.

[18] Jianakoplos N.A. and Bernasek A. (1998) *Economic Inquiry,* 36 (4), 620-630.

[19] Blau F.D. and Kahn L.M. (2000) *Journal of Economic Perspectives*, 14, 75-99.

[20] Moore Q., Shierholz H. (2004) '*Why did the convergence of male and female wages slow during the 1990s?*' Unpublished manuscript, Michigan Retirement Research Center, University of Michigan, Ann Arbor, MI.

[21] O'Neill J. (2003) *American Economic Review*, 93(2), 309-314.

[22] Madhusoodanan T.P. (1997) *Finance India*, 11(2), 285-304.

[23] Somasundaram V.K. (1998) *'A Study on Savings and Investment Pattern of Salaried Class in Coimbatore district'. Unpublished Thesis,* Bharathiyar University, Coimbatore, Tamilnadu.

[24] Rajarajan V. (2003) *Finance India*, 17(2), 565-576.

[25] Shobhana V.L. and Jayalakshmi J. (2005) *Working paper series.*

[26] Sharma B.C and Sharma D. (2004) *Indian Journal of Finance and Research*, 14(1&2).

[27] Rajarajen V. (2010) *Indian Institute of Finance*, 4, (3&4), 1274-1294.

[28] Hira T. and Mugenda O. (2000) *Journal of Financial Planning*, 13(2), 86-92.

[29] Deb M. and Chavali K. (2009) *Asia Pacific Business Review*, 5 (3), 45-55.

[30] Schubert R., Brown M., Gysler M. and Brachinger H.W. (1999) *American Economic Review,* 89 (2), 381-385.

[31] Masters R. and Meier R. (1988) *Journal of Small Business Management,* 26(1), 31-35.

[32] National Council of Alied Economic Research (2011) '*How Household Invest: Evidence from NCAER Household Survey',* sponsored by Securities Exchange Board of India (SEBI).

ISSN (Online) 2321 – 2004
ISSN (Print) 2321 – 5526

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH IN ELECTRICAL, ELECTRONICS, INSTRUMENTATION AND CONTROL ENGINEERING
Vol. 2, Issue 7, July 2014

Hardware efficient watermarking technique for finite state sequential circuit using STG

Jeebananda Panda[1], Ankur Bharadwaj[2], Neeta Pandey[3], Asok Bhattacharyya[4]

Associate Professor, ECE Department, Delhi Technological University, Delhi, India[1,3]

Assistant professor, ECE Department, JPIIT, Noida, India[2]

Professor, ECE Department, Delhi Technological University, Delhi, India[4]

Abstract: Intellectual Property Protection (IPP) is very important for a design created by IP owner. For this, IP owner embeds watermark in its design. One such type of technique is suggested by Oliviera, in which modification of State Transition Graph (STG) of a digital circuit takes place in such a way that it is not possible for the intruder to find that there is a watermark embedded in the circuit. It is also possible to prove the piracy of the design in court-of-law. A method for state reduction in the watermarked circuit has been proposed in this paper. The comparison of simulation of non-watermarked, watermarked circuit with existing technique and the modified reduced state watermarked circuit is done using ModelSim Simulator. The Detection of Piracy can be done by using a counter circuit.

Keywords: Intellectual Property Protection, State Transition Graph, Watermark, Finite state machine, Signature sequence

I. INTRODUCTION

There are various watermarking techniques for watermarking a sequential design. Finite State Machines are of two types: Completely specified and incompletely specified. Incompletely specified machines contain unused transitions which may be used for embedding watermark [3]. We are considering the watermarking of completely specified machines. In completely specified machines we can add extra input and output pairs to the original FSM [4] such that on the application of a particular input which only owner knows, piracy can be detected by observing the output sequence. But in complex designs finding such an input sequence is itself a tough task and it also adds to design overhead. Another method of watermarking includes embedding a signature in the design by state encoding [6].Oliviera [1] gave various techniques to embed digital watermark in sequential circuits, which include adding extra states in the original design. But, still the design overhead caused in the watermarking of a circuit is an important problem.

In this paper we propose a technique which can reduce the overhead caused due to watermarking of the sequential design. The synthesis results of the proposed technique have shown significant reduction in hardware overhead.
Different simulation tools are available for simulating a physical design before actually implementing them which made complex design of systems easy to test. But these designs can be easily stolen and used for unintended purposes without taking permission from IP owners. So, there has to be a method for protecting and tracking the ownership of such complex designs. Also, if necessary there should be a method of proving the ownership.

For this purpose, Oliviera[5] had suggested a watermarking algorithm in which State Transition Graph of a sequential circuit is modified in such a way that only owner knows its internal transitions and the user of that circuit can't even know that there is any watermark in the circuit. But, the watermarking algorithm used by Oliviera[5]adds some redundant states to the original STG which adds to the design overhead in complex designs. These redundant states can be reduced to certain extent depending upon the characteristics of State Transition Graph and choosing the signature sequence carefully. We have simulated watermark circuit with reduced number of redundant states and compared it with original watermarked circuit. We found that there was a reduction in number of flip-flops required to watermark same sequential circuit design.

II. WATERMARKING OF FINITE STATE MACHINES

Oliviera[5] proposed a technique for watermarking a completely specified FSMs. In this technique the STG is modified in such a way that authentication of the design can be proved by applying a signature sequence in addition to the input sequence of arbitrary length. The procedure for modification of STG[5] is explained below for reference-

1. Copy STG V for original STG Q in a way such that there is a corresponding qi ϵ Q state for every vi ϵ V state.
2. Create STG R with state ri, ($1 < I \leq k$) by copying state qti and all its outgoing edges, where qti refers to the state reached in STG Q at time t when the signature sequence is applied to STG Q.
   Here,
   k=(No. of bits in signature/No. of input bits)
   Note that r0 = q0
3. For each value of i such that $1 \leq i \leq k$, state qti has one of its incoming edges which now originates in state r(i - 1) and terminates at ri. For the input sequence

ISSN (Online) 2321 – 2004
ISSN (Print) 2321 – 5526

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH IN ELECTRICAL, ELECTRONICS, INSTRUMENTATION AND CONTROL ENGINEERING
Vol. 2, Issue 7, July 2014

corresponding to the signature, the edge originating in state r(i - 1) terminates at ri only. If the input sequence does not correspond to the signature, then the state reached from ri will be one of the states in the original STG Q for i < k while that in the duplicated STG V for i ≥ k so that the design conserves its functionality.

Thus the state transitions take place in STG Q or Q & R for i ≤ k and in STG Q, R and V for i ≥ k. In both cases the functionality obtained is the desired one. So, it is not possible to know for the intruder that STG has been modified. However, owner can detect the piracy by using the Piracy detection Circuit.

Design and Implementation of non – watermarked and watermarked circuit (As proposed by Oliviera [5]).
To elaborate the Oliviera [5] scheme, the STG by of Fig. 1 is considered. Following the steps outlined above, the STG of Fig. 1 modifies to STG of Fig. 2 for signature sequence '010'.

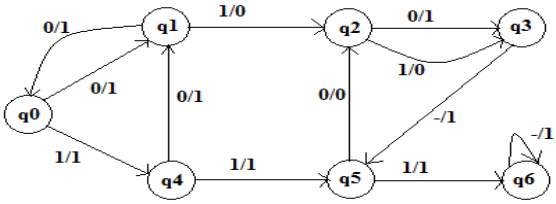


Fig.1 STG proposed by Oliviera[5].

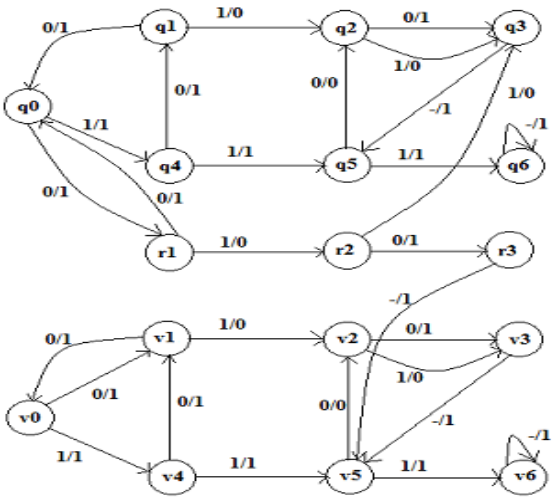


Fig.2 Watermarked STG consisting of 17 states [5]

It may be noted from Fig. 2 that all the states of the original STG are copied in the modified STG, which becomes a large hardware overhead in complex systems. However, It is not necessary to copy all the states of STG in some conditions. Given that the starting state is q0, the traversal is made via states r1, r2 and r3 on receiving the input sequence '010'. Subsequent inputs will make state transition to v5 for STG of Fig. 2 and thereafter vi states are traversed. The proposed technique is based on the fact that the state transitions, after watermark is detected and one duplicate state is reached, may be routed via an appropriate state of original STG. This transition may be used to set counter for the first time signature states are traversed and thereafter increasing the counter.

The above discussion indicates that the STG of Fig. 2 can be redrawn as Fig. 3 while maintaining the original functionality intact. The number of states is reduced from 17 states in Fig. 2 to 11 states in Fig. 3. Therefore the proposed scheme reduces hardware overhead caused by watermarking.

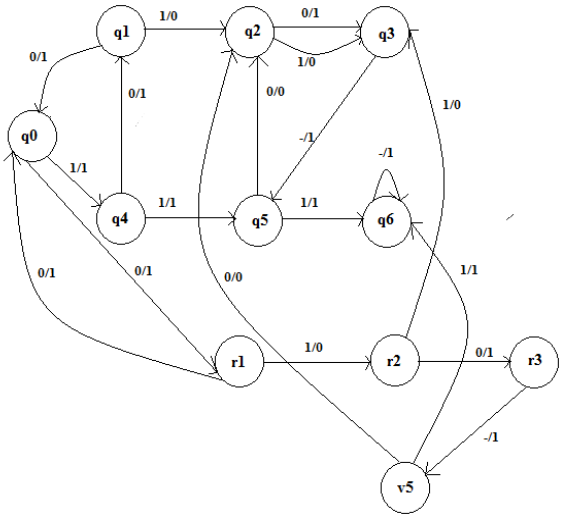


Fig.3 Proposed Reduced State watermarked circuit

Once state v5 is traversed in Figs. 3, there is no input sequence that can lead FSM to go to starting state q0 so, if until a reset is applied, circuit will not traverse ri states again even if signature is applied again and circuit will perform normally. This means that we can test the watermarked circuit only once after application of signature input, until we apply reset signal. This property was present in the original watermarked circuit of Oliviera [5] shown in Fig. 2 also.

But in case the original STG is such that after vi states circuit reaches on q0 state by applying some sequence of input, then there is a possibility to traverse R states again. To avoid this situation, following additional steps may be used:

(i) Use a test signal which becomes high when state vi is reached in the watermarked circuit.
(ii) When state q0 is reached and test signal is high, next state from q0 on application of the first bit of the signature will be q1 instead of r1. So the ri states are not traversed again once state V is reached.
(iii) (iii)This test signal can be disabled again when we apply low reset signal.

A close inspection of STG of Fig.2 reveals that state v5 is traversed after r3 state irrespective of the input sequence. This may not be true, in general, for every STG. To elaborate on this, the STG of a sequence detector [2] is considered. This circuit detects the sequence "11011" and it is an overlapping sequence detector. Fig.5 shows the watermarked circuit for a signature sequence "101". Here states v0 and v2 are traversed after watermarked state r3, both the states are therefore retained in reduced STG. The modified STG is drawn in Fig. 6 where the state transitions from v0 and v2 In this case when we reduce

ISSN (Online) 2321 – 2004
ISSN (Print) 2321 – 5526

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH IN ELECTRICAL, ELECTRONICS, INSTRUMENTATION AND CONTROL ENGINEERING
Vol. 2, Issue 7, July 2014

number of only two vi states are important i.e. v0 & v2. Now, the STG with reduced number of states is shown in Fig. 6.

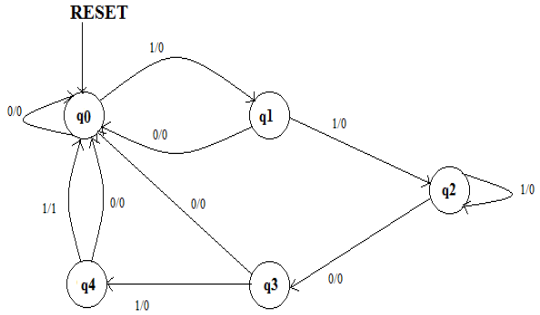


Fig.4 STG of Sequence Detector proposed by Subbaraman [2]

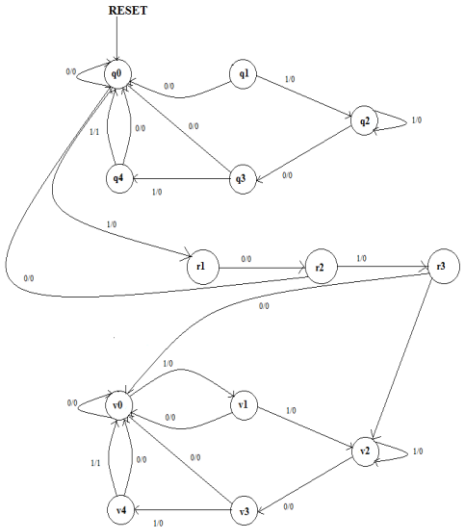


Fig.5 Watermarked STG of sequence Detector by Subbaraman[2]

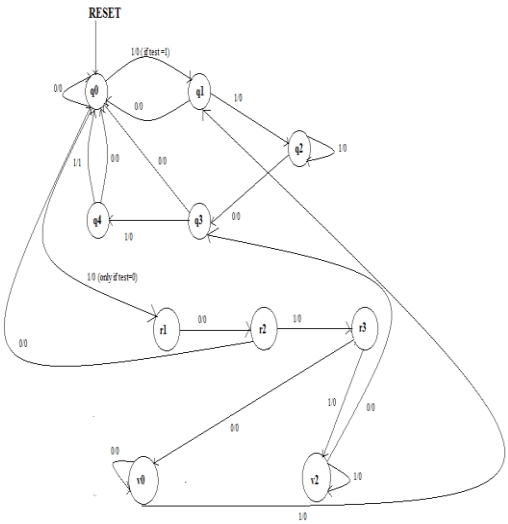


Fig.6 Reduced state watermarked STG of Sequence Detector (Proposed design)

In Fig. 6, when suppose state v0 is reached now, if we apply input '10' circuit will reach on q0 and if accidently signature is applied at this stage the R states will be traversed again, similarly if state v2 is reached after r3 and sequence '00' is applied, circuit will reach on q0 state and there is a possibility of traversing ri states. So, to eliminate this situation a test signal will be used. When state v0 or v2 is reached test signal will be '1' and remain high until reset is applied. Now if present state is q0 and test is '0' then next state will be r1 if input '1' is applied, if test is '1' then next state will be q1 and circuit will not go in STG R even if signature is applied accidently. To reduce the number of vi states, the last part of the signature sequence should be chosen in such a way that there are minimum numbers of transitions to the different vi states from final ri state.

III. RESULTS

Synthesis and Simulations are performed using Xilinx ISE 6.1i EDA tool with its built in synthesis tool IST and Modeltech's Modelsim 5.4a. The simulations and results of non-watermarked, watermarked and watermarked circuit with reduced states are shown in next section.

An input sequence "0101010101010….." which includes signature "010" as first three bits is applied to STG of Figs. 1-3 and the simulation results are depicted in Figs. 7-9 respectively. It may be noted that similar output characteristics are obtained for each case. Only the state transitions differ in watermarked and non- watermarked circuit. In case of non - watermarked circuit, for the given input sequence circuit traverses STG Q, but in case of watermarked circuit of Fig. 2 for the first three bits of input, its traverses STG R, and after that circuit enters in STG V which are replica of original STG and remain in STG V until reset is applied.

State assignment for Fig.1 – q0=000, q1= 001, q2= 010, q3=011, q4=100, q5=101, q6=110

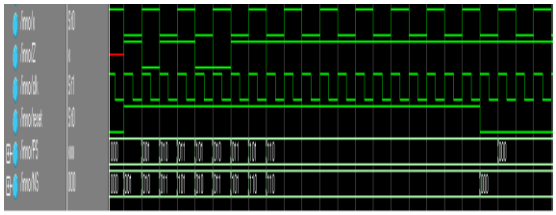


Fig. 7 Simulation of original STG of Fig.1

State assignment for Fig.2 q0=00000, q1=00001,q2=00010,q3=00011,q4=00100,q5=00101,q6=00110 r1= 00111, r2 =01000,r3=01001,v0=01010,v1=01011,v2=01100,v3=01101,v4=01110,v5=01111,v6=10000

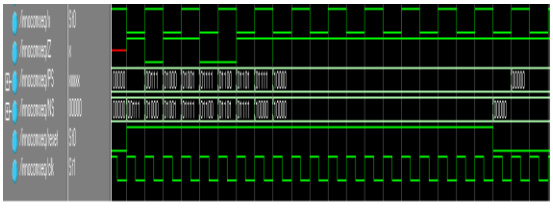


Fig. 8 Simulation of watermarked STG of Fig.2

Sate Assignment for Fig.3 – q0=0000, q1=0001, q2=0010,q3=0011,q4=0100,q5=0101,q6=0110,r1=0111,r2 =1000,r3=1001,v5=1010

ISSN (Online) 2321 – 2004
ISSN (Print) 2321 – 5526

IJIREEICE

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH IN ELECTRICAL, ELECTRONICS, INSTRUMENTATION AND CONTROL ENGINEERING
Vol. 2, Issue 7, July 2014

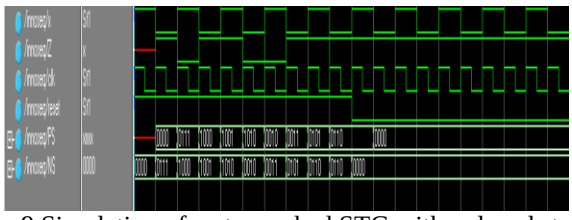


Fig. 9 Simulation of watermarked STG with reduced states of Fig. 3

In simulation of Fig.3, when the signature ”010” is applied the circuit traverses STG R as in the case of Fig.2. After that it enters the next vi state(v5) for any input applied to r3. Now, if the input is applied when present state is v5, the next state will be one of the qi states and then circuit remains in qi states until reset signal is applied.

Similarly the simulations of sequence detector of Figs. 4-6 are shown in Figs. 10-12.

State assignments for fig.4- q0=000,q1=001,q2=010,q3=011,q4=100

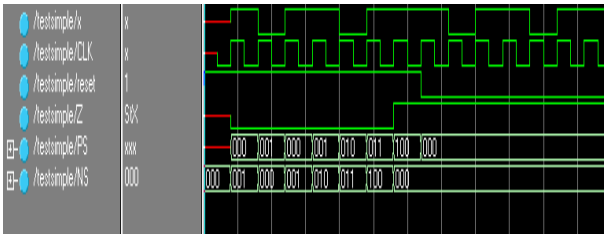


Fig. 10 Simulation of original sequence detector

State assignments for Fig.5 – q0=0000,q1=0001,q2=0010,q3=0011,q4=0100,r1=0101,r2 =0110,r3=0111,v0=1000,v1=1001,v2=1010 v3=1011, v4=1100

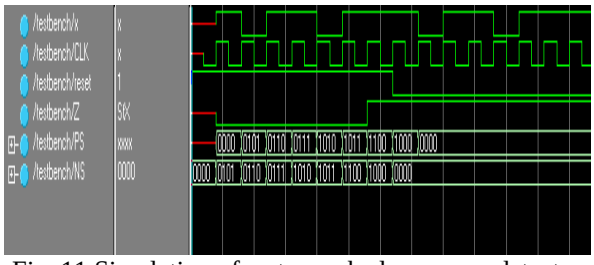


Fig. 11 Simulation of watermarked sequence detector

State assignments for Fig.6
q0= 0000, q1=0001, q2=0010, q3=0011, q4=0100,r1=0101,r2=0110,r3=0111,v0=1000, v2=1001

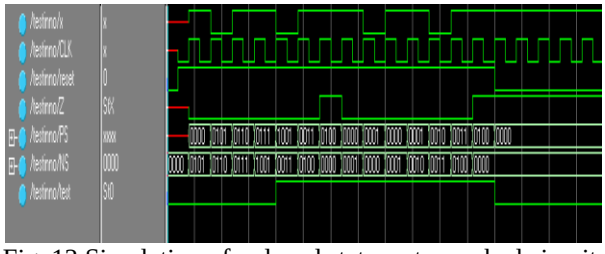


Fig. 12 Simulation of reduced state watermarked circuit.

In Fig. 12 it may be noted that, when test signal became high, circuit reaches in state v0. If test signal is high and q0 state is reached the next state from q0 with input ‘1’ is q1 not r1. So this proves the functionality of reduced state watermarked circuit explained in previous section.

TABLE I
SYNTHESIS REPORT OF STG USED BY OLIVIERA[5]

Details	Original Design (fig.1)[1]		Watermarked Design(fig. 2)[1]		Proposed Watermarked Design (fig. 3)	
IOs	4		4		4	
Flip Flops/Latches	FDC	8	FDC	15	FDC	12
Cell Usage	BELS	11	BELS	20	BELS	18
Clock Buffers	BUFGP	2	BUFGP	2	BUFGP	2
IO Buffers	IBUF	1	IBUF	1	IBUF	1
	OBUF	1	OBUF	1	OBUF	1

TABLE II
SYNTHESIS REPORT OF SEQUENCE DETECTOR IMPLEMENTED BY SUBBARAMAN[2]

Details	Original Design (fig.4)[6]		Watermarked Design(fig. 5) [6]		Proposed Watermarked Design (fig. 6)	
IOs	4		4		4	
Flip Flops/Latches	FDC	6	FDC	13	FDC	12
Cell Usage	BELS	9	BELS	21	BELS	19
Clock Buffers	BUFGP	2	BUFGP	2	BUFGP	2
IO Buffers	IBUF	1	IBUF	1	IBUF	1
	OBUF	1	OBUF	1	OBUF	1

The synthesis report shows that no. of latches/flip flops are reduces in case of Fig.3 as compared to Fig.2. The hardware requirement of watermarked circuit is much more as compared to non- watermarked circuit. The timing report for STG of Figs. 1-3 and Figs. 4-6 are respectively summarized in Tables 3 and 4.

TABLE III
TIMING SUMMARY OF STG USED BY OLIVIERA[5]

	Original Design (fig. 4)	Watermarked Design(fig. 5)	Watermarked Design (fig. 6)
Maximum Frequency	264.340 MHz	264.340 MHz	264.340 MHz

TABLE IV
TIMING SUMMARY OF SEQUENCE DETECTOR

	Original Design	Watermarked Design(fig. 2)	Watermarked Design (fig. 3)
Maximum Frequency	231.267 MHz	231.267 MHz	231.267 MHz

The maximum operating frequency in the reduced state watermarked circuit is same as that of original STG. This is an important result as this helps in maintaining the secrecy of watermarked design.

IV. DETECTION OF WATERMARK

The detection of watermark can be done using a Counter circuit as suggested by Subbararaman[2]. When the state

ISSN (Online) 2321 – 2004
ISSN (Print) 2321 – 5526

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH IN ELECTRICAL, ELECTRONICS, INSTRUMENTATION AND CONTROL ENGINEERING
Vol. 2, Issue 7, July 2014

transition occurs in STG Q count value becomes "00", as the signature sequence is applied count value increases for each state transition through ri states. After final ri state the next state is a vi state and counter retains its previous value indicating that whole signature sequence has been traversed.

V. CONCLUSION

Simulation and Synthesis results of the above method proved that there is no change in output response of the original, watermarked circuit as used by Oliviera[5] and Subbaraman[2] and the proposed reduced state watermarked circuit. But the number of states is considerably reduced in case of reduced state watermarked circuits thereby reducing the hardware overhead due to watermarking. Also the maximum operating frequency is same in all the cases. So we should explore two possibilities for state reduction:-

Firstly, the signature sequence should be chosen in a way that there is minimum number of transitions to vi states after final state of STG R. Secondly, the vi states to be copied from STG Q should be chosen in such a way that there is no possible state transition to initial state of STG Q, when we apply further input sequence after the arrival of vi states.

REFERENCES

[1] Arlindo L. Oliveira, "Techniques for the Creation of Digital Watermarks in Sequential Circuit Design," IEEE Transactions on Computer-aided Design of Integrated Circuits and Systems, Vol. 20, No. 9, September 2001.

[2] ShailaSubbaraman , and P. S. Nandgawe, "Intellectual Property Protection of Sequential Circuits Using Digital Watermarking", First International Conference on Industrial and Information Systems, ICIIS 2006, 8 - 11 August 2006, Sri Lanka

[3] J. Toruoglu, and E. Charbon, "Watermarking Based Copyright Protection of Sequential Functions", IEEE journal of Solid State Circuits, Vol. 35, No. 3,February 2000, pp434-440.

[4] Amr T. Abdel-Hamid, Sofi'eneTahar and El MostaphaAboulhamid, "IP Watermarking Techniques- Survey and Comparison ", Proceedings of The 3rd IEEE International Workshop on System-on-Chip for Real-Time Applications, ISBN 0- 7695-1929-6103 copyright 2003 IEEE.

[5] Arlindo L. Oliveira, "Robust Rechniques for WatermarlingSequential Circuit Designs", DAC 99, New Orleans, Louisiana c 1999 ACM 1-58113-109-7/99/06.

[6] M. Lewandowski, R. Meana, M. Morrison and S. Katkoori, "A Novel Method for Watermarking Sequential Circuits", International Symposium on Hardware-Oriented Security and Trust, 2012 IEEE.

BIOGRAPHIES

Jeebananda Panda Born on 15 th Feb, 1968 in Odisha , India . He got graduated in Electrical engineering and subsequently in electronics and Communication engineering in 1988 and 1989 respectively. He got his M.E degree in Applied Electronics specialization from Bharathiyar University in 1992. Presently working as an Associate Professor in the Department of Electronics and Communication Engineering, Delhi Technological University, Delhi, India. He is presently working for Ph.D in Dept. of Electronics in Engineering, Faculty of Technology, Delhi University. His field of research is Watermaking of Digital Data.

Ankur Bharadwaj Born on 04 January, 1990 in Meerut, Uttar Pradesh, India. Completed M.Tech in VLSI Design and Embedded Technology in year 2013 from Delhi Technological University, Delhi, India. Currently Ph.D from Jaypee Institute of Information Technology, Noida, India. Working as an Assistant Professor in Jaypee Institute of Information Technology, Noida since 2013.

Neeta Pandey Completed her M.E. in Microelectronics from Birla Institute of Technology and Sciences, Pilani and Ph.D. from Guru Gobind Singh Indraprastha University Delhi. She has served in Central Electronics Engineering Research Institute, Pilani, Indian Institute of Technology, Delhi, Priyadarshini College of Computer Science, Noida and Bharati Vidyapeeth's College of Engineering, Delhi in various capacities. At present, she is working as an Assistant Professor in ECE Department, Delhi Technological University. Her research interests are in analog and digital VLSI Design.

Asok Bhattacharyya Obtained M.Tech and Ph.D. degree from Institute of Radio Physics, Calcutta University, India in 1970 and 1981 respectively. He joined Delhi College of Engineering in May 1974. He has retired as Professor in E&C Engineering. He has worked in different fields- Digital System Design, Analog System Design, Easily testable and diagnosable Digital systems/ Fault tolerant Computing and Medical Image Processing area. He has authored two research monographs.

Optik xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

Image enhancement via Median-Mean Based Sub-Image-Clipped Histogram Equalization

Kuldeep Singh [a,*], Rajiv Kapoor [b]

[a] Central Research Lab, Bharat Electronics Ltd, Ghaziabad 201010, India
[b] Department of Electronics & Communication, Delhi Technological University, Delhi 110042, India

ARTICLE INFO

Article history:
Received 18 September 2013
Accepted 24 April 2014
Available online xxx

Keywords:
Histogram equalization
Image quality measure
Image information content
Bi-histogram equalization

ABSTRACT

This paper presents a robust contrast enhancement algorithm based on histogram equalization methods named Median-Mean Based Sub-Image-Clipped Histogram Equalization (MMSICHE). The proposed algorithm undergoes three steps: (i) The Median and Mean brightness values of the image are calculated. (ii) The histogram is clipped using a plateau limit set as the median of the occupied intensity. (iii) The clipped histogram is first bisected based on median intensity then further divided into four sub images based on individual mean intensity, subsequently performing histogram equalization for each sub image. This method achieves multi objective of preserving brightness as well as image information content (entropy) along with control over enhancement rate, which in turn suits for consumer electronics applications. This method avoids excessive enhancement and produces images with natural enhancement. The simulation results show that MMSICHE method outperforms other HE methods in terms of various image quality measures, i.e. average luminance, average information content (entropy), absolute mean brightness error (AMBE) and background gray level.

1. Introduction

Contrast enhancement and brightness preservation are two prime focus areas of researchers in the field of consumer electronics products. Nowadays mobile phones have been widely used to take pictures in daily life. Mobile phones have limited hardware capability for digital photography. In this scenario, post-processing using software tools is needed to improve the quality of the acquired image. Histogram equalization (HE) is most extensively utilized contrast enhancement technique due to its simplicity and ease of implementation [1]. The idea behind HE is to flatten the probability distribution and stretching the dynamic range of gray levels, which in result improves the overall contrast of the image [2]. HE utilizes the cumulative density function (CDF) of image for mapping the gray levels of original image to the enhanced image. HE is not suitable for most consumer electronics applications such as TV, cameras etc., as it tends to change the mean brightness of the image to the middle level of the gray level range, which in turn produces annoying artifacts and intensity saturation effects. Various methods [2–6] have been suggested in literature to overcome the shortcomings in histogram equalization method. Kim [2] in 1997 was the first one to propose an algorithm named brightness preserving bi histogram equalization (BBHE) which preserves the mean brightness of the image and improves the contrast.

BBHE bisects the histogram based on the input mean brightness and equalizes the two sub histograms independently. In 1999 Wan et al. [3] proposed an algorithm named dualistic sub image histogram equalization (DSIHE) and claimed that it is better than BBHE in terms of preservation of brightness and information content (entropy) of the image. DSIHE separates the histogram based on median value instead of mean, which implies that each histogram contains almost equal number of pixels. Chen and Ramli introduced minimum mean brightness error bi-histogram equalization (MMBEBHE) for preserving the mean brightness "optimally" [4]. This method is an extension of BBHE, which iteratively calculates the absolute mean brightness error (AMBE) for gray levels 0 to $L-1$ and bisects the histogram based on the intensity value X_m, which yields minimum AMBE.

Chen and Ramli proposed another approach named recursive mean-separate histogram equalization (RMSHE) [5]. This technique iteratively performs the BBHE in which the histogram is divided into two parts based on the average input brightness and BBHE is performed to each sub histogram independently. Sim et al. [6] proposed a similar technique to RMSHE known as recursive sub-image histogram equalization (RSIHE). This technique performs

* Corresponding author. Tel.: +91 9910101592.
E-mail addresses: kuldeep.er@gmail.com (K. Singh), rajivkapoor@dce.ac.in (R. Kapoor).

http://dx.doi.org/10.1016/j.ijleo.2014.04.093
0030-4026/

the division of histogram based on the median value of brightness instead of mean brightness. Finding the optimal value of iteration factor is a big challenge for producing significant enhancement results. Kim and Chung [7] presented a method named Recursively Separated and Weighted Histogram Equalization (RSWHE) similar to RSIHE and RMSHE methods. In addition, RSWHE modify the histogram by weighting process using normalized power law function. These techniques do not provide a mechanism for adjusting the level of enhancement. Clipped histogram based techniques [8–10] were proposed as a solution for controlling the enhancement rate as well as preserving the original brightness. These methods control maximum value of the histogram by clipping histograms higher than the prespecified threshold. Different approaches are proposed for the determination of clipping threshold or plateau limit.

Abdullah-Al-Wadud et al. [11] proposed a new class of histogram partitioning named as dynamic histogram equalization (DHE). The DHE partitions the original histogram based on local minima and then, assigns a new dynamic range to each sub-histogram. The major drawback of this method is that it remaps the histogram peaks by allocating new dynamic range, which significantly changes the mean brightness. Ibrahim and Kong [12] proposed a method brightness preserving dynamic histogram equalization (BPDHE) similar to DHE, which is extension of the DHE. BPDHE applies Gaussian-smoothing filter before the histogram partitioning process is carried out. The BPDHE uses the local maxima as the separating point rather than the local minima used in DHE. Ibrahim and Kang claimed that the local maxima are better for mean brightness preservation. Sheet et al. [13] proposed a modification of the BPDHE technique named as Brightness Preserving Dynamic Fuzzy Histogram Equalization (BPDFHE). This method uses Fuzzy histogram computation for smoothing operation of histogram before partitioning of image into sub histograms. The authors of BPDFHE method proved the superiority of the algorithm in terms of less computational time and brightness preservation. These dynamic methods are suitable only for the images having significant peaks in the histograms. Quadrants Dynamic Histogram Equalization (QDHE) method was proposed by Ooi and Isa [14] for better contrast enhancement. QDHE partitions the histogram into four sub-histograms using the median value of intensity and then clips the histogram according to the mean of intensity occurrence of the input image and finally a new dynamic range is assigned to each sub-histogram before each sub-histogram is equalized. Ooi and Isa [15] also proposed Adaptive Contrast Enhancement Methods with Brightness Preserving which comprised of two methods named as dynamic quadrants histogram equalization plateau limit (DQHEPL) and bi-histogram equalization median plateau limit (BHEPL-D). DQHEPL is extension of RSIHE, divides the histogram into four sub histograms, and then assigns a new dynamic range and finally implements clipping process. BHEPL-D is the extension of the BHEPL except that it clips the histogram using the median of the occupied intensity. Chang and Chang [16] presented a simple approach for contrast enhancement named as Simple Histogram Modification Scheme (SHMS). This method modifies the histogram by changing the values of two boundary values of the support of the histogram.

A new method named as background brightness preserving histogram equalization (BBPHE) [17] was proposed by Tan et al. The partition method used by BBPHE is based on background levels and non-background levels range. After partition, each sub-image is equalized independently, and then combined into the final output image. It is claimed that the background levels are only stretched within the original range, hence, the over enhancement can be avoided by BBPHE. Singh and Kapoor [19] proposed Exposure based Sub Image Histogram Equalization (ESIHE) method for enhancement of low exposure images where the image exposure threshold is used for sub dividing image.

Although various techniques were proposed to cater specific problem of contrast enhancement a new robust algorithm is being proposed here which addresses problems of preservation of mean brightness, entropy and control on the enhancement rate simultaneously. This algorithm also puts emphasis on natural enhancement of images. The authors believe that the MMSICHE technique that achieves the multiple objectives of entropy maximization, brightness preservation and control over enhancement is a better approach to natural image enhancement.

This paper is organized as follows: Section 2 describes the proposed MMSICHE method. Section 3 gives experimental results, and Section 4 concludes the paper.

2. Median-Mean Based Sub-Image-Clipped Histogram Equalization

This section, presents the algorithm of MMSICHE. The algorithm consists of three steps, namely median and mean calculation, Histogram Clipping and Histogram Subdivision & Equalization. Following subsections present description of each step of the algorithm.

2.1. Median and mean calculation

The median of the image is denoted as an intensity value X_e where the cumulative density function is 0.5 [3]. Two mean intensity values (X_{ml} and X_{mu}) are calculated for two individual sub histogram divided based on median value. The values of X_e, X_{ml} and X_{mu} are calculated before histogram clipping process. Eq. (1) computes the total number of samples N for a given image

$$N = \sum_{k=0}^{L-1} h(k) \tag{1}$$

$h(k)$ is histogram of image and L is total number of gray levels. For calculating X_e consider a variable $z(k)$ as computed in (2)

$$z(k) = z(k-1) + h(k) \quad \text{for} \quad k = 0, 1, \ldots, L-1 \quad \text{and} \quad z(0) = h(0) \tag{2}$$

Median variable can be calculated as per (3) using (1) and (2)

$$X_e = k \quad \text{where} \quad z(k) \geq \frac{N}{2} \tag{3}$$

Eqs. (4) and (5) expresses the calculation of mean variables X_{ml} and X_{mu}

$$X_{ml} = \sum_{k=0}^{X_e-1} P_l(k) \times k \tag{4}$$

$$X_{mu} = \sum_{k=X_e}^{L-1} P_u(k) \times k \tag{5}$$

where $P_l(k)$ and $P_u(k)$ are individual PDF of two sub histograms divided based on median. Eqs. (6) and (7) depict the calculation of these PDFs

$$P_l(k) = \frac{h(k)}{N_l} \quad \text{for} \quad k = 0, 1, \ldots, X_e - 1 \tag{6}$$

$$P_u(k) = \frac{h(k)}{N_u} \quad \text{for} \quad k = X_e, X_e + 1, \ldots, L-1 \tag{7}$$

N_l and N_u are total numbers of pixels in the lower and upper histogram respectively.

Please cite this article in press as: K. Singh, R. Kapoor, Image enhancement via Median-Mean Based Sub-Image-Clipped Histogram Equalization, Optik - Int. J. Light Electron Opt. (2014), http://dx.doi.org/10.1016/j.ijleo.2014.04.093

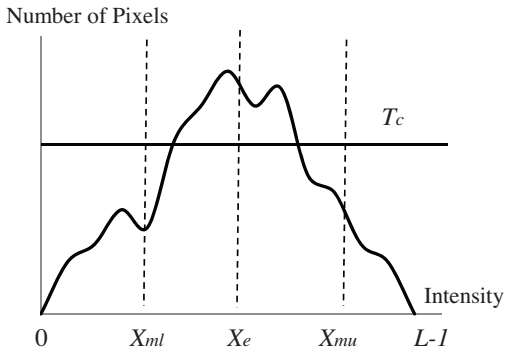


Fig. 1. Process of histogram clipping.

2.2. Histogram clipping

The idea behind the histogram clipping is to control the enhancement rate consequently to result in a natural appearance of the image. For limiting the enhancement rate, we need to limit the first derivative of the histogram or the histogram itself [9].

The histogram bins having the value greater than the clipping threshold are limited to the threshold. The clipping threshold is calculated as the median of occupied intensity. The formula for clipping threshold T_c is presented in (8), and (9) calculates the clipped histogram

$$T_c = median[h(k)] \tag{8}$$

$$h_c(k) = T_c \quad \text{for} \quad h(k) \geq T_c \tag{9}$$

$h(k)$ and $h_c(k)$ are the original and clipped histogram. This method is computationally efficient and less time consuming. The clipping process is graphically illustrated in Fig. 1.

2.3. Histogram sub division and equalization

The original histogram is first bisected based on median intensity value X_e as calculated in (3). These individual sub histograms are further divided into two small sub histograms where the individual mean X_{ml} and X_{mu} as calculated in (4) and (5) acts as separating point of sub histograms. The Histogram Sub Division process results in four sub images W_{Ll}, W_{Lu}, W_{Ul} and W_{Uu} ranging from gray level 0 to X_{ml}, $X_{ml}+1$ to X_e, X_e+1 to X_{mu} and $X_{mu}+1$ to $L-1$. $P_{Ll}(k)$, $P_{Lu}(k)$, $P_{Ul}(k)$ and $P_{Uu}(k)$ are corresponding PDF of these sub images as defined in Eqs. (10)–(13)

$$P_{Ll}(k) = \frac{h_c(k)}{N_{Ll}} \quad \text{for} \quad 0 \leq k \leq X_{ml} \tag{10}$$

$$P_{Lu}(k) = \frac{h_c(k)}{N_{Lu}} \quad \text{for} \quad X_{ml}+1 \leq k \leq X_e \tag{11}$$

$$P_{Ul}(k) = \frac{h_c(k)}{N_{Ul}} \quad \text{for} \quad X_e+1 \leq k \leq X_{mu} \tag{12}$$

$$P_{Uu}(k) = \frac{h_c(k)}{N_{Uu}} \quad \text{for} \, X_{mu}+1 \leq k \leq L-1 \tag{13}$$

N_{Ll}, N_{Lu}, N_{Ul} and N_{Uu} are total number of pixels in sub images W_{Ll}, W_{Lu}, W_{Ul} and W_{Uu} respectively. $C_{Ll}(k)$, $C_{Lu}(k)$, $C_{Ul}(k)$ and $C_{Uu}(k)$ are corresponding CDF of individual sub images and CDFs can be defined as Eqs. (14)–(17)

$$C_{Ll}(k) = \sum_{k=0}^{X_{ml}} P_{Ll}(k) \tag{14}$$

$$C_{Lu}(k) = \sum_{k=X_{ml}+1}^{X_e} P_{Lu}(k) \tag{15}$$

$$C_{Ul}(k) = \sum_{k=X_e+1}^{X_{mu}} P_{Ul}(k) \tag{16}$$

$$C_{Uu}(k) = \sum_{k=X_e+1}^{L-1} P_{Uu}(k) \tag{17}$$

The next step of MMSICHE is to equalize all the four sub histograms individually. The transfer functions for histogram equalization based on Eqs. (10)–(17) can be defined as Eqs. (18)–(21)

$$F_{Ll} = X_{ml} \times C_{Ll} \tag{18}$$

$$F_{Lu} = (X_{ml}+1) + (X_e - X_{ml}+1) \times C_{Lu} \tag{19}$$

$$F_{Ul} = (X_e+1) + (X_{mu} - X_e+1) \times C_{Ul} \tag{20}$$

$$F_{Uu} = (X_{mu}+1) + (L - X_{mu}+1) \times C_{Uu} \tag{21}$$

F_{Ll}, F_{Lu}, F_{Ul} and F_{Uu} are the transfer functions used for equalizing the sub histograms individually. The final step involves the integration of all sub images into one complete image for analysis. The MMSICHE-ed output image is produced by the combination of all four transfer functions.

3. Simulation results

In this section, the simulation results of proposed method MMSICHE are compared with existing histogram equalization based methods, i.e. WTHE, QDHE, RSWHE, SHMS, BHEPL and BHEPL-D. In order to analyze and compare the existing methods we use five test images: *Tank*, *U2*, *Field*, *Copter* and *Hands*. To evaluate the performance of MMSICHE, four image quality measures (IQM) are chosen, i.e. average luminance, absolute mean brightness error (AMBE), average information content (entropy) and background gray level (BGL) [18]. Average luminance μ is the mean gray level or brightness of the image. For better results, average luminance should be close to that of the original image. Absolute mean brightness error (AMBE) is the absolute difference between the input and output mean brightness. The formula for calculation of μ and AMBE is expressed in (22) and (23). X_μ is the mean value of the input image, while Y_μ is the mean value of the output image.

$$\mu = \sum_{l=0}^{L-1} l \times P(l) \tag{22}$$

$$\text{AMBE} = |X_\mu - Y_\mu| \tag{23}$$

AMBE measure is measure of excessive brightness change and directly related to average luminance. For best performance, the AMBE value should be as low as possible for better performance. Average information content (entropy) is a measure of richness of the details of the image. The larger value of the entropy indicates that more information content is available in the images. Eq. (24) defines Entropy $Ent(p)$

$$Ent(p) = -\sum_{l=0}^{L-1} P(l) \log_2 P(l) \tag{24}$$

Background gray level is the quality measure for natural enhancement. The BGL value close to the original image guarantees natural enhancement.

Please cite this article in press as: K. Singh, R. Kapoor, Image enhancement via Median-Mean Based Sub-Image-Clipped Histogram Equalization, Optik - Int. J. Light Electron Opt. (2014), http://dx.doi.org/10.1016/j.ijleo.2014.04.093

Table 1
Average luminance.

Images	Original	WTHE	QDHE	RSWHE-M	SHMS	BHEPL	BHEPL-D	MMSICHE
Tank	132.38	212.49	162.65	124.61	131.81	154.35	33.31	134.54
U2	32.51	108.28	57.78	26.80	131.33	35.17	33.79	37.16
Field	106.39	132.81	140.17	95.22	129.02	120.29	33.25	112.76
Copter	219.46	180.22	217.31	205.43	130.52	208.73	215.62	215.96
Hands	49.13	172.07	60.58	42.44	172.47	46.78	44.07	52.21
Average	107.97	161.17	127.69	98.90	139.03	113.06	72.00	110.52

Table 2
Absolute mean brightness error (AMBE).

Images	WTHE	QDHE	RSWHE-M	SHMS	BHEPL	BHEPL-D	MMSICHE
Tank	80.10	30.27	7.77	0.57	21.96	99.07	2.16
U2	75.77	25.27	5.72	98.82	2.66	1.28	4.64
Field	26.42	33.78	11.17	22.63	13.89	73.14	6.37
Copter	39.24	2.15	6.97	88.94	10.74	3.84	3.50
Hands	122.94	11.45	6.69	123.34	2.35	5.06	3.08
Average	68.89	20.58	7.66	66.86	10.32	36.47	3.95

Table 3
Average information content (entropy).

Images	Original	WTHE	QDHE	RSWHE-M	SHMS	BHEPL	BHEPL-D	MMSICHE
Tank	5.49	4.65	5.48	4.65	5.38	5.47	2.91	5.50
U2	5.64	4.07	5.44	4.07	5.41	5.61	5.60	5.55
Field	6.56	5.28	6.53	5.28	6.50	6.53	3.86	6.56
Copter	6.31	3.75	6.24	3.75	6.11	6.21	6.19	6.22
Hands	3.99	3.20	3.91	3.20	3.60	3.96	3.97	3.98
Average	5.60	4.19	5.52	4.19	5.40	5.55	4.51	5.56

3.1. Performance assessment based on image quality measures

Tables 1–4 show a matrix of all four IQMs for 5 test images, where rows represent the test images and the columns represent various methods used for comparison. Average luminance and AMBE are two measures, which reflect the brightness preservation capability. Table 1 shows average luminance measures for all the images using various methods. Table 1 reveals that the average of all images is best for MMSICHE method in terms of its closeness to the average of original image The average of MMSIHE is 107.97 close to the average of original image (110.52). Specifically for Copter, Hands and Field image, the MMSICHE method produces the images having average luminance very close to the original image. Table 2 shows the matrix of AMBE for all the images using various methods. Similar to average luminance MMSICHE provides best results in terms of AMBE among all methods. MMSICHE algorithm has least AMBE value for *Copter*, *Hands* and *Field* image and further as shown in the last row the average of AMBE for all images is 3.95, which is significantly less than all other techniques.

Better results in terms of average luminance and AMBE reflect that MMSICHE is well suited for mean brightness preservation required for consumer electronics devices. Table 3 presents the entropy measures for all the techniques used in this work. MMSICHE produces highest or close to the highest entropy for all the five images thus becomes best suitable approach for bringing out information contents of the images. Specifically for *Tank*, *Field* and *Hands* images the entropy values are almost equal to original image. The average of entropy produced by MMSICHE method for all images is 5.56 that is very close to average entropy (5.60) for original images, however average entropy of other methods is very less in comparison with the original image. The entropy closer to original image guarantees natural enhancement. Background gray level results shown in Table 4 significantly outperform other techniques for all the images. The BGL values from MMSICHE method for most images are very close to the original image and ensure the natural enhancement. However, the MMSICHE method's BGL average for all the images is exactly equal to the original image.

3.2. Assessment of visual quality & natural appearance

Qualitative assessment of contrast enhancement is necessary along with quantitative assessment based on IQMs. By visual quality inspection the judgment of annoying artifacts, over enhancement and unnatural enhancement can be done. Wide varieties of standard images ranging from low contrast to high contrast, dark background to bright background are chosen to show the robustness of new method.

Table 4
Background gray level (BGL).

Images	Original	WTHE	QDHE	RSWHE-M	SHMS	BHEPL	BHEPL-D	MMSICHE
Tank	147	134	195	134	175	186	0	148
U2	22	23	20	23	101	16	16	27
Field	116	108	163	108	154	133	0	118
Copter	234	221	226	221	185	238	239	228
Hands	23	24	24	24	159	12	10	23
Average	108	102	125	102	154	117	53	108

Please cite this article in press as: K. Singh, R. Kapoor, Image enhancement via Median-Mean Based Sub-Image-Clipped Histogram Equalization, Optik - Int. J. Light Electron Opt. (2014), http://dx.doi.org/10.1016/j.ijleo.2014.04.093

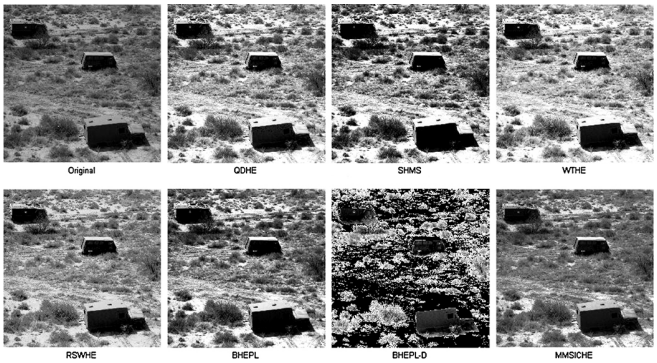


Fig. 2. Enhancement results of *Field* image.

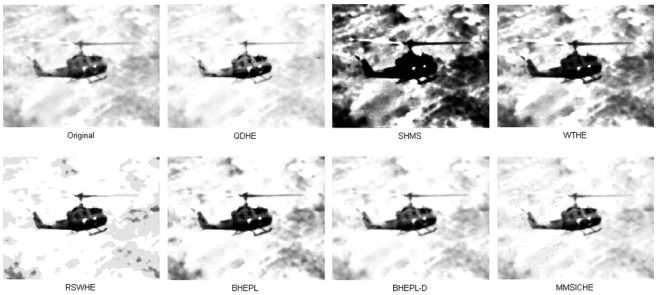


Fig. 3. Enhancement results of *Copter* image.

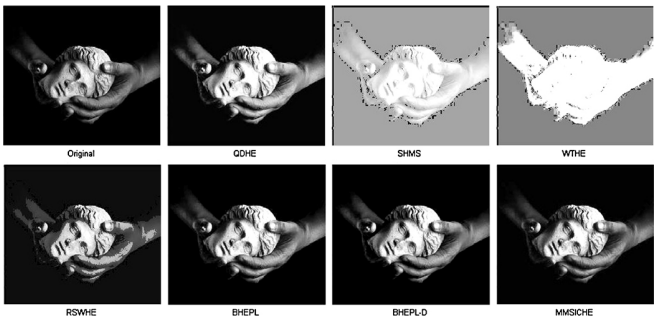


Fig. 4. Enhancement results of *Hands* image.

The analysis of visual results from Figs. 2–6 shows that the MMSICHE method produces the best results for all the images in terms of control on over enhancement as well as natural appearance. QDHE, WTHE and SHMS methods have produced over enhanced images in Fig. 2 of Field image. However, the BHEPL-D method has introduced noise and produced annoying artifacts.

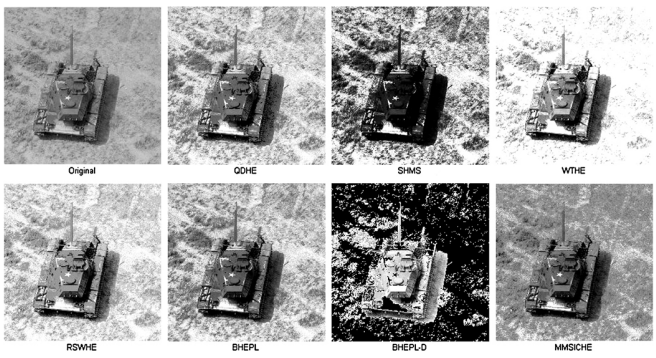


Fig. 5. Enhancement results of *Tank* image.

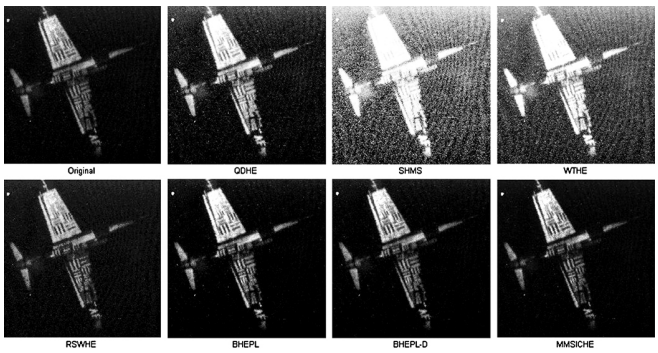


Fig. 6. Enhancement results of *U2* image.

The acceptable and natural enhanced images are produced by the BHEPL, RSWHE and MMSICHE. Although the MMSICHE results in Fig. 2 for *Field* image are visually comparable to other methods but it yields the highest entropy value almost equal to original image.

The Copter image results in Fig. 3 can be analyzed by viewing the copter object. The object has been darkened by WTHE and SHMS method and overall contrast has been reduced by these methods. MMSICHE method has produced image with similar results of other methods. Similar to Copter image both WTHE and SHMS has deteriorated the Hands image in Fig. 4 badly. RSWHE has produced over enhanced image particularly on the finger and thumb portion. MMSICHE-ed *Hands* image is naturally enhanced image. The concrete results in terms of contrast enhancement can be clearly observed in Fig. 5 of Tank image. The original image is a low contrast image and MMSICHE yields contrast enhanced image along with natural appearance.

Over-enhancement and intensity saturation phenomena have been produced by RSWHE and WTHE. SHMS output do not provide clear vision of object in case of *Tank* image and BHEPL-D has deteriorated the image. The results in Fig. 6 of *U2* image clearly show the superiority of MMSICHE method over other methods. It is clearly noticeable that the outputs of SHMS, WTHE and QDHE has noise amplification, however the MMSICHE-ed image is free from noise as well as provides good contrast enhancement.

3.3. Summary of assessment and discussion

From the assessment of all the IQMs and visual inspection, it can be concluded that: (i) MMSICHE method is well suited for brightness preservation in comparison to other methods. (ii) MMSICHE technique is best among other methods in terms of bringing out the information content, i.e. provides highest entropy. (iii) MMSICHE produces images with contrast enhancement and with natural appearances. The objective of this paper is to maximize entropy, preserve original brightness and control the over enhancement. The important factors for accomplishing the objective can be summarized as: (i) Bisecting the image based on the median value plays the role in maximizing entropy and natural enhancement. (ii) Subdividing the image into four sub images based on mean intensity played the role for brightness preservation. (iii) Histogram clipping approach provided the feature of control on over enhancement. The combination of all three above-mentioned process termed as MMSICHE meets the objective of the paper and produces images, which are, not only quantitatively better but also better in terms of quality in comparison to other methods.

4. Conclusion

In this paper, HE based technique MMSICHE has been proposed for preservation of brightness and entropy along with control on

Please cite this article in press as: K. Singh, R. Kapoor, Image enhancement via Median-Mean Based Sub-Image-Clipped Histogram Equalization, Optik - Int. J. Light Electron Opt. (2014), http://dx.doi.org/10.1016/j.ijleo.2014.04.093

over enhancement. In this method clipping of histogram provides control on over enhancement. The bisection of histogram first based on median then based on mean plays the role of preservation of mean brightness and entropy simultaneously. Simulation results clearly show that MMSICHE outperforms other HE based methods based on image quality measures. Visual quality assessment results also prove the supremacy of MMSICHE over other methods in terms of natural enhancement.

References

[1] R.C. Gonzalez, R.E. Woods, Digital Image Processing, 2nd ed., Prentice-Hall, Englewood Cliffs, NJ, 2002.
[2] Y.T. Kim, Contrast enhancement using brightness preserving bi-histogram equalization, IEEE Trans. Consum. Electron. 43 (1997) 1–8.
[3] Y. Wan, Q. Chen, B.M. Zhang, Image enhancement based on equal area dualistic sub-image histogram equalization method, IEEE Trans. Consum. Electron. 45 (1999) 68–75.
[4] S.D. Chen, A.R. Ramli, Minimum mean brightness error bi-histogram equalization in contrast enhancement, IEEE Trans. Consum. Electron. 49 (2003) 1310–1319.
[5] S.D. Chen, A.R. Ramli, Contrast enhancement using recursive mean-separate histogram equalization for scalable brightness preservation, IEEE Trans. Consum. Electron. 49 (2003) 1301–1309.
[6] K.S. Sim, C.P. Tso, Y.Y. Tan, Recursive sub-image histogram equalization applied to gray scale images, Pattern Recogn. Lett. 28 (2007) 1209–1221.
[7] M. Kim, M.G. Chung, Recursively separated and weighted histogram equalization for brightness preservation and contrast enhancement, IEEE Trans. Consum. Electron. 54 (2008) 1389–1397.
[8] Q. Wang, R.K. Ward, Fast image/video contrast enhancement based on weighted thresholded histogram equalization, IEEE Trans. Consum. Electron. 53 (2007) 757–764.
[9] T. Kim, J. Paik, Adaptive contrast enhancement using gain-controllable clipped histogram equalization, IEEE Trans. Consum. Electron. 54 (2008) 1803–1810.
[10] C.H. Ooi, N.S.P. Kong, H. Ibrahim, Bi-histogram with a plateau limit for digital image enhancement, IEEE Trans. Consum. Electron. 55 (2009) 2072–2080.
[11] M. Abdullah-Al-Wadud, et al., A dynamic histogram equalization for image contrast enhancement, IEEE Trans. Consum. Electron. 53 (2007) 593–600.
[12] H. Ibrahim, N.S.P. Kong, Brightness preserving dynamic histogram equalization for image contrast enhancement, IEEE Trans. Consum. Electron. 53 (2007) 1752–1758.
[13] D. Sheet, H. Garud, A. Suveer, M. Mahadevappa, J. Chatterjee, Brightness preserving dynamic fuzzy histogram equalization, IEEE Trans. Consum. Electron. 56 (2010) 2475–2480.
[14] C.H. Ooi, N.A.M. Isa, Quadrants dynamic histogram equalization for contrast enhancement, IEEE Trans. Consum. Electron. 56 (2010) 2552–2559.
[15] C.H. Ooi, N.A.M. Isa, Adaptive contrast enhancement methods with brightness preserving, IEEE Trans. Consum. Electron. 56 (2010) 2543–2551.
[16] Y.C. Chang, C.M. Chang, A simple histogram modification scheme for contrast enhancement, IEEE Trans. Consum. Electron. 56 (2010) 737–742.
[17] T.L. Tan, K.S. Sim, C.P. Tso, Image enhancement using background brightness preserving histogram equalization, Electron. Lett. 48 (2012) 155–157.
[18] S.D. Chen, A new image quality measure for assessment of histogram equalization-based contrast enhancement, Digit. Signal Process. 22 (2012) 640–647.
[19] K. Singh, R. Kapoor, Image enhancement using exposure based sub image histogram equalization, Pattern Recogn. Lett. 36 (2014) 10–14.

Please cite this article in press as: K. Singh, R. Kapoor, Image enhancement via Median-Mean Based Sub-Image-Clipped Histogram Equalization, Optik - Int. J. Light Electron Opt. (2014), http://dx.doi.org/10.1016/j.ijleo.2014.04.093

IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 16, Issue 4, Ver. III (Jul – Aug. 2014), PP 100-106
www.iosrjournals.org

Implementation and Result Analysis of Polyalphabetic Approach to Caesar Cipher

Prachi Patni[1]
[1](*Computer Science & Engineering Department, Government College of Engineering, Aurangabad, India)*

Abstract: *In the modern world as there is drastic hike in use of internet for our daily work there is need to keep our information safe and secure so that an intruder can't misuse it. Cryptography was established to solve this problem. Cryptography is an art of transforming information (Plain Text) using encrypting algorithms into a form that is not readable (Cipher Text) without access to specific decoding algorithms.*
In this paper the author presents a novel approach to cryptographic techniques and illustrates the result and analysis of the proposed algorithm and points out that it is with improved security from many kind of attacks.
This paper is partitioned in following sections: 1st section contain basic introduction about cryptography and Caesar Cipher, 2nd section includes proposed system, 3rd contain performance analysis where proposed system is compared with other techniques, 4th include Conclusion and Future Scope and last section contains References.
Keywords: *Cryptography, Caesar, Cipher*

I. Introduction

Internet can be called as backbone of this modern era which includes interchanging of large amount of data between various communication channels.This modern era is dominated by paperless transactions in offices by means of E-mail messages, E-cash transactions, etc. In various business and commercial sectors, there may be some secret information like confidential information banking transactions, government information, credit information is transferred over web using social network, E-mails etc. So there is a need to develop a scheme that guarantee to protect the information from the attacker. Cryptology is at the way of providing such guarantee. The word cryptology is derived from two Greek words: kryptos, which means "secret or hidden" and logos, implies for "description". Cryptology comprises of two competing skills – concealment and solution. The concealment part of cryptology is known as cryptography. Cryptography is often called "code making." The solution part of cryptology is called cryptanalysis. Cryptanalysis is often called "code breaking". [12]

1.1 Cryptography

The purpose of cryptography is to convert plaintext (message) to unreadable cipher text.

1.1.1 Terminology used in cryptography:[6]

i) Plain Text: This is original message or actual secret message which person wants to send to other party.
ii) Cipher Text: This is encrypted message which is output of encryption algorithm. Cipher text message cannot be understood by intruder because of its non-readable format. Is can only be decrypted by authentic user by using key.
iii) Encryption Algorithm: This is the process of reconstructing plaintext into cipher text with use of key.
iv) Key: This is also given as an input to encryption algorithm. It may be alpha numeric, numeric or may be a special symbol.
v) Decryption Algorithm: This is the process of converting cipher text to pain text. It is a reverse method of encryption algorithm. Encryption algorithm takes place at the sender end and decryption algorithm takes place at the receiver end.

1.1.2 Goals of Cryptography

Cryptography renders lots of security goals to ensure the secrecy of data, unmodified data and so on. Following are the various goals of cryptography [13].

i) Authentication:
- It must be achievable for the recipient of a message to make sure of its source.
- An intruder should not be able to act as someone else.

ii) Integrity:
- It should be achievable for the recipient of a message to certify that the message has not been altered by an intruder.
- An intruder should not be able to switch a fake message for an original one.

iii) Non-Repudiation:
- The person who is sender/receiver of the message must not be able to deny later that he/she sent a message.

iv) Confidentiality:
- It ensures that information can only be understood by those who have permission to access the message.

1.1.3 There are three ways by which plain text can be converted to cipher text. [1]

i) Transposition technique: This technique includes rearranging of elements to change its appearance.

ii) Substitution technique: This technique includes replacing an element of plaintext into an element of ciphertext.

iii) Transposition- Substitution technique: In this both techniques are used.

The cryptography is divided into two main classes depending upon the security key they used for encryption and decryption of text.

1.1.4 Types of Cryptography

i) Symmetric Encryption: In symmetric key cryptography same key is used for encryption and decryption purpose i.e. key can be fetched from the decryption algorithm. The encryption algorithm produces the key and then sends it to receiver section where decryption takes place. It is much effective and faster than asymmetrical key cryptography [9]. Symmetric key encryption main drawback is that both the users need to transfer their keys in a secure way because if the key is compromised then whole system will be compromised.

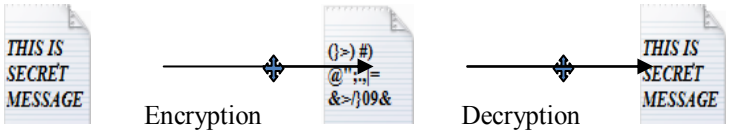


Figure 1. Symmetric Encryption

ii) Asymmetric Encryption: In Asymmetric key cryptography different keys are used for encryption and decryption. It is also called as public key cryptography. It consist of two keys namely public key and private key. Public key is known to the everyone and is used for encryption. Private key is known only the sender/reciever of that key and is used for decryption. The public key and the private key are correlated to each other by any mathematical means[9]. Figure 2 shows working of asymmetrical key cryptography.

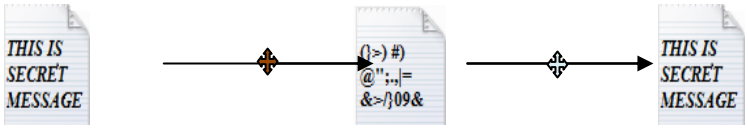


Figure 2. Asymmetric Encryption

1.1.5 Types of attacks on cryptography

i) Cipher text only attack (COA): In this type of attack it is assumed that only set of ciphertexts is available to the cryptanalyst. It is the weakest type of attack of all because cryptanalyst's lack of information.

ii) Known-plaintext attack**:** The attacker knows or can guess the plaintext for some parts of the cipher text. The attacker can now decrypt the rest of the cipher text blocks using this information. This can be accomplished by determining the key that was used during encryption of the data, or making use of some shortcut.

iii) Chosen-plaintext attack: The attacker is able to have any text he likes encrypted with the unknown key. The attacker requires determining the key used for encryption.

1.2 Caesar Cipher

Plaintext letters: zxcvqwertasdfgmnblkjhpoiuy

Ciphertext letters: JFOOFJJGFKJGLFMFGFFKKNKNKNV

Consider an example where a plaintext is given with its respective key. If anybody wants to remember that key it is nearly impossible, so other option is to write the key but it can also be lost, stolen or forgotten. It is necessary to create a key which need not to be written [1]. Therefore various algorithms were proposed to solve this problem. Caesar cipher is one among them. It is called by many names like shift cipher, Caesar's code or Caesar shift. It is one of the simplest and most widely known classical encryption techniques. It is an example of a substitution cipher method [13]. It was used by Julius Caesar to communicate with his army. Caesar decided that he would be shifting each letter by three places left, and so he informed all of his generals of his

decision, and then he sent them that encrypted messages [13]. One of the strengths of the Caesar cipher is its ease of use.

It is a type of substitution cipher in which each letter in the plaintext is replaced by a letter by shifting some positions. For example, suppose shift is 3, A would become D, B would be replaced by E, and so on. As with all Monoalphabetic substitution ciphers, the Caesar cipher can be easily broken [1]. The encryption can also be act as a substitute for modular arithmetic by first replacing the letters by numbers, according to the scheme, A=0, B=1… Z=25.

Encryption of a letter x by a shift n can be representing mathematically as:

$$E_n(x) = (x + n) mod\ 26 \quad (1)$$

Similarly, decryption can be represented as follows

$$D_n(x) = (x - n) mod\ 26 \quad (2)$$

The key can be remembered easily because there is of the pattern in it. Sender and receiver just needed to remember the shift.

II. Proposed Algorithm

In this section, the main algorithm which is used in the proposed system is described. The proposed system, which is used to encrypt the text, is divided into the following 3 main phases:

Phase-1: Creating Ciphertext1: In this phase we open a file and read it stream by stream. The stream is converted into block of 8. Now the block is XORed with array of key generated by Random Number Generator. Now the block created will be used as key for the second stream.

Phase-2: Applying Caesar Cipher: The Ciphertext1 is encrypted again using algorithm Caesar Cipher.

Phase-3: Creating Ciphertext2.

Each above phase consists of number of sub-phases. The detailed description for each of above phase is described as follows.

The flow for proposed system is as depicted in Figure 3

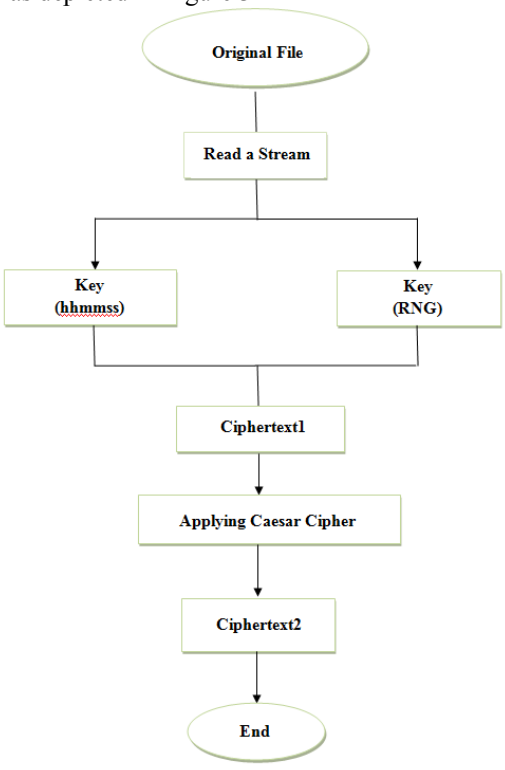


Figure 3: Flow Diagram for Proposed System

This proposed algorithm we created two Ciphertexts because if the file contains two same words they should be encrypted differently. This algorithm also contains two keys. The first reason for having two keys is that this will helps in increasing security. As we know that if the security is compromised then the whole system will be compromised. So it is most necessary and important to make security of key as unbreakable as possible.

PHASE 1:

Step 1: The first Step is reading text stream by stream from a file. Stream must be converted into block of 8. So length of the first stream is calculated.
• If the length of the string is greater than 8, 8 characters from the starting is taken to create a block.
• If the length of the string is found to be less than 8 then padding by 0 is done.
In this way a string is created of size 8.
For example if the first string is "hello" then it will become "hello000" for converting it into ciphertext1. Else if the string is "encrypting" then it will take only "encrypti"
Step 2: Keys are created
• Key 1: The first key is created by using "time" and "date".
For e.g. if today's date is: 13-06-2014 and time is 22: 56:54 so the required key will be 22565420140613. So if the intruder has plaintext and algorithm and try's to create a new file even after 1 second then also the text will be change and the authentic user will know that the file is modified.
• Key 2: The user will enter a random number and according to that number a Random Number Generator will create a number Between 1 to 1000.
For e.g. if user enter 34 as a random number then the Random Number Generator which is an inbuilt function in C# will produce a number say 991.
Both the keys are encrypted in different pattern so that if the intruder can able to guess the first key he/she will not able to compromise the system because the second key will protect the integrity of the system. After encrypting the first and second key the key is written to the output file.
Step 3: Random array is XORed with block of plaintext
By using key1 and key2 as input to Random Number Generator an array of 8 numbers is created. For example [1, 2, 3, 4, 3, 2, 0]. If the block is first block of the file then the string is XORed with random array generated let give name as "A". "A" must not exceed from 25 so for each ciphertext1 which will behave like a key for another blocks they are divided by 25. For second block "A" will be XORed with the second block let give name it as "B". "B" is divided by 25. Then "B" will be XORed with third block and so on.
PHASE-2: Applying Caesar Cipher
Step 4: The Ciphertext1 is again converted into the real length of string.
Step 5: Generate timekey for the Caesar Cipher

In this step timekey for the Caesar cipher is generated. The timekey is generated by the KEY1 by summing each and every digit of Key1.
For e.g. if Key1 is 22565420140613 then the timekey created for Caesar cipher is 2+2+5+6+5+4+2+0+1+4+0+6+1+3 = 41.
Step 6: Generation of Secret key by timekey + Key2 + String Length for Caesar Cipher.

In this step the secret key for the Caesar cipher is created. The Secret key is created by combining three parts. The first part is timekey which is created in the previous Step (Step 5). The second part is Key2 which is created in phase 1 and the third part if finding the length of the string. So the Secret Key will be sum of timkey, Key1, String Length
Step7: Split the string into odd and even position characters.

As the step is clear itself. The string is divided into two arrays. The first array is of the characters which are placed on the odd position in the string and second array is of characters which are placed on the even position in the string.
Step 8: Reverse odd and even position characters.
Step 9: Apply Caesar cipher on odd position characters from right to left and decrease key by 1.
Step 10: Apply Caesar cipher on even position characters from left to right and decrease key by 1.
Step 11: Merge even and odd position characters.
PHASE 3: Creating Final Ciphertext.
Step 12: From the string generated in the previous step file 1st middle and last character and subtract them by 26.
Step13: Repeat all the steps for each and every word till all the word are converted into ciphertext2.
Step 14: End.
The decryption algorithm will work in reverse order of the encryption algorithm. First step is same for encryption as well as decryption i.e. reading the text string by string. Secondly 1st middle and last character is found then they are added with 26. Then split the ciphertext in even and odd position arrays. Applying Caesar cipher to both the arrays. Generate plaintext1 and with the help of Key1 and Key2 generate final plaintext.

III. Performance Analysis

In this paper, the popular algorithms including AES (Rijndael), DES, 3DES were implemented, and their performance was compared by encrypting input files of varying contents and sizes. The algorithms were implemented in C#.NET.
Table1 and Table2 show the comparison between various algorithms [7].

Table 1: Comparison between various encryption algorithms

Input Size (in Kbytes)	AES	DES	3DES	Proposed Method
49	56	29	54	45
59	38	33	48	47
100	90	49	81	55
247	112	47	111	63
321	164	82	167	74
694	210	144	226	100
899	258	240	299	122
963	208	250	283	150

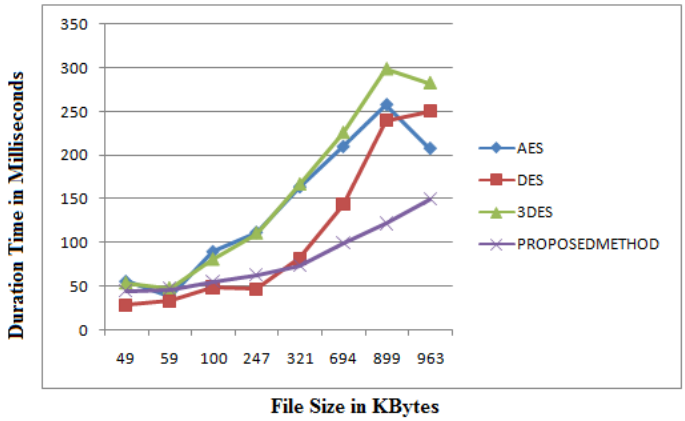


Figure 4: Graph showing encryption time for different file size

Table 2: Comparison between various decryption algorithms

Input Size (In Kbytes)	AES	DES	3DES	Proposed Method
49	63	50	53	50
59	58	42	51	52
100	60	57	57	58
247	76	72	77	65
321	149	74	87	76
694	142	120	147	103
899	171	152	171	120
963	164	157	177	130

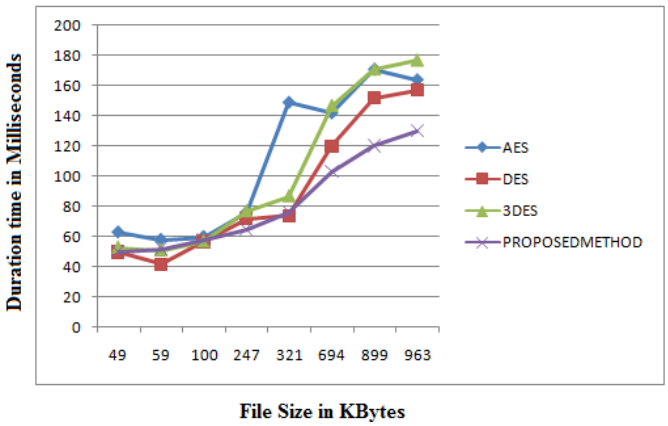


Figure 5: Graph showing encryption time for different file size

The task of performance analysis of the proposed system is accomplished by executing some experimental testing. For finding the performance of the proposed method some factors are considered. By comparing these factors with the other algorithms analysis is done

1. Tunability: It could be advantageous to be able to dynamically specify the encrypted portion and the encryption parameters correspond to different requirements. In this paper, tunability can have only one value either ‘yes’ or ‘no’.

2. Computational Speed: In many applications it is required that encryption process as well decryption process to be fast
3. Key Length Value: In the encryption process the key management is the important aspect that shows how the data is encrypted. The symmetric key algorithm uses a variable key length therefore they use longer key. Hence, there is a need of key management.
4. Encryption Ratio The encryption ratio is the measure of the amount of data that is to be encrypted. Encryption ratio must be small to reduce the complexity.

Table 3 shows comparison between various techniques on different factors [8]

Table 3: Performance Comparison of Image Encryption Schemes

Algorithm	PlainText	CipherText	Key	Encryption Ratio	Speed	Tunability	Security Against Attacks	Advantages
AES	128	128	128 /192 /256	High	Fast	No	1) Chosen- plain text 2) Known-plain text	1) Reliable 2) Longer Key length supported
DES	64	64	56	High	Fast	No	1) Brute-force attack	1) Less no of computation. 2) Simple and fast. 3) Cryptanalysis is difficult
3DES	64	64	168	Moderate	Fast	No	1) Chosen- plain text 2) Known-plain text Brute-force attack	1) More Reliable 2) Longer key length 3) Cryptanalysis is difficult
BLOWFISH	64	64	32-448	High	Fast	Yes	1) Dictionary attacks	1) Fast and secure. 2) Compact.
RSA	MIN. 512	MIN. 512	512-1024	High	Fast	No	1) Timing attacks	1) Offering high performance 2) Delivering broad platform support
Proposed System	64	64	112-168	Moderate	Fast	No	1) Chosen - plain text 2) Known - plain text 3) Chosen - cipher text	1) Fast and secure 2) Contain two keys 3) Less no. of computations.. 4) Cryptanalysis is difficult

IV. Conclusions And Future Scope

In this paper, the author proposed and implemented an approach for text encryption and decryption. Proposed technique for secure transmission of Plaintext file has covered many limitations and security threats of older techniques of similar kind. Increasing levels of security provided in proposed algorithm including the block transformation, XOR technique, randomized bits and use of modified Caesar cipher. The algorithm itself has many advantages for increasing security. The use of two key gives double security. The polyalphabetic approach to Caesar cipher makes proposed algorithm both secure and easy to implement. The use of different key generation algorithm secure proposed application from various attacks.

From the implementation and testing results it is concluded that the proposed method is suitable for many types of text files like Excel, text, and word document. The algorithm is independent of machine i.e. if the other system has the application installed then he/she can encrypt the file and send encrypted file via internet and receiver can decrypt it. The computational speed of algorithms depends on the size of the text file to be encrypted or decrypted. Another main feature of this method is that it satisfies the properties of Shannon's Confusion and diffusion and Kerckoff's law so without knowing keys it makes decryption nearly impossible

Future Scope

- Security of the key can be improved.
- Encryption time and Decryption time can be decreased.
- In case of Caesar cipher any other algorithm can be used

References

[1]. Prachi Patni, A Poly-alphabetic approach to Caesar Cipher Algorithm, International Journal of Computer Science and Information Technologies (IJCSIT), Volume 4, 2013, 954-959, ISSN: 0975-9646.

[2]. Amit Joshi and Bhavesh Joshi, A Randomized Approach for Cryptography in Emerging Trends in Networks and Computer Communications (ETNCC), 22-24 April 2011.

[3]. S G Srikantaswamy and Dr. H D Phaneendra, Improved Caesar Cipher with Random Number Generation Technique and Multistage Encryption, International Journal on Cryptography and Information Security (IJCIS), Vol.2, No.4, December 2012.

[4]. Ramandeep Sharma, Richa Sharma and Harmanjit Singh, "Classical Encryption Techniques" published in International Journal of Computers & Technology, Volume 3. No. 1, AUG, 2012.
[5]. O.P. Verma, Ritu Agarwal, Dhiraj Dafouti and Shobha Tyagi, Performance Analysis of Data Encryption Algorithms, IEEE Delhi Technological University, India, 2011.
[6]. Somdip Dey, SD-AREE: An Advanced Modified Caesar Cipher Method to Exclude Repetition from a Message published in International Journal of Information & Network Security (IJINS), Vol.1, No.2, June 2012, pp. 67~76.
[7]. Diaa Salama Abd Elminaam, Hatem Mohamed Abdual Kader, and Mohiy Mohamed Hadhoud, Evaluating The Performance of Symmetric Encryption Algorithms, International Journal of Network Security, Vol.10, No.3, PP.213-219, May 2010.
[8]. AL. Jeeva, Dr. V. Palanisamy, K. Kanagaram, Comparative Analysis Of Performance Efficiency And Security Measures Of Some Encryption Algorithms, International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622, Vol. 2, Issue 3, May-Jun 2012, pp.3033-3037.
[9]. Kashish Goyal and Supriya Kinger, Modified Caesar Cipher for Better Security Enhancement published in International Journal of Computer Applications (0975 – 8887)(IJCA), Volume 73– No.3, July 2013.
[10]. William Stallings, Cryptography and Network Security, Fourth Edition (Prentice-Hall) pp.80-81.
[11]. http://www.cs.trincoll.edu /~crypto/historical/caesar.html (Savarese, C and Hart, B, The Caesar Cipher, Last updated: 04/26/2010 03:46:57).
[12]. http://www.nku.edu (Fall 2006 Chris Christensen).
[13]. CRYPTOGRAPHY, https://en//.wikipedia.org/wiki/cryptography.

Implementation of Modified Synchronous Reference Theory for Control of Grid Connected Distributed Generation Systems

Alka Singh
Department of Electrical Engineering
Delhi Technological University
Bawana Road, Delhi, India
alkasingh.dr@gmail.com

Abstract—This paper deals with the control of distribution energy resource (DER) connected to a grid connected system feeding non-linear loads. The DER is controlled to provide power quality improvement capabilities viz. power factor correction, harmonic reduction and load balancing. The developed controller is based on modification of Synchronous Reference Theory. It possesses unique features such as filtration and fundamental voltage extraction from the polluted grid voltages to eliminate harmonics and unbalance. Additionally, the control is implemented over the square of DC link voltage. The distributed generation source is realized as a voltage source converter. Both simulation and experimental results prove the effectiveness of control algorithm with non-linear loads. The control algorithm works well for power factor correction, harmonic reduction and operation under unbalanced load condition. The dynamics of the system under load change and load unbalancing are well depicted.

Keywords-power factor correction; distributed generation; load balancing; utility grid

I. INTRODUCTION

A number of benefits of renewable energy resources exist viz. high reliability, modular in nature and environment friendly operation. These advantages have ensured increased acceptance of distributed resources in the world [1-12]. The renewable energy sources (wind, solar etc.) are currently being interfaced to the conventional systems using voltage source inverters (VSC's)[2-4]. There are several advantages of distributed resources over conventional sources so a high emphasis is nowadays paid to renewable energy sources

(RES) and their interconnection to grid. These sources are connected at the low or medium voltage levels. A primary energy source generally in the form of wind energy, photo-voltaic cell, fuel cell, battery etc is connected at the primary of the converter which converts this power and supplies it to the load in islanded mode. A number of such voltage source converters may be connected in parallel and to the utility grid also. Control of such distributed energy sources [5-13] is an important control aspect.

Very recently the concept of multifunctional Distributed Generators has been introduced which is controlled to provide two equally important tasks of providing power delivery as well as power quality improvement. Active power filters [1,12, 14] provide reactive power compensation and power quality improvement only. The concept is to flexibly control the distributed generator resource as a source and extract maximum power from it and also provide compensation current from it. Lately, power converters in both voltage and current control mode [7,8] have been used for utility interface for distributed generation sources. These converters are mainly interfaced with L or higher order LC / LCL filters [9-11]. Repetitive control, hysteresis regulation, predictive control are frequently used and multi-loop feedback control with LC/ LCL filters is gaining popularity[4-11]. However, several points of concern relate to LCL filters viz. the selection of the outer and inner control variables for improving system stability, avoiding resonance condition and study the effect on system performance under parameter variation. Fig.1 shows a number of distributed generation sources connected to the utility grid.

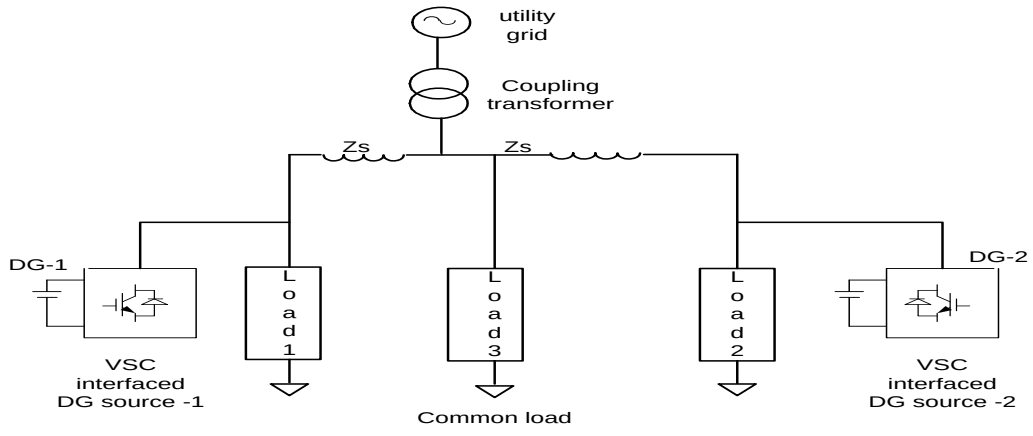


Fig. 1. A number of distributed generation sources connected to utility grid

II.SYSTEM DESCRIPTION

The diagram shows the distributed energy source feeding variety of loads (linear/ non-linear) in grid connected mode. A three-leg insulated gate bipolar transistors (IGBT's) based VSC is used and controlled as DStatcom. It is connected to the point of common coupling (PCC) through a set of interfacing inductors (R,L).

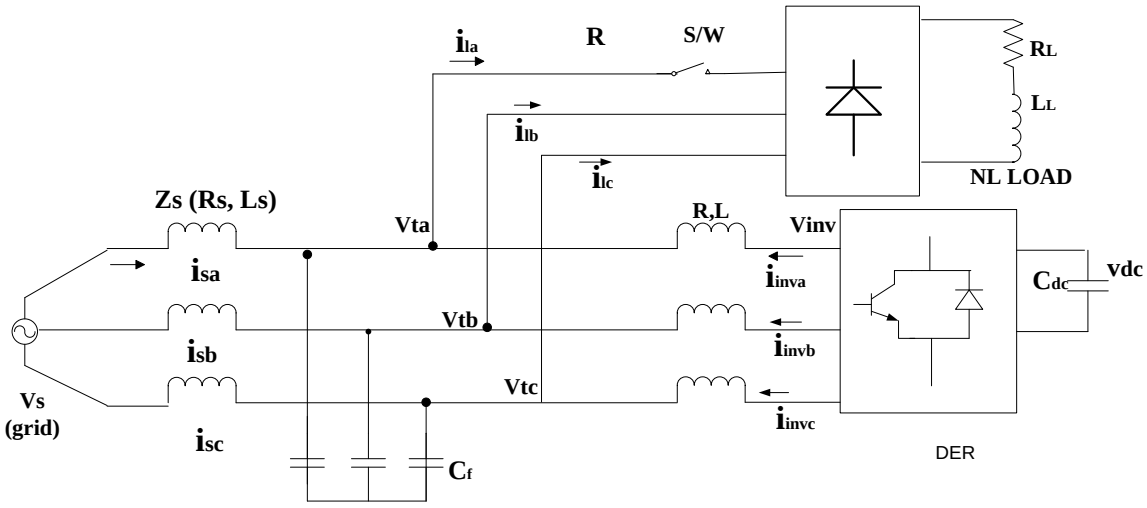


Fig. 2. Schematic diagram of utility grid and Distributed energy source feeding loads

The implemented system requires VSC, interfacing inductors, scaling circuits for sensing the input signals and driver circuit for output signals. Active power transfer flow can also take place by placing an active source (battery or renewable energy source) at DC link. Three Hall Effect current sensors (LEM LA 25- NP) are used to sense load currents for phase 'A', 'B' and 'C'. To sense the source currents for phase 'A' and 'B', another set of Hall Effect current sensors (LEM LA 25- NP) are used. The third phase source current is estimated by considering that the algebraic sum of three phase source currents is zero. Two voltage sensors (LV 25-P) are used to sense phases 'A' and 'B' voltages. The third phase voltage is estimated considering the algebraic sum of three-phase source voltages is zero. The DC link voltage is sensed using the voltage sensor LV 25P/ SP5. Scaling circuits are designed to use in between the sensed signals and ADC (Analog to Digital Converter) of DSP to have a zero offset with the sensed signals. The control algorithm is realized in MATLAB environment with

real time blocks. Hex codes are loaded in the DSP-dSPACE 1104 to generate six switching signals for IGBT's through PWM (pulse width modulation) pins of a DSP. A Semikron three-phase VSC (model MD B6CI 800/415-35F) with six IGBT based switches is used for experiments. Table I lists the parameters of the system used for simulation as well as experiments.

III . MATHEMATICAL MODEL

For the system configuration shown in Fig.2 , v_{ta}, v_{tb} and v_{tc} denote the terminal voltages at the PCC and v_{sa}, v_{sb} and v_{sc} are the voltages at the three–phase ac source (grid), then the following differential equations can be written.

$$v_{sa} = v_{ta} - R_s i_{sa} - L_s \frac{di_{sa}}{dt} \quad (1)$$

$$v_{sb} = v_{tb} - R_s i_{sb} - L_s \frac{di_{sb}}{dt} \quad (2)$$

$$v_{sc} = v_{tc} - R_s i_{sc} - L \frac{di_{sc}}{dt} \quad (3)$$

The equations relating the inverter voltages with the PCC voltages are written as

$$v_{ta} = v_{inva} - Ri_{inva} - L \frac{di_{inva}}{dt} \quad (4)$$

$$v_{tb} = v_{invb} - Ri_{invb} - L \frac{di_{invb}}{dt} \quad (5)$$

$$v_{tc} = v_{invc} - Ri_{invc} - L \frac{di_{invc}}{dt} \quad (6)$$

where R, L denote the interfacing resistor and inductor values and i_{inva}, i_{invb}, i_{invc} and v_{inva}, v_{invb}, v_{invc} denote the inverter currents and voltages in three phases. At the DC link the equation (7) is valid

$$\frac{C dv_{dc}}{dt} = i_{dc} \quad (7)$$

The equations shown above can be depicted in the form of a single closed loop control for the control of grid current (i_s) as shown in Fig.3

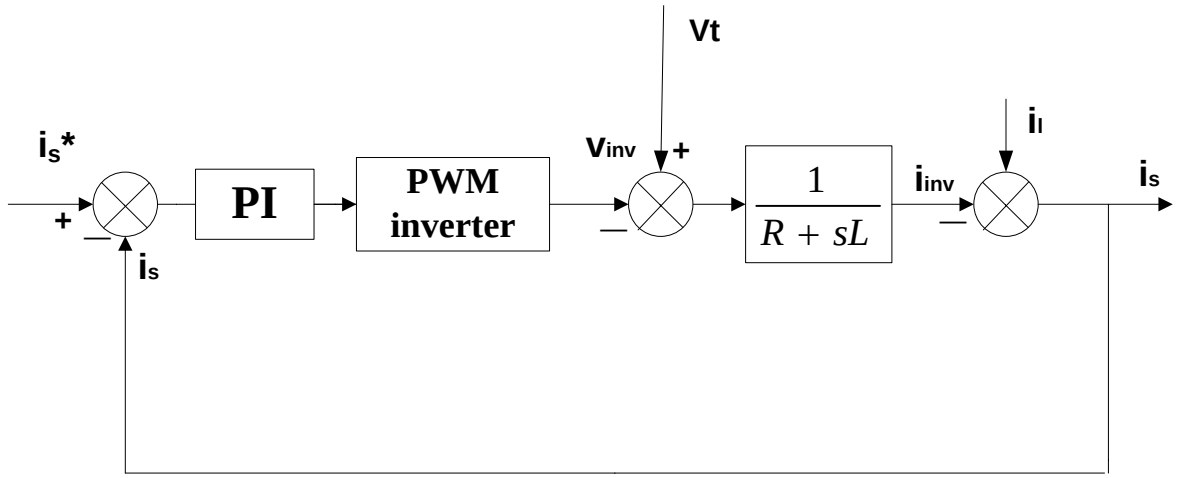


Fig.3 Simplified closed loop analysis of the system

The PWM inverter is represented as $K_{inv}/(1 + sT_{inv})$ R and L represent the interfacing resistance and inductance respectively. For further simplification T_{inv}=0.

The current controller is represented in the form of PI controller with transfer function,

$$Gc(s) = Kp + (Ki/s) \quad (8)$$

The closed loop transfer function is derived to be

$$\frac{i_s}{i_s^*} = \frac{K_{inv}(K_p s + K_i)}{s^2 L + s(R + K_{inv}K_p) + K_{inv}K_i} \quad (9)$$

Comparing with the standard closed loop transfer function,

$$\frac{i_s}{i_s^*} = \frac{w_n^2}{s^2 + 2s\zeta w_n + w_n^2} \quad (10)$$

The values of damping ratio ζ, and angular frequency w_n are found to be influenced by parameters K_p, K_i , K_{inv} and parameters R and L. The desired damping ratio is obtained by varying the gain constants.

III. CONTROL ALGORITHM

The control algorithm used for the control of distributed generation source in this paper is based on modified synchronous reference frame theory. The synchronisation is achieved by deriving unit templates from the supply grid. Since, it is observed that the supply voltages have harmonics and are unbalanced, hence, the use of filters and extraction of fundamental voltage component of the three phases is undertaken as a preliminary step. These filtered supply voltages are used as PCC voltages for the generation of reference current generation, as is discussed below. Fig.4 shows the control algorithm developed using synchronous reference frame (SRF) theory. A noticable difference is that the control is developed over the square of DC link voltage. Although the DC link voltage is sensed, the PI controller is developed over the error between the actual and reference value of square of DC link voltage. The SRF based controller utilizes the abc-dq transformation (Park's transform) to transform the current and voltage waveforms into a reference frame that rotates synchronously with the grid voltage.

Filtered fundamental PCC voltages (v_{saf}, v_{sbf}, v_{scf}) are obtained by filtering voltages (v_{sa},v_{sb}, v_{sc}) through band pass filters, whose cutoff frequency is between 30Hz to 60Hz. The symmetrical component method developed by Fortesque [1] is used to obtain fundamental voltage components (v_{sa1},v_{sb1}, v_{sc1}) from the filtered voltages as below.

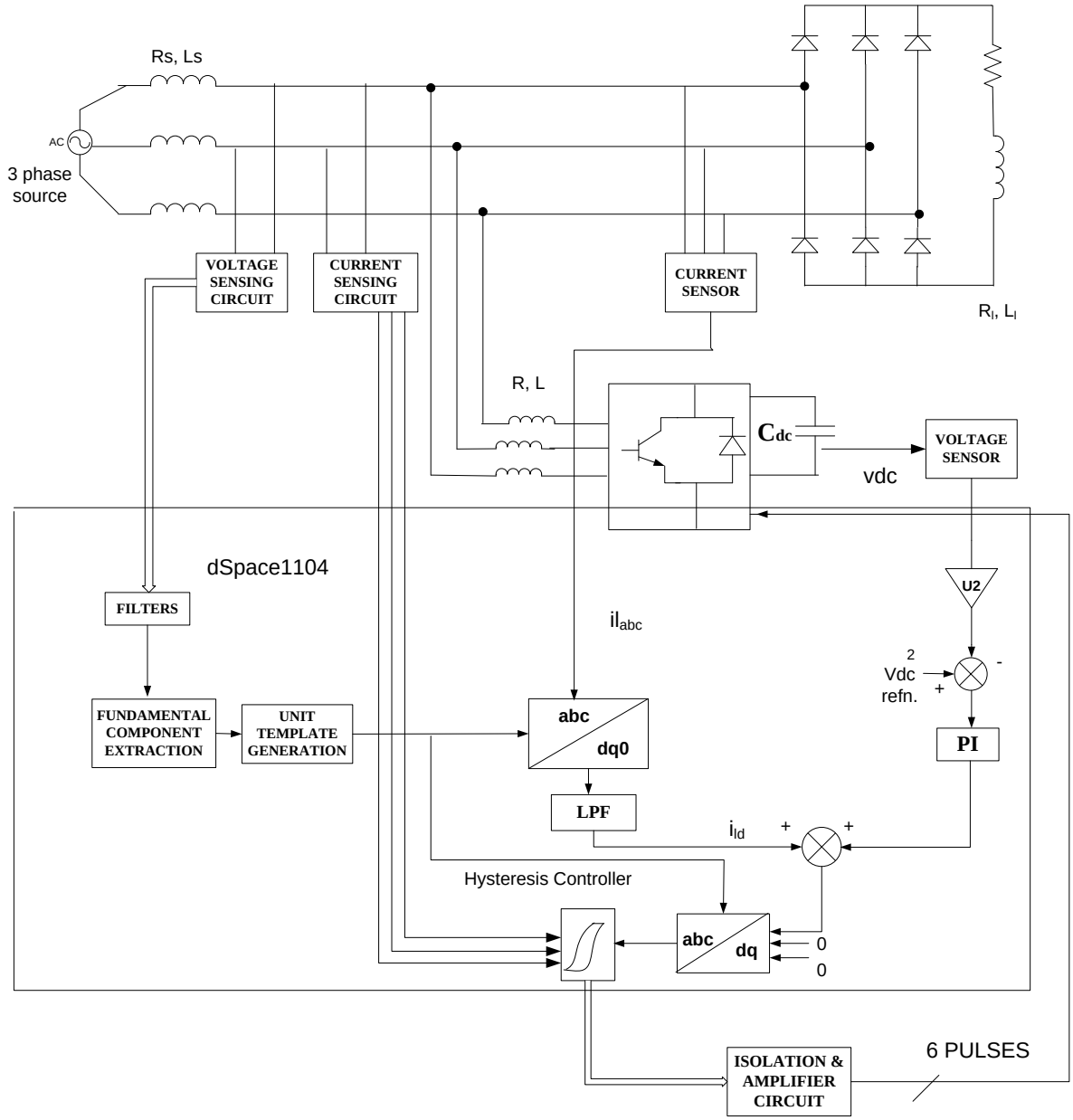


Fig.4 Complete control scheme for the system

$$v_{sa1} = \frac{1}{3}(v_{saf} + \lambda v_{sbf} + \lambda^2 v_{scf}) \tag{11}$$

$$v_{sb1} = \frac{1}{3}(v_{sbf} + \lambda a v_{scf} + \lambda^2 v_{saf}) \tag{12}$$

$$v_{sc1} = \frac{1}{3}(v_{scf} + \lambda v_{saf} + \lambda^2 v_{sbf}) \tag{13}$$

where $\lambda = e^{j120} = -0.5 + 0.866j$ and $\lambda^2 = e^{j240} = -0.5 - 0.866j$

The amplitude of the PCC voltage is now obtained using

$$V_{tm} = \{2/3(v_{sa1}^2 + v_{sb1}^2 + v_{sc1}^2)\}^{1/2} \quad (14)$$

The unit in-phase vectors (u_a, u_b, u_c) are calculated as:

$$u_a = v_{sa1}/ V_{tm} \quad (15)$$

$$u_b = vs_{b1}/ V_{tm} \quad (16)$$

$$u_c = v_{sc1}/ V_{tm} \quad (17)$$

The quadrature unit current vectors (w_a, w_b, w_c) are derived from in-phase unit current (u_a, u_b, u_c) vectors as:

$$w_a=(-u_b + u_c)/ \{(3)^{1/2}\} \quad (18)$$

$$w_b= (u_a(3)^{1/2} + (u_b-u_c)) / \{2(3)^{1/2}\} \quad (19)$$

$$w_c= (-u_a(3)^{1/2} + (u_b-u_c))/ \{2(3)^{1/2}\} \quad (20)$$

The basic SRF theory involves the transformation of currents in synchronously rotating d-q frame. It uses the PLL block to evaluate θ, transformation angle to convert currents from αβ to dq frame

$$i_\alpha = i_{la}(\sqrt{2/3})\, cos\theta - i_{lb}(\sqrt{1/6}) - i_{lc}(\sqrt{1/6})$$

$$i_\beta = i_{lb}(\sqrt{1/2}) - i_{lc}(\sqrt{1/2})$$

$$i_d = i_\alpha cos\theta + i_\beta sin\theta \quad (21)$$

$$i_q = -i_\alpha sin\theta + i_\beta cos\theta$$

Instead of the PLL block, the unit templates obtained from filtered PCC volatges are used to estimate i_d, i_q values of load currents. The obtained i_d and i_q currents are now passed over Low pass filter (LPF) to extract the dc component i_{ddc} and i_{qdc}. A PI controller is realised over the square of DC link voltage to regulate it to its reference value.

$$i_{d(n)} = i_{d(n-1)} + K_{pd}\{ v_{dc1}^2(n) - v_{dc1}^2(n-1) \} + K_{id}\, v_{dc1}(n) \quad (22)$$

where $v_{dc1}(n) = v_{dcref}^2(n) - v_{dc}^2(n)$ denotes the error in square of DC link voltage calculated over reference and sensed value obtained from the voltage sensor. K_{pd} and K_{id} are the proportional and integral gains of the PI controller. The output of PI controller accounts for the losses of VSC and this component is added to the i_{ddc} for power factor correction.

Once the reference d,q components of currents are obtained, the three phase reference supply currents (i_{sar}, i_{sbr}, i_{scr}) can be obtained. These reference supply currents are compared with the sensed supply currents (i_{sa}, i_{sb} and i_{sc}) to estimate the three phase current error components. The current errors are passed over carrier-less PWM controller to generate gating signals for the six IGBTs (Insulated Gate Bipolar Transistors) of VSC. The control scheme is direct and based on supply (grid) currents instead of the inverter currents in the three phases.

IV. Performance and Results

A MATLAB based model of distributed energy resource is developed and tested under distorted and unbalanced voltage of ac mains to validate the proposed control algorithm. The performance of modified synchronous detection control algorithm in time domain is simulated and analyzed using a three-leg VSC configuration with nonlinear loads. Simulation as well as experimental results are presented for two cases viz. power factor correction and load balancing along with power factor correction.

A) Power Factor Correction and Load Change

Fig. 5 shows the simulated results showing the filtered PCC voltage (v_t), supply currents(i_s), load currents (i_l) , currents (i_c) injected by DStatcom and DC link voltage (v_{dc}). The results are obtained with a non linear load of R_l=30Ω connected at the end of diode rectifier. Load is varied to R_l=50Ω for time duration from t=0.2sec to t=0.28sec. The dynamics of the currents and DC link are evident from Fig.5 and it is observed that the DC link voltage settles down to a reference value of 200V in less than two cycles. The THD of the supply currents is 4.3% when the load currents have a THD of 28%.

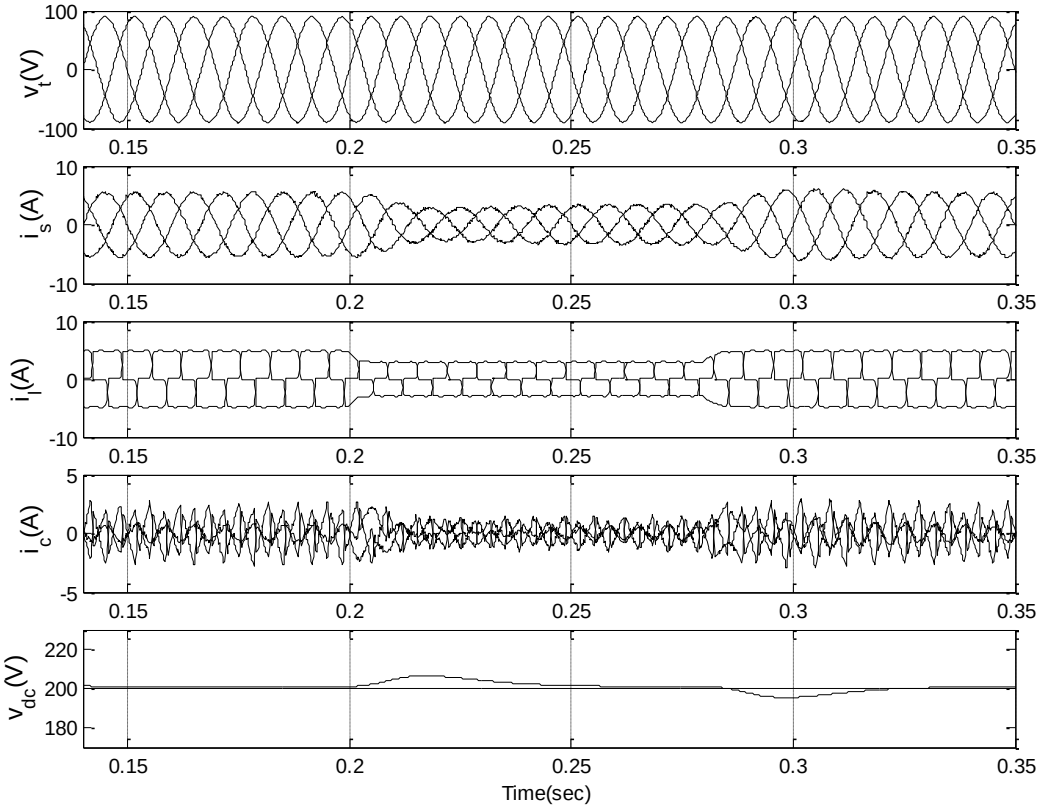


Fig. 5 Simulated results for power factor correction in steady state as well as load removal (t=0.2sec) and addition (t=0.28sec)

b) Power Factor Correction under Unbalanced Load Operation

Fig.6 shows the simulated results for the system under steady state conditions till t=0.2sec when phase 'c' of the load is suddenly disconnected from t=0.2sec to t=0.28sec. The simulated results show the control algorithm not only maintains the DC link voltage but also the grid currents are balanced and

sinusoidal during this interval. The Distributed energy source injects compensating currents of higher magnitude (as seen from simulation results). The normal balanced loads with R_l=30 Ω and L=60mH are restored at t=0.28sec. The THD of the supply currents is 4.4% when the load currents have a THD of 29%.

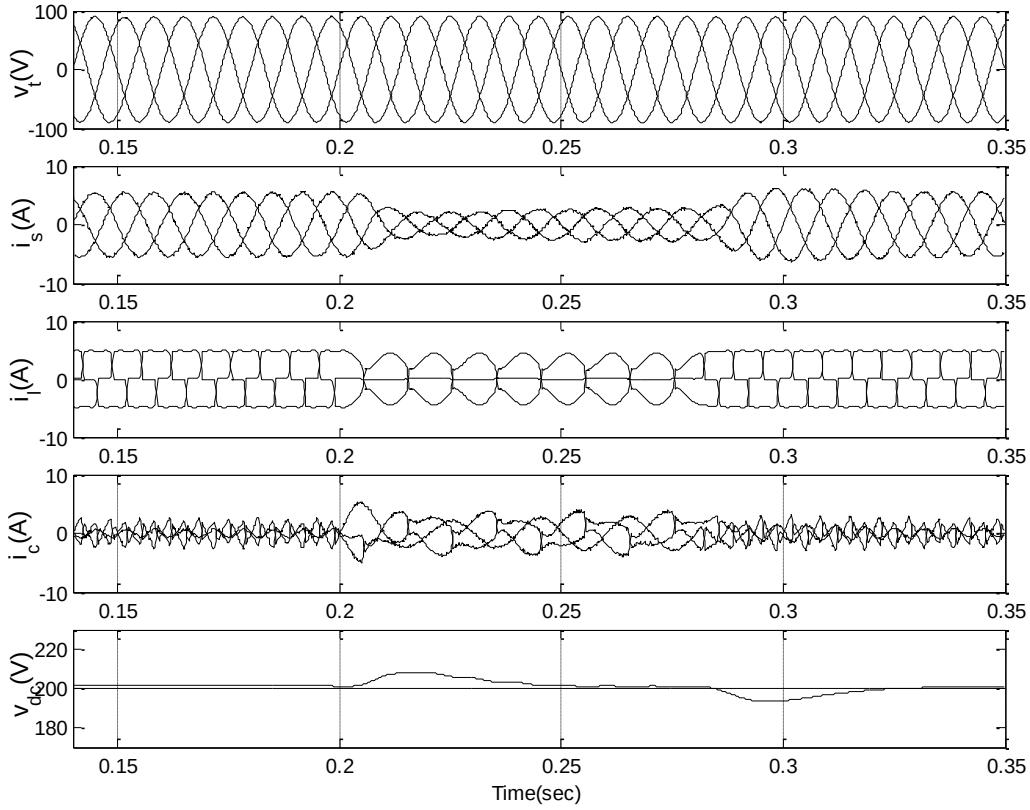


Fig. 6 Simulated results for power factor correction and load balancing in steady state as well as one phase load removal (t=0.2sec) and restoration (t=0.28sec)

V. EXPERIMENTAL RESULTS FOR PROPOSED SCHEME

The experimental results for the same system configuration (shown in Table I) and load are shown in Fig 7-8. Fig.7a shows the waveforms of phase 'a' supply current, load current, compensator current and Dc link voltage in steady state.

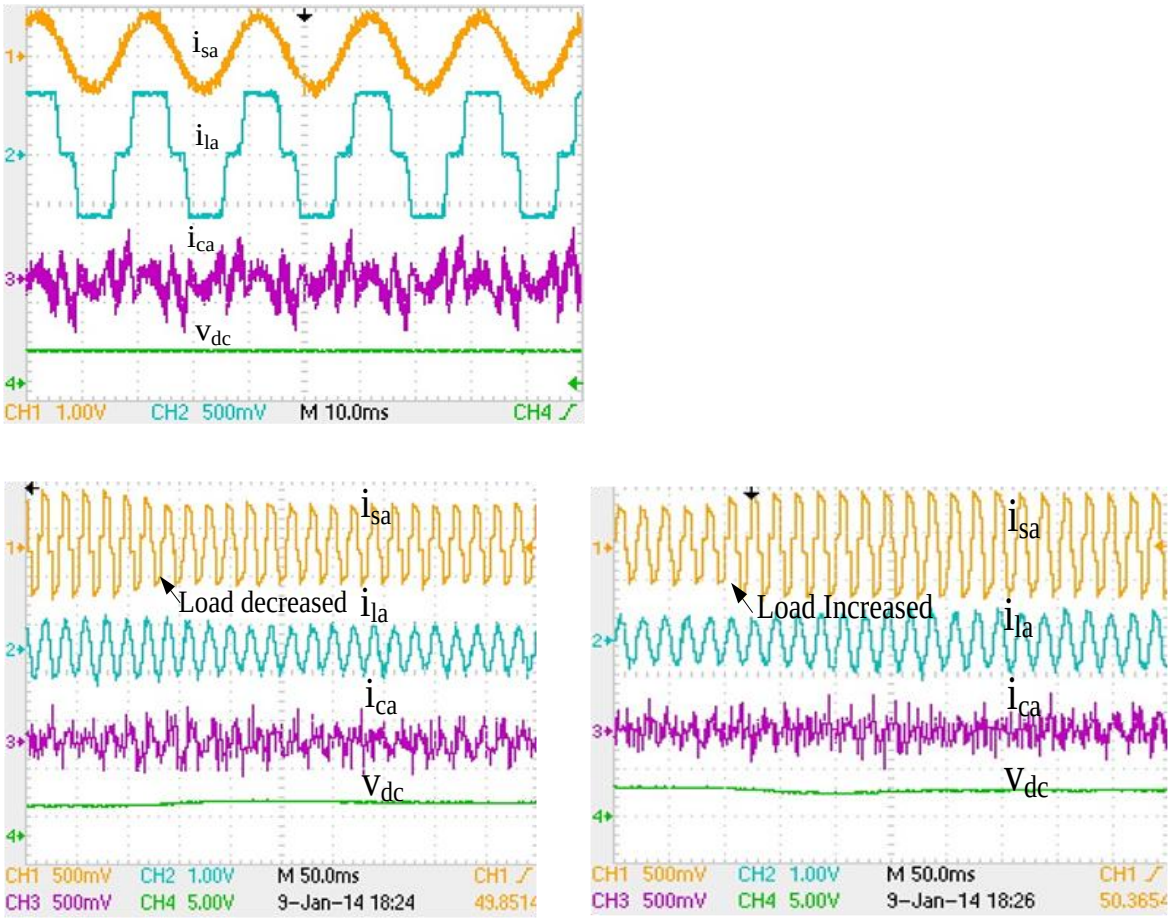


Fig. 7 a Plot of isa, ila, ica, Vdc (10A/div, 5A/div, 5A/div for the currents)
b Plot of ila, ilb, ilc, Vdc (5A/div, 5A/div, 5A/div for the currents) for load decrease
c Plot of ila, ilb, ilc, Vdc (5A/div, 5A/div, 5A/div for the currents) for load increase

Fig.*7b and 7c* show the dynamics of load decrease and load increase respectively. Fig 8a shows the Power Analyser 434 (Fluke) waveforms and THD of the supply current which is 4.4%. The unfiltered PCC voltages are depicted in Fig.8b that show a THD of 2.8% and the magnitudes are also slightly different. It is due to this reason only that filtered, positive sequence voltage components are first extracted and then used for further analysis and reference current generation.

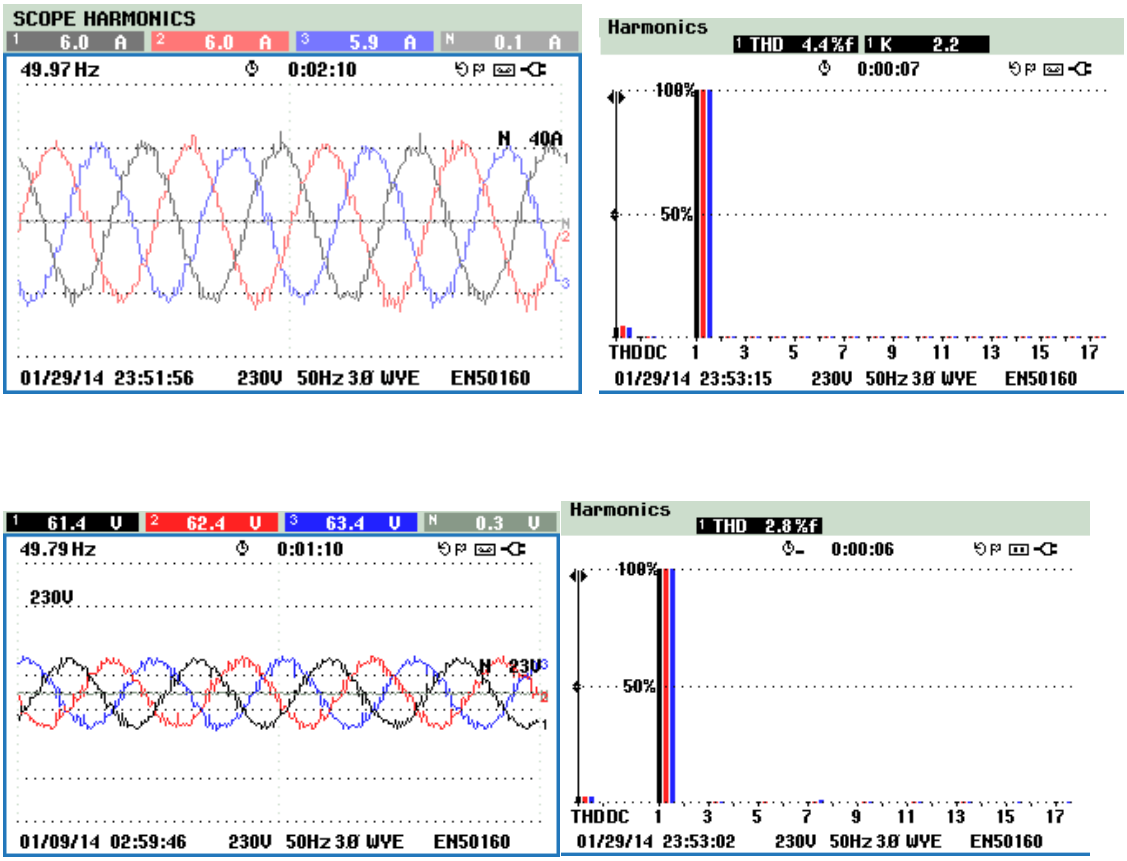
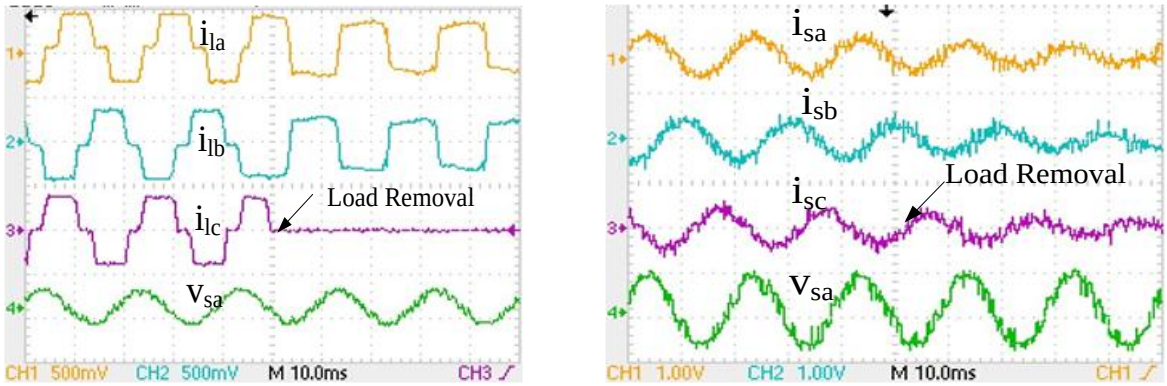
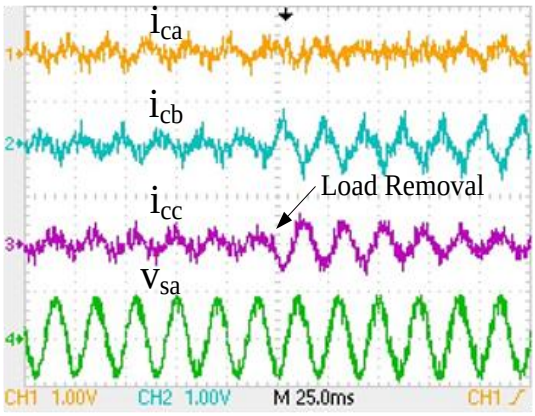


Fig.8a Fluke 434 results showing three phase supply currents (isa, isb, isc) and THD of 4.4%
8b Supply voltages showing 2.8% THD

Experimental verification for power factor correction and load unbalancing are presented next in Fig. 9. The load current in phase ‘c’ is suddenly disconnected and this is evident from the plot Fig. 9b for ilc. Fig. 9c shows the variation of grid supply currents in all the three phases before and after the instant of unbalancing.





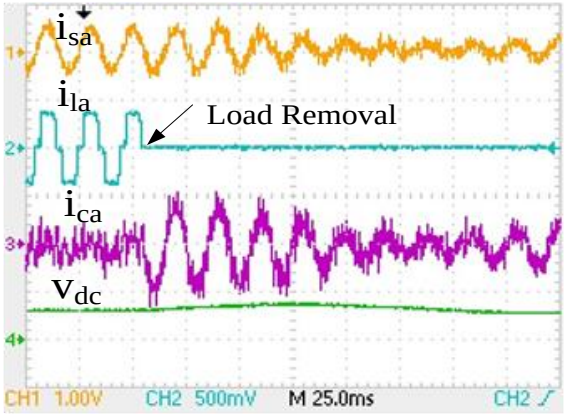


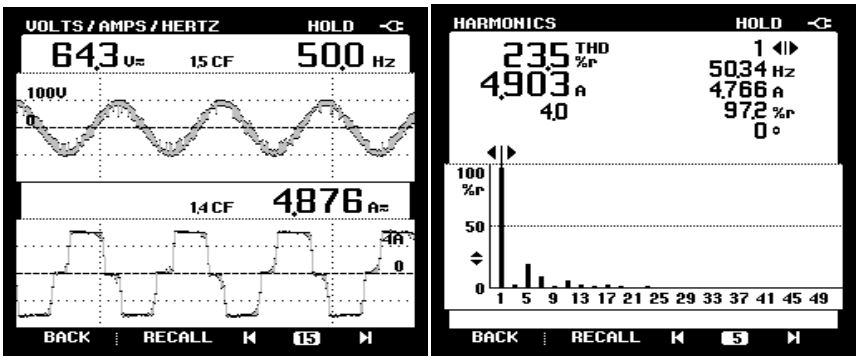
Fig.9a Load current s (5A/div) and vsa
9b Supply voltages (5A/div) and vsa
9c Inverter currents (5A/div) and vsa
9d Source, load, inverter currents (5A/div) and DC link voltage

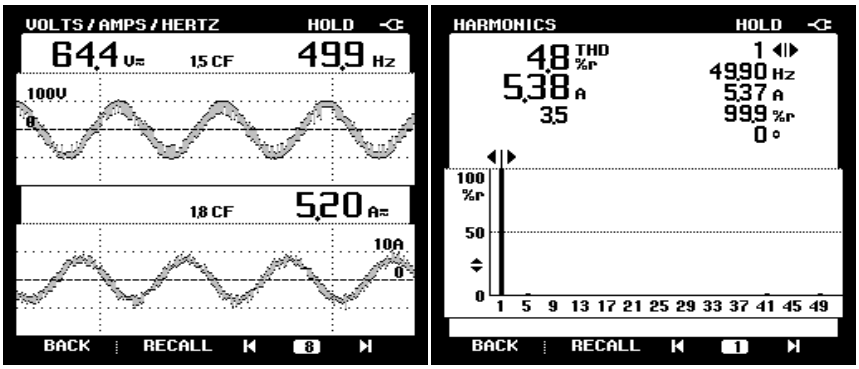
Fig.9c shows the variation of compensator currents in all the three phases before and after load in phase'c' is disconnected for some time. It is observed from Fig9 a-c that the distributed source is controlled to provide necessary compensation currents during unbalanced load conditions. The DER acts as a shunt compensator and regulates the grid currents to equal, balanced and sinusoidal even when one phase of load is switched off. It provides large compensation currents in all the three phases during the interval. Also, it is observed from Fig. 9d that the DC link settles down to its reference value of 200V quite rapidly.

V. RESULT COMPARISON

Conventional SRF scheme has a few drawbacks viz. it cannot be used to obtain optimum results under unbalanced grid supply voltages. The voltage signals are processed by the PLL to generate θ. This transformation angle (θ) is sensitive to voltage harmonics and unbalances. Hence filtered PCC voltages are used in the modified control algorithm for reference current generation. If the unfiltered voltages are used in control algorithm, the reference supply current generated would themselves be non-sinusoidal and it would be difficult to achieve three-phase sinusoidal, balanced supply currents. The use of positive, negative and zero sequence components has been used in this paper, which is simple and effective method for extraction.

The processing of current signals in both the schemes is similar. These are transformed into d-q frame and then filtered. Then, compensating current transformed back to a-b-c frame and fed to hysteresis current controller for switching pulse generation. The results obtained show that under normal operating conditions, both the control strategies are suitable for compensation, but under the distorted supply conditions and unbalanced operating condition, the performance of the conventional scheme deteriorates. The experimental results for the conventional scheme are shown in Fig.10





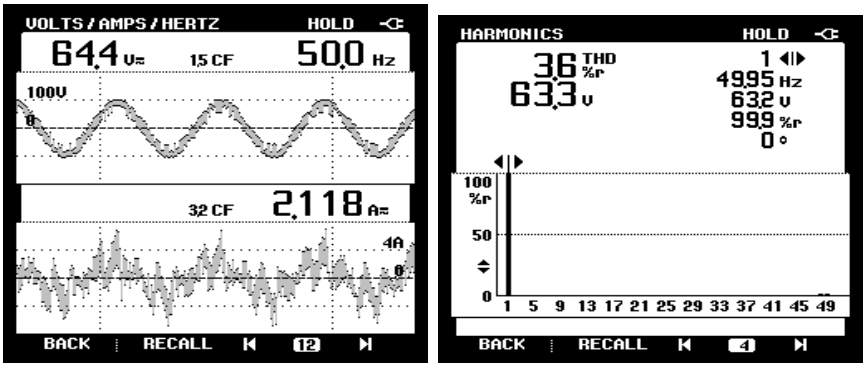


Fig.10a,b vsa and Load current (ila) (5A/div) and its THD 23.5%
10c,d vsa and Supply current (isa) (5A/div) and its THD 4.8%
10e vsa and Compensator current (ica) (5A/div)
10f THD 3.6% in the grid voltage

Fig.10 shows the results with conventional SRF scheme. It is evident that the supply voltages have a high THD level of 3.6% (Fig.10f). With the controller realised on vdc and the gains optimised, the performance of the shunt compensator is observed. The load current has a high THD content of 23.5%, the controller injects a compensating currents to make the supply current sinusoidal. Under the optimised PI gains, the lowest THD observed in supply currents is of the order of 4.8%. This is higher than the THD level of 4.4% observed with the vdc^2 controller and the modified SRF theory. Fig.11-12 show the waveforms and THD for isa, ila with modified SRF theory and conventional SRF theory respectively. Fig.11-12 show the simulation results with both the schemes.

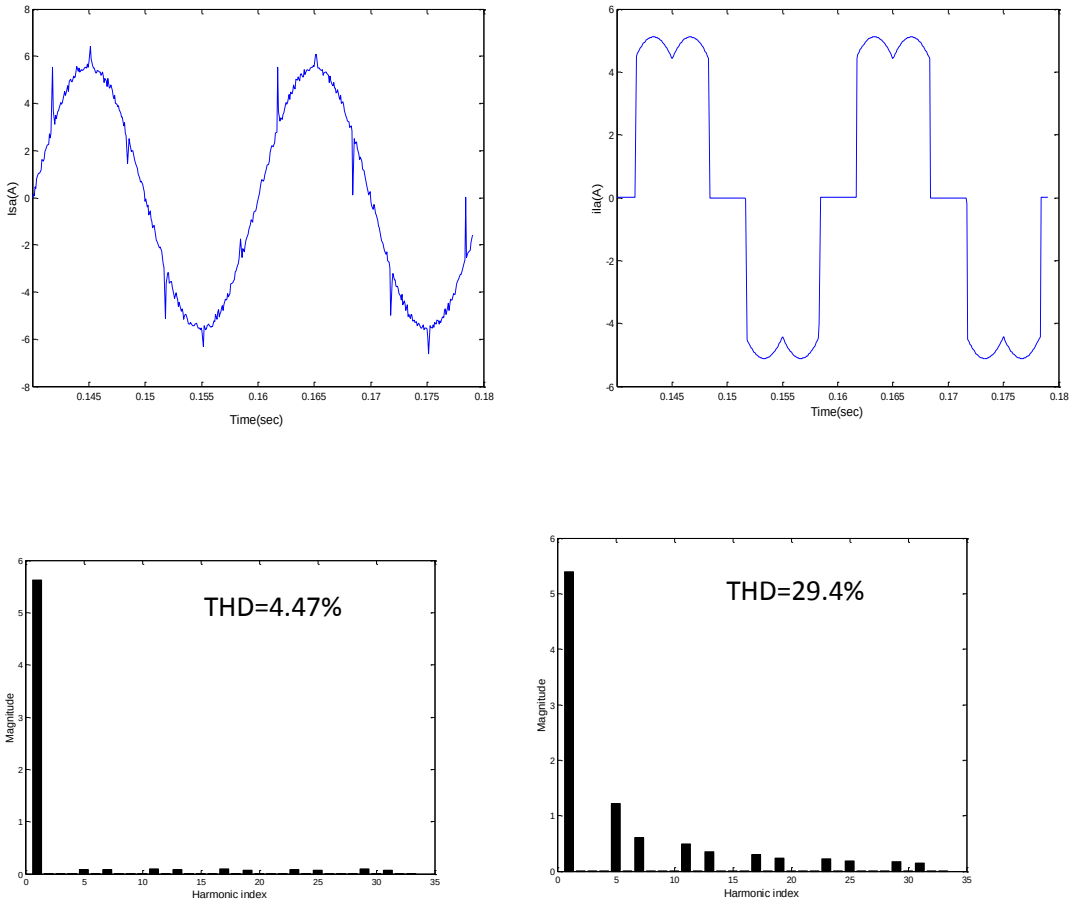


Fig.11 Waveforms and THD for isa and ila for modified SRF theory

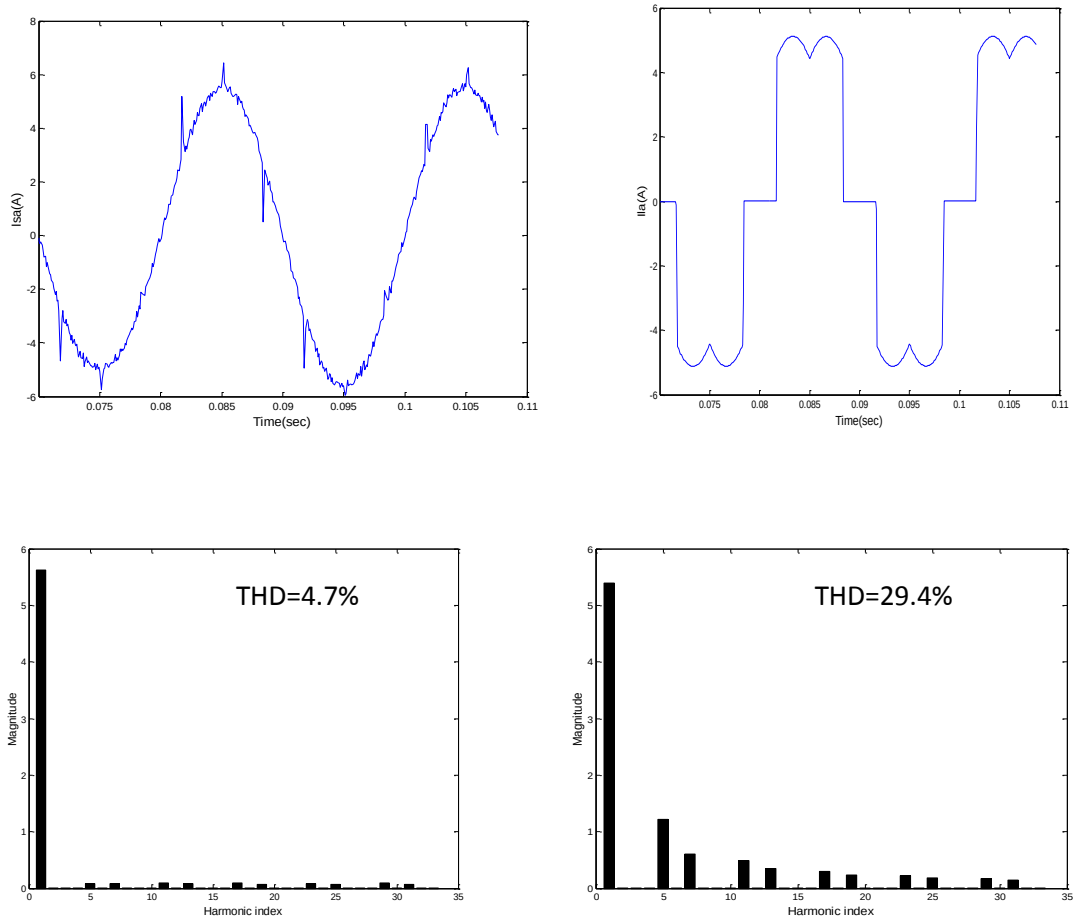


Fig.12 Waveforms and THD for isa and ila for conventional SRF theory

Comparison Results are highlighted in the Table I below

	Parameters (Kp, Ki)	Overshoot (V)	Undershoot (V)	Settling Time (sec)	THD
Conventional SRF	0.35, 0.1	~ 3V	~ 2V	< 2cycles	4.9% approx
Modified SRF	0.00028, 0.00034	~ 4V	~ 3V	< 2cycles	4.5% approx

Table I : Comparison of Control schemes

VII. CONCLUSIONS

This paper discusses and implements a control algorithm based on modification of synchronous reference frame theory for interfacing a distributed energy source to grid. The control algorithm is investigated to inject only reactive power and relieve the grid from supplying it. The control scheme is modified to include fundamental voltage component of the PCC voltage which is having harmonics, and unbalancing of up-to 5%. A single closed loop control for the system is investigated for control of grid currents. The control is developed and experimentally tested on v_{dc}^2 rather than the usual control on v_{dc} alone. Two benefits – viz lower possible reduction in THD level and lower range of PI controller gain settings were observed with this control. Simulation results and experimental results for the prototype system are obtained for power factor correction, load unbalancing and reduction of THD in grid currents. The experimental results obtained are very close to the simulation results. The proposed control scheme is able to achieve a THD of 4.4% in grid currents when the load current is 28.7%, thus meeting IEEE 519 standards. The distribution shunt compensator has been utilized to provide reactive power compensation only, though real power injection into the system is also possible with an active source at the DC link.

Acknowledgment

The author wishes to thank the Department of Science and Technology, Government of India for the Sponsored Project, No. SR/FTP/ETA-20/2010.

TABLE I: SYSTEM PARAMETERS (Both Simulation and Experimental)

Parameters	Value
Grid voltage *E* (RMS)	63.5 V
DC-link voltage V_{dc}	200 V
DC link capacitor C_{dc}	1600μF
Inverter-side filter L	3.5 mH
Filter capacitor C_f	3 μF
Rectifier resistor R_l	30 Ω
Load Inductor L_l	60mH

Sampling frequency f_s	10 kHz
Switching frequency f_c	10 kHz

Conflicts of Interest

The author(s) declare(s) that there is no conflict of interests regarding the publication of this article.

References

[1] H. Akagi, E.H. Watanabe and M. Aredes, *Instantaneous Power Theory and Applications to Power Conditioning*, IEEE Press, Wiley InterScience, 2007.

[2] F. Blaabjerg, Z. Chen and S. Kjaer, "Power electronics as efficient interface in dispersed power generation systems", *IEEE Trans. on Power Electronics* , vol. 19, No.5, Sep 2004, pp1184-1194.

[3] *Recommended Practices and Requirements for Harmonic Control in Electrical Power Systems*, IEEE Standard 519-1992, 1993.

[4] R. Teodorescu, M. Liserre, and P. Rodríguez, "Grid converters for photovoltaic and wind power systems," John Wiley & Sons, 2011, pp. 289–311.

[5] F. Blaabjerg, Remus Teodorescu, M. Lissere and A. V. Timbus, "Overview of Control and Grid Synchronization for Distributed power generation systems", *IEEE Trans. on Industrial Electron.* , vol. 53, No.5, Oct 2006, pp1398-1409.

[6] Y. Chung, W. Liu, D. A. Cartes, E. Collins, "Control Methods of Inverter Interfaced Distributed Generators in a Microgrid system", *IEEE Trans. on Industry Applications*, Vol. 46, No.3, May 2010, pp1078-1088.

[7] J. He, Y. W. Li, and M. S. Munir, "A flexible harmonic control approach through voltage-controlled DG–grid interfacing converters," *IEEE Trans. Ind. Electron.*, vol. 59, pp. 444-455, Jan. 2012.

[8] J. He and Y. W. Li, "Hybrid voltage and current control approach for DG-Grid interfacing converters with LCL filters," *IEEE Trans. Ind. Electron.*, vol. 60, pp. 1797-1809, May 2013.

[9] P. C. Loh and D. G. Holmes, "Analysis of multiloop control strategies for LC/CL/LCL-filtered voltage-source and current-source inverters," *IEEE Trans. Ind. Appl.*, vol. 41, pp. 644-654, Mar./Apr. 2005.

[10] J. He and Y. W. Li, "Generalized closed-loop control schemes with embedded virtual impedances for voltage source converters with LC or LCL filters," *IEEE Trans. Power Electron.*, vol. 27, pp. 1850-1861, Apr. 2012.

[11] N. He, D. Xu, Y. Zhu, J. Zhang, G. Shen, Y. Zhang, J. Ma, and C. Liu, "Weighted average current control in a three-phase grid inverter with an LCL filter," *IEEE Trans. Power Electron.*, vol. 28, pp. 2785-2797, Jun. 2013.

[12] Y. Li, D. M. Vilathgamuwa, and P. C. Loh, "Microgrid power quality enhancement using a three-phase four-wire grid-interfacing compensator," *IEEE Trans. Ind. Appl.*, vol. 41, pp. 1707-1719, Nov./Dec. 2005.

[13] D. N. Zmood, and D. G. Holmes, "Stationary frame current regulation of PWM inverters with zero steady-state error," IEEE Trans. Power Electron., vol. 18, no. 3, pp. 814-822, May 2003.

[14] A. Chandra, B. Singh, B. N. Singh, K. Al-Haddad, "An improved control algorithm of shunt active filter for voltage regulation, harmonic elimination, power-factor correction, and balancing of nonlinear loads," IEEE Trans. Power Electron., vol.15, no.3, May 2000, pp495-507.

Title: Ligand based virtual screening for identifying potent inhibitors against viral neuraminidase: An *in silico* approach

Author: Vinita Mishra Sangeeta Kashyap Yasha Hasija

PII: S1658-3655(14)00065-X
DOI: http://dx.doi.org/doi:10.1016/j.jtusci.2014.04.007
Reference: JTUSCI 85

To appear in:

Received date: 14-2-2014
Revised date: 2-4-2014
Accepted date: 2-4-2014

Please cite this article as: V. Mishra, S. Kashyap, Y. Hasija, Ligand based virtual screening for identifying potent inhibitors against viral neuraminidase: An *in silico* approach, *Journal of Taibah University for Science* (2014), http://dx.doi.org/10.1016/j.jtusci.2014.04.007

Ligand based virtual screening for identifying potent inhibitors against viral neuraminidase: An *in silico* approach

Vinita Mishra[1], Sangeeta Kashyap[2], Yasha Hasija[1,*]

1. Department of Biotechnology, Delhi Technological University (Formerly Delhi College of Engineering), Shahbad Daulatpur, Main Bawana Road, Delhi-110042, India

2. Department of Biotechnology, Meerut Institute of Engineering and Technology, Meerut, N.H.-58, Baghpat Road Bypass Crossing, Meerut-250 005, India

*Corresponding author's E-mail: yashahasija@gmail.com

ABSTRACT

Various inhibitors have been developed for neuraminidase but resistance against these drugs in many viral strains makes it an advantageous and interesting task to discover compounds which can be more promising in preventing viral infection through neuraminidase. Virtual screening methods have been proved as an efficient *in* silico approach for drug discovery processes. In the present study, we used ligand based virtual screening process for identifying potent inhibitors against viral neuraminidase enzyme. The approach utilized in this study has been successful in identifying 15 compounds which may be potential inhibitors of neuraminidase. These compounds were screened via three screening platforms (MVD, PyRx, Fred) by setting oseltamivir as reference compound, which is a FDA approved drug against influenza virus. These compounds were then filtered by their *in silico* ADME/T (**A**bsorption, **D**istribution, **M**etabolism, **E**xcretion, and **T**oxicity) values and only 12 of them were found to have comparatively better results. The results of the present study are reported herein so that researchers, who are having required laboratory facilities for synthesizing drugs, can utilize findings of this study for developing new drugs against influenza with better efficacy.

Keywords

Neuraminidase (NA); Haemagglutinin (HA); Neuraminidase inhibitors (NAIs); Molecular docking; ZINC database; Virtual Screening (VS); Molegro Virtual Docker (MVD); PyRx; FRED

1. INTRODUCTION

Some of the reported worst epidemics in the history of humans have been caused by influenza viruses. Influenza viruses contain two major glycoproteins attached on their surface: haemagglutinin (HA) and neuraminidase (NA). Haemagglutinin helps virus in the attachment and penetration of host cells via sialic acid binding sites [1, 2] and Neuraminidase enzymatically cleaves bonding between haemagglutinin and sialic acid from cell surface

glycoconjugates and thus facilitates the release of progeny virions from infected cells, spreads the new virus particles, and prevents the aggregation of progeny virions [3, 4] . According to different antigenic properties of haemagglutinin and neuraminidase molecules, influenza type A viruses can be classified into 16 subtypes by haemagglutinin (H1–H16) and 9 subtypes by neuraminidase (N1–N9) [5]. Numerous combinations of haemagglutinin and neuraminidase subtypes have been found in influenza type A viruses on avian species, i.e. bird flu. Among them, H5N1 has received extensive attention in recent years since it has caused a considerable number of human lives worldwide [6].

In theory, both haemagglutinin and neuraminidase can be considered as therapeutic targets for preventing the replication and spread of influenza viruses in host cells. Although the crystal structure of haemagglutinin was already resolved in early 1980s [7], no tightly binding compounds have been discovered for it. As for neuraminidase, many inhibitors with high potencies have been developed. For example, zanamivir and oseltamivir (Tamiflu) are two successful drugs currently in use [8, 9]. Nevertheless, resistance against these drugs has subsequently been developed by influenza viruses, still making the development of new classes of neuraminidase inhibitors a significant and urgent task [10, 11]. Neuraminidase was chosen as a suitable drug target because NA plays a major role in influenza virus propagation, and the amino acids residues of the active site interacting directly with the substrate or surrounding the central active site of the enzyme are strictly conserved [12].

Virtual Screening (VS) is considered as computational approach of High Throughput Screening (HTS) and refers to the *in silico* evaluation of properties of different molecular scaffolds such as binding affinity, interaction energy etc. Different applications of machine learning to virtual screening have been presented in the literature including both ligand-based similarity searching and structure-based docking. The main purpose of such applications is to prioritize databases of molecules as active against a particular protein target. Some case studies presented in the same perspective suggest that VS has already played a significant role in the discovery of some compounds that are now in the clinical trial or even in the market [13-18].

2. MATERIALS AND METHODS

2.1. Selection and Preparation of Target/ Receptor Protein for Docking

Target neuraminidase protein (PDB ID: 2HU0) was obtained from the Protein Data Bank (PDB) which is a repository for the 3-D structural data of large biological molecules [19]. This protein was selected as target because the observation of the open conformation for the 150-loop in the group-1 structures suggest that, for these enzymes, this conformation of the loop is intrinsically lower in energy than the closed conformation. Group-1 neuraminidases (N1 and N8) initially bind to oseltamivir in this open conformation but eventually

adopt the closed conformation. It thus seems that oseltamivir binding to group-1 neuraminidases favors the higher energy or closed conformation of the 150-loop that it probably accesses via a relatively slow conformational change. It should therefore be possible to design new inhibitors for group-1 neuraminidases that are selective for the open 150-loop conformation and would thereby have the potential to bind more strongly than oseltamivir or zanamivir [20].

2.2. Cavity Detection and Selection

Possible active site(s)/ cavities for the target protein (2HU0) of interest were determined with the help of Molegro Virtual Docker (MVD) [21]. A number of cavities were detected for the same protein but only one cavity was selected as active site on the basis of prior information about the active site residues, available in literatures. This cavity was utilized in further docking studies.

2.3. Ligand Screening From ZINC Database (Version - 12)

Ligands were screened from ZINC database on the basis of structural similarity with known neuraminidase inhibitors (oseltamivir, zanamivir, peramivir and laninamivir) [22-25]. 70% structural similarity was used as cut off for database screening. On the basis of oseltamivir, zanamivir, peramivir and laninamivir 30, 57, 3, and 58 analogs were found respectively. Out of 58 analogs of laninamivir 57 were common for both laninamivir and zanamivir and one was different (ZINC ID 71974042), common analogs were considered only once.

3. RESULTS AND DISCUSSION

3.1. Active Site Analysis

The active sites of all influenza neuraminidases contain three arginine residues – Arg118, Arg292 and Arg371 – that bind the carboxylate of the substrate sialic acid, one arginine residue, Arg152, interacts with the acetamido substituent of the substrate, and one glutamine residue, Glu276, forms hydrogen bonds with the 8-and 9-hydroxyl groups of the substrate [20]. Binding pocket of NA, as predicted by MVD and supported by various literatures, is given in Figure 1 along with their amino acid residues.

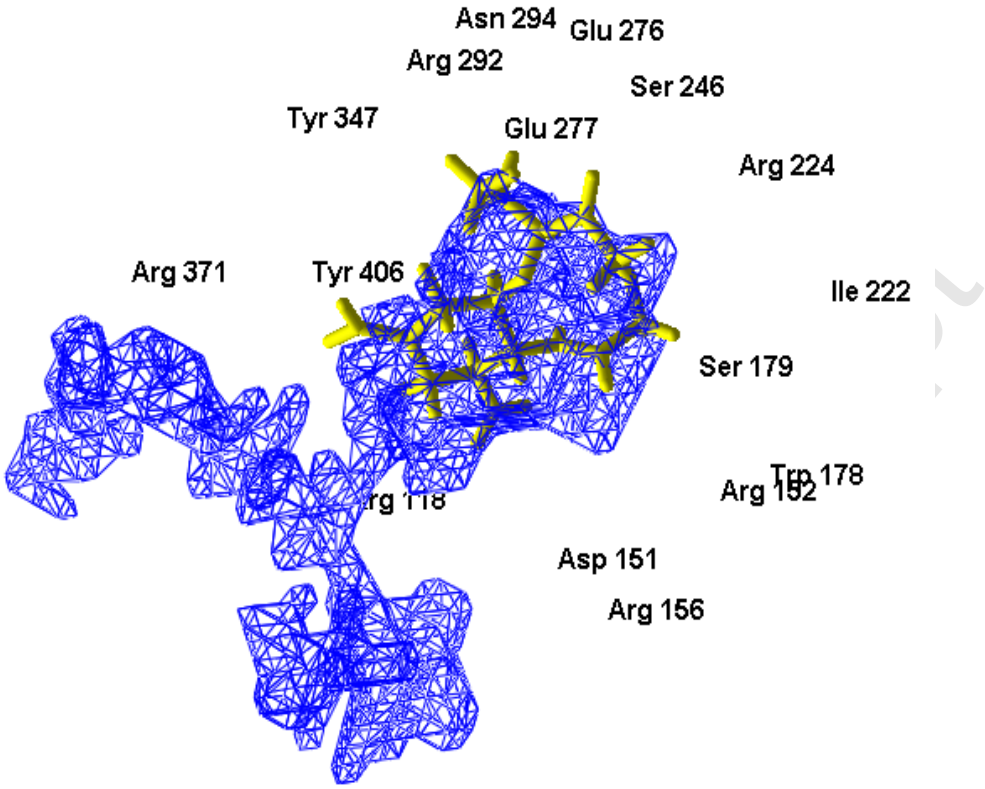


Figure 1: Active site of NA predicted by MVD (Active Site Residues: Arg118, Asp151, Arg152, Arg156, Trp178, Ser179, Ile222, Arg224, Ser246, Glu276, Glu277, Arg292, Asn294, Arg371, Tyr347, Tyr406).

3.2. Selection, Validation and Evaluation of Docking Protocol

In molecular docking, validation of docking protocol is a necessary step in order to ensure that ligands bind within the binding pocket in the correct conformation which is done by validating the size and center of the coordinates of the grid box across the binding pocket [26]. In this work, docking protocol was validated by redocking co-crystallized structure of neuraminidase in complex with oseltamivir (PDB ID: 2HU0). During redocking various algorithms were used which are available in MVD and it was observed that most suited algorithm was MolDock SE (Simplex Evolution) search algorithm. All binding conformations of redocked oseltamivir within the binding pocket of NA produced by MolDock SE algorithm were similar to binding mode of the co-crystallized ligand, and the root mean square deviation (RMSD) for these conformations were below 2 Å [27]. The RMSD of best pose was 1.75 Å, interaction energy (between target protein and ligand) was -100.69 kcal/ mol and hydrogen bond energy was –11.69 kcal/ mol, Figure 2.

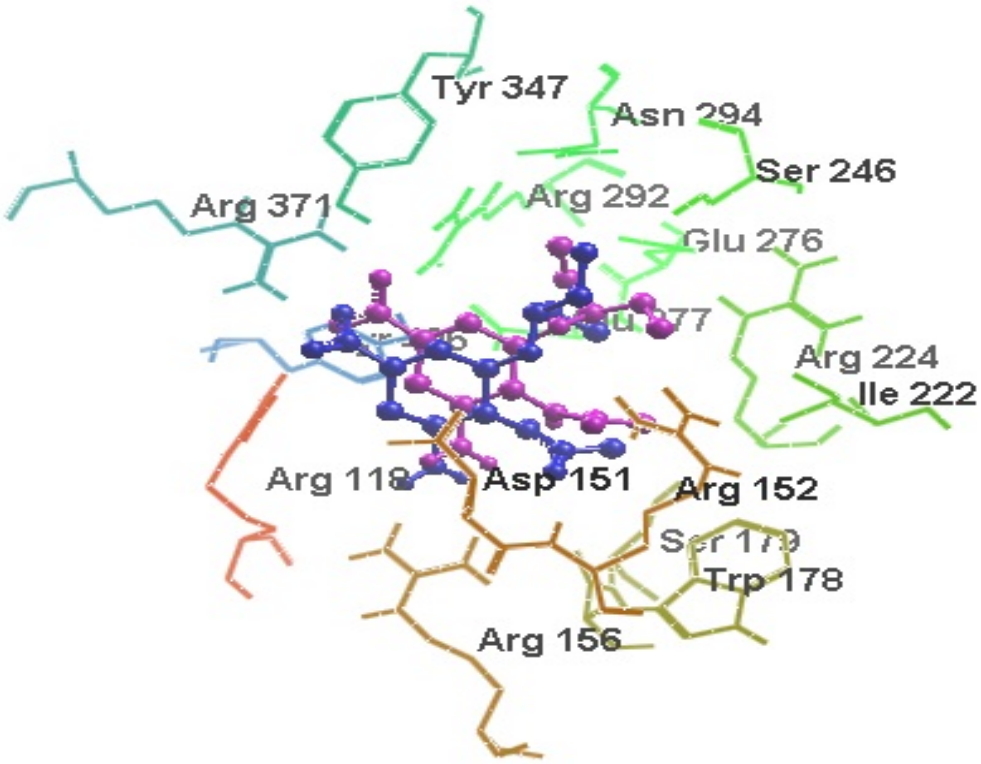
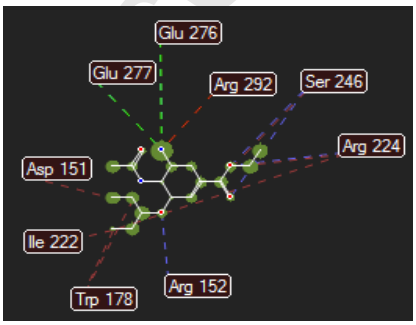


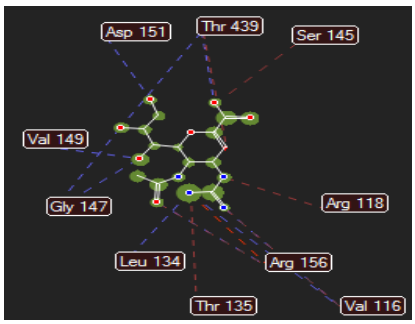
Figure 2: Superimposed image of highest ranking docked conformer (blue) and co-crystallized neuraminidase with oseltamivir (2HU0) (purple).

3.3. Analysis of Interactions and Binding Poses of Known Inhibitors

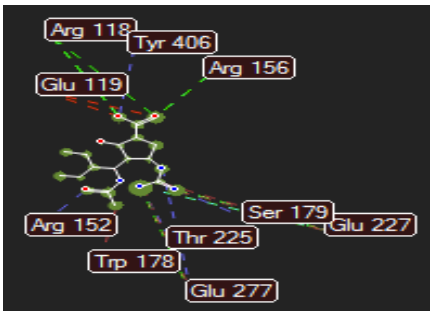
All the 4 known inhibitors were re-docked within the binding pocket of neuraminidase and their interactions were analyzed with the help of MVD, Figure 3(A)-(D).



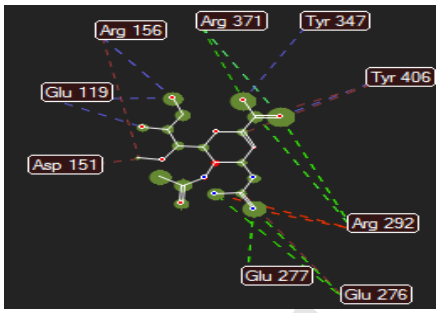
(A)



(*B)*



(C)



(*D)*

Figure 3 (A-D): The docked poses of Oseltamivir, Zanamivir, Peramivir, and Laninamivir respectively within the binding pocket of neuraminidase showing their different interactions (hydrogen bonds, electrostatic and steric interactions are shown in blue, green and red dotted lines respectively).

3.4. Docking Study for Screening of Unknown Inhibitors

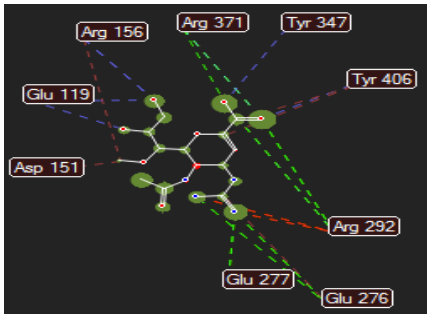
All the 91 different analogs of known inhibitors that were screened from ZINC database were docked within the binding pocket of neuraminidase with the help of Molegro Virtual Docker, PyRx (Autodock Vina) and FRED [21, 28-30]. Only those compounds were selected after screening that showed positive result from all the three docking platform and were having better binding affinity, docking score, similarity with the known inhibitors and other scoring functions (Table 1). Poses for the unknown inhibitors were evaluated on the basis of various factors such as by comparing their hydrogen bonding interaction patterns and overall interaction overlay with the known inhibitors, etc.

3.5. ADME/T Studies

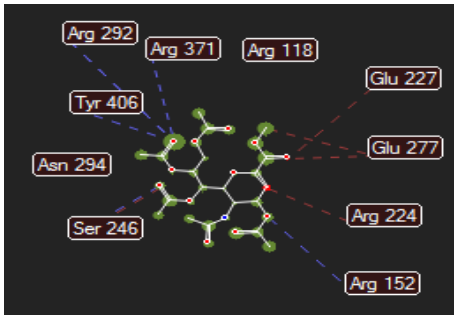
APOD (Abbreviated Profile of Drugs) calculates ADME/T values of a compound with respect to a reference compound (RC) on the basis of their molecular properties such as Molecular weight (MW), Hydrogen bond donor (HD), Hydrogen bond acceptor (HA), Lipophilicity (LP), and Polar solvent accessibility (PSA) and creates a comparative chart of ADME/T values for both the compounds. ADME/T (Absorption, Distribution, Metabolism, Excretion, and Toxicity) values were calculated with the help of APOD for all the 15 compounds that were screened after ligand based virtual screening and it was found that only 12 of them (Table 2) showed better results than oseltamivir (RC1) in terms of Distribution, Metabolism and Excretion, whereas Absorption and Toxicity values of these compounds were found same as of oseltamivir (RC1). Interaction maps of the compounds that showed better results after ADME/T filtering is shown in figure 4 A-L.

Table 2: Comparative chart for ADME/T values of screened compound along with reference compound oseltamivir (RC1)

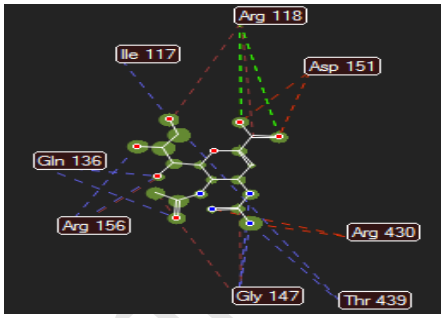
Ligand Name (ZINC ID)	**Properties of Ligand**					**ADME/T values**				
	MW	**HA**	**HD**	**LP**	**PSA**	**A**	**D**	**M**	**E**	**T**
RC1 (ZINC03929508)	313.418	6	4	0.85	93.85	1	3	5	3	0
VS2 (ZINC22047629)	473.431	13	1	0.12	171.39	1	4	5	4	0
VS3 (ZINC22047634)	473.431	13	1	0.12	171.39	1	4	5	4	0
VS4 (ZINC 26284236)	226.277	4	0	1.46	60.44	1	4	6	3	0
VS5 (ZINC 29559740)	346.34	11	8	-3.24	170.86	1	4	6	3	0
VS6 (ZINC 33676598)	473.431	13	1	0.21	171.39	1	4	5	4	0
VS7 (ZINC 33676599)	473.431	13	1	0.21	171.39	1	4	5	4	0
VS8 (ZINC 33676602)	473.431	13	1	0.21	171.39	1	4	5	4	0
VS9 (ZINC 35645304)	473.431	13	1	0.21	171.39	1	4	5	4	0
VS10 (ZINC 37033736)	332.313	11	9	-3.67	181.86	1	4	6	3	0
VS11 (ZINC 40641191)	332.313	11	9	-3.51	193.48	1	4	6	3	0
VS12 (ZINC 44136660)	430.41	12	3	-0.99	158.682	1	4	5	4	0
VS15 (ZINC 06692561)	290.248	9	5	-3.01	160.98	1	4	5	2	0



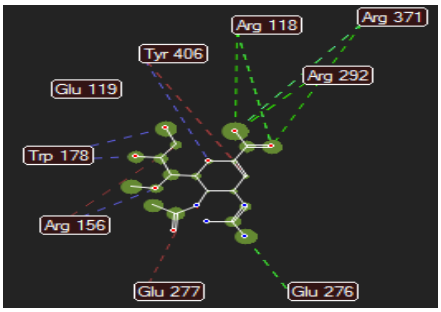
(A): VS2 (ZINC22047629) ligand



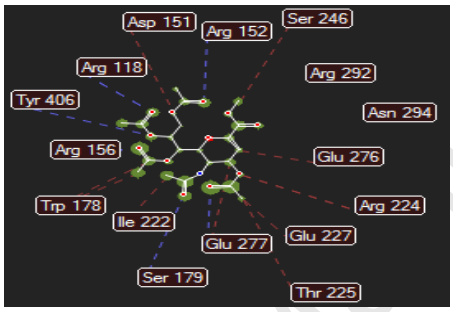
(B): VS3 (ZINC22047634) ligand



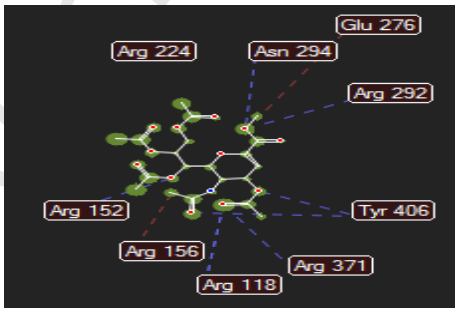
(C): VS4 (ZINC 26284236) ligand



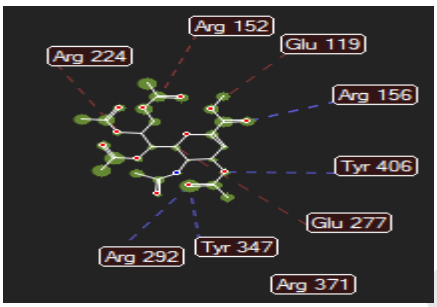
(D): VS5 (ZINC 29559740) ligand



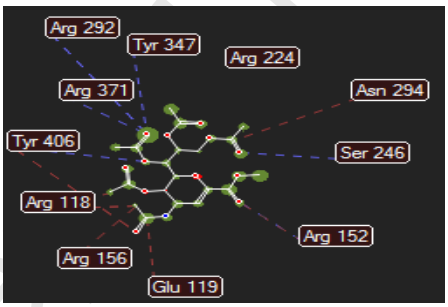
(E): VS6 (ZINC 33676598) ligand



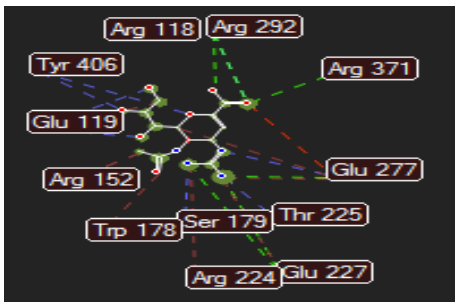
(F): VS7 (ZINC 33676599) ligand



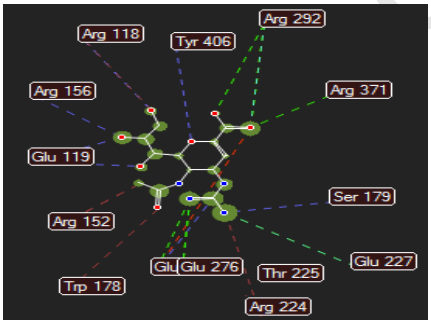
(G): VS8 (ZINC 33676602) ligand



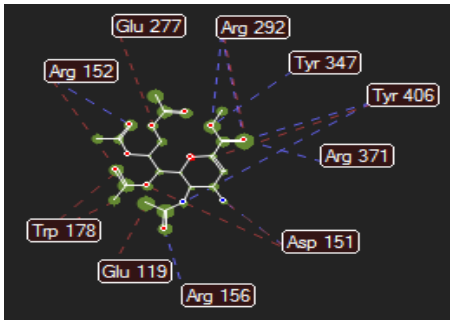
(H): VS9 (ZINC 35645304) ligand



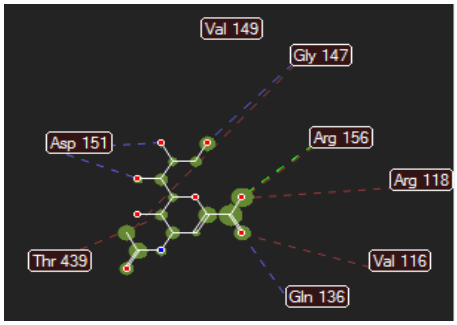
(I):VS10 (ZINC 37033736) ligand



(J):VS11 (ZINC 40641191) ligand



(K): VS12 (ZINC 44136660) ligand



(L): VS15 (ZINC 06692561) ligand

Figure 4 (A-L): The docked structures of virtually screened compounds after ADME/T filtration within the binding pocket of neuraminidase showing their overall interactions

4. CONCLUSIONS

Virtual screening methods are widely used for reducing cost and time of drug discovery process. In this study, we used only oseltamivir standards (various scores and binding energy) as a cutoff for screening of new and potential inhibitors because Oseltamivir is a FDA approved drug and was discovered via structure based approach [20, 31], and when we set Oseltamivir as a standard for screening other known inhibitors (Zanamivir, Peramivir, and Laninamivir) the possibility of success of this approach in identification of true positive inhibitors was 100%, because it was able to screen all the three other known inhibitors as true positive, whereas on using other inhibitors as standard the rate of success had decreased. This approach was successful in identifying 15 compounds which may behave as potential inhibitors. These compounds were screened via three screening platforms (MVD, PyRx, FRED). The docked poses of these compounds resembles similar orientation as observed with neuraminidase ligand (Oseltamivir). These ligands were docked deeply inside the binding pocket of NA forming interactions with ARG118, ASP151, ARG152, ARG156, TRP178, SER179, ILE222, ARG224, SER246, GLU276, and GLU277 [20]. These compounds were further filtered by their *in silico* ADME/T (absorption, distribution, metabolism, excretion, and toxicity) values and only 12 of them were found as comparatively better than or as potential as oseltamivir. Therefore, this study shows the importance of this hypothesis in screening of small molecule libraries and their use to intensify drug discovery process prior synthesis.

5. ABBREVIATIONS

VS: Virtual Screening, HTS: High Throughput Screening, RC: Reference Compound, NA: Neuraminidase, HA: Haemagglutinin, ADME/T: Absorption, Distribution, Metabolism, Excretion, and Toxicity, RMSD: Root Mean Square Distance, APOD: Abbreviated Profile of Drugs, LE1 (Ligand Efficiency 1): MolDock Score divided by Heavy Atoms count, LE3 (Ligand Efficiency 3): Rerank Score divided by Heavy Atoms count, HA: Hydrogen Bond Acceptor, HD: Hydrogen Bond Donor, MW: Molecular Weight, LP: Lipophilicity, PSA: Polar Solvent Accessibility.

REFERENCES

1. J.N.S.S. Couceiro, J.C. Paulson, L.G. Baum. 1993. Influenza virus strains selectively recognize sialyoligosaccharides on human respiratory epithellium, the role of the host cell in selection of hemaglutinin receptor specificity. Virus Res. 29: 155 - 165.

2. D.C. Wiley, J.J. Skehel. 1987. The structure and function of the hemagglutinin membrane glycoprotein of influenza virus. Annu. Rev. Biochem. 56: 365–394.
3. J.A. Griffin, S. Basak, R.W. Compans. 1983. Effects of hexose starvation and the role of sialic acid in influenza virus release. Virology. 125: 324–334.
4. C. Liu, M.C. Eichelberger, R.W. Compans, G.M. Air. 1995. Influenza type A virus neuraminidase does not play a role in viral entry, replication, assembly, or budding. J. Virol. 69: 1099–1106.
5. World Health Organization. 1980. A revision of the system of nomenclature for influenza viruses: a WHO memorandum. Bull. World Health Organ. 58: 585–591.
6. C. Bender, H. Hall, J. Huang, A. Klimov, N. Cox, A. Hay, V. Gregory, K. Cameron, W. Lim, K. Subbarao. 1999. Characterization of the surface proteins of influenza A (H5N1) viruses isolated from humans in 1997–1998. Virology. 254: 115–123.
7. W. Weis, J.H. Brown, S. Cusack, J.C. Paulson, J.J. Skehel, D.C. Wiley. 1988. Structure of the influenza virus haemagglutinin complexed with its receptor, sialic acid. Nature. 333: 426–431.
8. M. von Itzstein, W.Y. Wu, G.B. Kok, M.S. Pegg, J.C. Dyason, B. Jin, T.V. Phan, M.L. Smythe, H.F. White, S.W. Oliver, P.M. Colman, J.N. Varghese, D.M. Ryan, J.M. Woods, R.C. Bethell, V.J. Hotham, J.M. Cameron, C.R. Penn. 1993. Rational design of potent sialidase-based inhibitors of influenza virus replication. Nature. 363: 418–423.
9. C.U. Kim, W. Lew, M.A. Williams, H. Liu, L. Zhang, S. Swaminathan, N. Bischofberger, M.S. Chen, D.B. Mendel, C.Y. Tai, W.G. Laver, R.C. Stevens. 1997. Influenza neuraminidase inhibitors possessing a novel hydrophobic interaction in the enzyme active site: design, synthesis, and structural analysis of carbocyclic sialic acid analogues with potent anti-influenza activity. J. Am. Chem. Soc. 119: 681–690.
10. P.A. Reece. 2007. Neuraminidase inhibitor resistance in influenza viruses. J. Med. Virol. 79: 1577–1586.
11. G. Orozovic, K. Orozovic, J.D. Järhult, B. Olsen. 2014. Study of oseltamivir and zanamivir resistance-related mutations in influenza viruses isolated from wild mallards in sweden. 9 (2): e89306.
12. W.P. Burmeister, R.S. Daniels, S. Dayan, J. Gagnon, S. Cusack, R.W. Ruigrok. 1991. Sequence and crystallization of influenza virus B/Beijing/1/87 neuraminidase. Virology. 180: 266–272.
13. G. Melagraki, A. Afantitis, H. Sarimveis, P.A. Koutentis, J. Markopoulos, O. Igglessi-Markopoulou. 2007. A novel QSPR model for predicting _ (lower critical solution temperature) in polymer solutions using molecular descriptors. J. Mol. Model. 13: 55-64.
14. C.G. Gadhe, S.H. Lee, T. Madhavan, G. Kothandan, D. Choi, S.J. Cho. 2010. Ligand based CoMFA, CoMSIA and HQSAR analysis of CCR5 antagonists. Bull. Kor. Chem. Soc. 31: 2761-2770.

15. M. Goodarzi, M.P. Freitas, N. Ghasemi. 2010. QSAR studies of bioactivities of 1- (azacyclyl)-3-arylsulfonyl-1H-pyrrolo[2.; 3-b]pyridines as 5-HT6 receptor ligands using physicochemical descriptors and MLR and ANN-modeling. Eur. J. Med. Chem. 45: 3911-3915.
16. V.D. Mouchlis, T.M. Mavromoustakos, G. Kokotos. 2010. Molecular docking and 3D-QSAR CoMFA studies on indole inhibitors of GIIA secreted phospholipase A2. J. Chem. Inf. Model. 50: 1589-1601.
17. V.D. Mouchlis, T.M. Mavromoustakos, G. Kokotos. 2010. Design of new secreted phospholipase A2 inhibitors based on docking calculations by modifying the pharmacophore segments of the FPL67047XX inhibitor. J. Comput. Aided. Mol. Des. 24: 107-115.
18. V. Mishra, C.V.S.S Prasad. 2011. Ligand based virtual screening to find novel inhibitors against plant toxin Ricin by using the ZINC database. Bioinformation. 7: 46-51.
19. H.M. Berman, K. Henrick, H. Nakamura, J. Markley, P.E. Bourne, J. Westbrook. 2007. Realism about PDB. Nature Biotechnology. 25: 845-846.
20. R.J. Russell, L.F. Haire, D.J. Stevens, P.J. Collins, Y.P. Lin, G.M. Blackburn, A.J. Hay, S.J. Gamblin, J.J. Skehel. 2006. The structure of H5N1 avian influenza neuraminidase suggests new opportunities for drug design. Nature. 443: 45-49.
21. R. Thomsen, M.H. Christensen. 2006. MolDock: A New Technique for High-Accuracy Molecular Docking. J. Med. Chem. 49: 3315-3321.
22. H. Ikematsu, N. Kawai, N. Iwaki, S. Kashiwagi. 2014, In vitro neuraminidase inhibitory activity of four neuraminidase inhibitors against influenza virus isolates in the 2011-2012 season in Japan. J. Infect. Chemother. 20 (2): 77-80.
23. C. Tran-To Su, X. Ouyang, J Zheg, C.K. Kwoh. 2013. Structural analysis of the novel influenza A (H7N9) viral Neuraminidase interactions with current approved neuraminidase inhibitors Oseltamivir, Zanamivir, and Peramivir in the presence of mutation R289K. BMC Bioinformatics. **14** (Suppl 16): S7.
24. Y. Yu, S. Garg, P.A. Yu, H.J. Kim, A. Patel, T. Merlin, S. Redd, T.M. Uyeki. 2012. Peramivir use for treatment of hospitalized patients with influenza A (H1N1) pdm09 under emergency use authorization, October 2009-June 2010. Clin Infect Dis. 55 (1): 8-15.
25. M. Kakuta, S. Kubo, M. Tanaka, S. Tobiume, T. Tomozawa, M. Yamashita. 2013. Efficacy of a single intravenous administration of laninamivir (an active metabolite of laninamivir octanoate) in an influenza virus infection mouse model. Antiviral Res. 100 (1): 190-5.
26. S.V. Lim, M.B.A. Rahman, B.A. Tejo. 2011. Structure-based and ligand-based virtual, screening of novel methyltransferase inhibitors of the dengue virus. BMC Bioinformatics. **12** (Suppl 13) (2011) S24.
27. C.P. Vianna, W.F.d. Azevedo. 2012. Identification of new potential Mycobacterium tuberculosis shikimate kinase inhibitors through molecular docking simulations. J. of Mole. Model. 18: 755-764.

28. R.B. Jacob, T. Andersen, O.M. McDougal. 2012. Accessible high-throughput virtual screening molecular docking software for students and educators. PLoS Comput Biol. 8 (5): e1002499.
29. M. McGann. 2012. FRED and HYBRID docking performance on standardized datasets. J. Comp.-Aid. Mol. Design. 26(8): 897-906.
30. M. McGann. 2011. **FRED Pose Prediction and Virtual Screening Accuracy.** *J. Chem. Inf. Model. 51 (3):* 578-596.
31. C.U. Kim, W. Lew, M.A. Williams, H. Liu, L. Zhang, S. Swaminathan, N. Bischofberger, M.S. Chen, D.B. Mendel, C.Y. Tai, W.G. Laver, R.C. Stevens. 1997. Influenza neuraminidase inhibitors possessing a novel hydrophobic interaction in the enzyme active site: design, synthesis, and structural analysis of carbocyclic sialic acid analogues with potent anti-influenza activity. J. Am. Chem. Soc. 119: 681–90.

Table 1: Virtually screened compounds on the basis of various scoring functions

Ligand Name	ZINC ID	E-Inter (protein – ligand Interaction energy in KJ/Mol)	Energy (KJ/Mol) or MolDock Score	HBond Energy (KJ/Mol)	LE1	LE3	MW	Rerank Score (arbitrary units)	Steric Interaction (KJ/Mol)	VdW (LJ12-6) (KJ/Mol)	Similarity (w.r.t Oseltamivir)	No of H-Bonds	Binding Affinity (PyRx) (Kcal/Mol)	Pose score (FRED)
RC1	ZINC03929508	-125.705	-104.972	-11.1686	-4.77146	-3.4799	313.412	-76.5578	-106.988	-17.9726	32.0808	7	-6.5	-26.88
RC2	ZINC03918138	-140.869	-120.679	-20.0388	-5.24689	-3.84836	332.31	-88.5123	-121.083	-37.8168	18.6352	10	-7.2	-34.30
RC3	ZINC03981610	-126.69	-125.005	-15.4094	-5.43498	-3.76797	328.407	-86.6634	-96.7167	-15.7241	21.2183	9	-6.7	-34.26
RC4	ZINC03985629	-127.088	-103.138	-25.5124	-4.2974	-2.96609	346.336	-71.1861	-84.5985	5.40812	33.1139	13	-7.2	-30.97
VS1	ZINC22012077	-132.346	-127.995	-25.9599	-5.33314	-4.25503	338.402	-102.121	-106.386	-35.0829	20.6627	14	-6.7	-33.00
VS2	ZINC22047629	-146.301	-121.596	-8.52048	-3.68472	-2.24388	473.428	-74.0479	-137.781	-1.00072	23.528	4	-7.1	28.79
VS3	ZINC22047634	-151.225	-136.077	-21.2061	-4.12354	-2.93814	473.428	-96.9585	-130.019	-40.8673	27.4237	11	-7.4	-28.84
VS4	ZINC26284236	-130.935	-121.342	-17.9446	-5.27575	-3.54091	332.31	-81.441	-108.726	3.32461	18.7158	11	-7.0	-34.42
VS5	ZINC29559740	-125.638	-114.955	-22.3746	-4.7898	-4.18879	346.336	-100.531	-90.4354	-30.5518	22.693	9	-7.2	-30.00
VS6	ZINC33676598	-152.037	-151.737	-22.3478	-4.59809	-2.77607	473.428	-91.6103	-129.689	-1.22205	43.5508	12	-6.6	-27.55
VS7	ZINC33676599	-155.43	-142.93	-18.8045	-4.3312	-2.97458	473.428	-98.1612	-136.626	-45.4967	35.0065	9	-7.1	-30.14
VS8	ZINC33676602	-148.948	-136.9	-14.2166	-4.14848	-3.02718	473.428	-99.8971	-134.731	-44.0123	27.7475	10	-7.4	-27.38
VS9	ZINC35645304	-140.523	-123.43	-19.629	-3.7403	-1.72349	473.428	-56.8752	-120.894	38.7106	18.3914	11	-6.8	-28.34
VS10	ZINC37033736	-133.937	-116.808	-29.2808	-5.07863	-3.58839	332.31	-82.533	-90.8218	6.01711	21.5714	16	-6.9	-32.62
VS11	ZINC40641191	-134.246	-123.034	-29.5452	-5.34932	-3.01232	332.31	-69.2834	-92.6921	34.3449	20.395	14	-6.8	-33.16
VS12	ZINC44136660	-138.916	-132.874	-24.0434	-4.42914	-3.38108	430.406	-101.432	-114.873	-42.5319	22.5911	11	-7.3	-33.06
VS13	ZINC65739846	-127.233	-114.735	-7.59441	-5.46358	-4.61402	297.373	-96.8943	-109.426	-31.7841	25.2443	5	-6.6	-33.19
VS14	ZINC65739850	-127.434	-115.528	-15.1424	-5.50135	-3.71704	297.373	-78.0578	-102.989	-4.83545	21.8996	9	-6.5	-35.68
VS15	ZINC06692561	-127.657	-101.861	-15.4738	-5.09303	-4.82693	290.247	-96.5387	-108.535	-34.8038	24.9761	7	-6.7	-32.36

Materials Letters ∎ (∎∎∎∎) ∎∎∎–∎∎∎

Contents lists available at ScienceDirect

Materials Letters

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

Microencapsulated cardanol derived benzoxazines for self-healing applications

Pratibha Sharma [a,b], Swapnil Shukla [c], Bimlesh Lochab [c,**], Devendra Kumar [b], Prasun Kumar Roy [a,*]

[a] Centre for Fire, Explosive and Environment Safety, DRDO, Timarpur, Delhi 110054, India
[b] Department of Applied Chemistry and Polymer Technology, Delhi Technological University, Delhi 110042, India
[c] Department of Chemistry, School of Natural Sciences, Shiv Nadar University, Gautam Buddha Nagar 203207, Uttar Pradesh, India

ARTICLE INFO

Article history:
Received 12 May 2014
Accepted 6 July 2014

ABSTRACT

Benzoxazine (Bz) monomer was synthesized from renewable cardanol, a by-product from cashew-nut industry, using a solventless approach. The monomer was encapsulated in poly(styrene) (PS) shells by solvent evaporation technique to obtain spherical microcapsules. The microcapsule dimensions and core content could be tailored by optimizing the operating parameters, particularly stirring speed and PS concentration. Successful demonstration of this simple and versatile methodology widens the scope for large-scale application of benzoxazines in the field of self-healing.

1. Introduction

Polybenzoxazines belong to a class of phenolic thermosetting resins, which exhibit interesting properties, mainly zero-volume shrinkage, low water absorption, high thermal stability and chemical resistance, high glass transition temperature (T_g), long shelf life and self-curing characteristics. Benzoxazine precursors hold enormous potential as healing agents in self-healing compositions [1], provided they are encapsulated in fragile microcontainers, prior to their inclusion in the polymeric matrix. Surprisingly, attempts towards usage of oxazine monomers for this purpose have not been reported till date. Presently, dicyclopentadiene (DCPD) is the most common healing agent used for self-healing applications, which undergo ring opening metathesis polymerization in the presence of organometallic catalysts [2]. Although, compositions based on encapsulated DCPD have reached the stage of commercial maturity, in view of the exceptionally high cost of the organometallic catalyst, it is desirable to explore alternate healing agents.

Benzoxazines, interestingly, do not require any catalyst for their curing, and hence are ideal candidates for the aforementioned purpose. What makes these materials even more attractive is the possibility of deriving these monomers from renewable sources like cardanol; a by-product of the cashew nut processing industry [3–5]. Benzoxazines undergo thermally activated ring-opening polymerization (ROP) reaction at elevated temperatures and it is possible to alleviate the curing conditions by making structural modifications. Recently, Lochab et al. have successfully reduced the Bz–C curing temperature from 242 to 161 °C by blending with oxazines containing higher and acidic functionalities [6]. Other functional groups such as allylic [7], methylol [8] etc. has also been reported to lower the curing temperature.

Unfortunately, the possibility of curing of Bz monomers under acidic conditions imposes a restriction on the technique used for its encapsulation. The conventional micro-encapsulation procedure involves in-situ emulsion polymerization of urea or melamine with formaldehyde on the surface of the hydrophobic dispersed phase of the precursor [9]. The condensation process requires low pH (<3); [2,10] but the acidic conditions may induce curing in Bz [11]. We believe that this issue can be addressed by encapsulating Bz in soluble polymers by solvent evaporation, which surprisingly has not been attempted till date.

In this paper, we demonstrate a technique for physical encapsulation of cardanol derived Bz in PS shells. We adopt a "green" approach for the synthesis of cardanol based benzoxazines (Bz–C), where the role of cardanol is extended to that of a reactive diluent. This is followed by its encapsulation in PS microcapsules. Of particular interest is to investigate the effect of operating parameters to arrive at optimized conditions for preparation of microcapsules.

2. Preparation benzoxazine monomers from cardanol

Benzoxazine monomer was prepared from cardanol as per the procedure reported previously [3]. In brief, a mixture of cardanol

* Corresponding author. Tel.: +911123907191; fax: +911123819547.
** Corresponding author. Tel.: +120 2663801.
E-mail addresses: bimlesh.lochab@snu.edu.in (B. Lochab), pk_roy2000@yahoo.com, prasunroy2000@gmail.com (P. Kumar Roy).

http://dx.doi.org/10.1016/j.matlet.2014.07.048
0167-577X/

(100 g, 0.33 mol), paraformaldehyde (19.8 g, 0.66 mol) and aniline (30.1 mL, 0.33 mol) was slowly heated and maintained under isothermal conditions at 80 °C and 90 °C for 1 h and 2 h respectively. Post-cooling, 500 mL of water was added and the organic layer was collected after extraction with chloroform (2 × 100 mL). After drying over sodium sulphate, and solvent removal under reduced pressure, the residue was dried at 70 °C under vacuum to yield Bz–C in quantitative yield as a red brown oil.

Microencapsulation of benzoxazine in poly(styrene): Bz–C was encapsulated in PS microcapsules by solvent evaporation technique using the procedure reported in the literature [12]. PS was prepared by emulsion polymerization process and the properties are presented in the supplementary section [13]. A solution of benzoxazine monomer in chloroform (20% w/v) was added drop wise to 100 mL aqueous PVA solution (2.5% w/v) under continuous stirring (500 rpm). Separately, a solution of PS in chloroform (2–10% w/v) was injected through a hypodermic syringe into the reaction vessel, maintained at 60 °C, at different stirring rates (400–600 rpm). Post-evaporation of chloroform, the reaction mixture was cooled, and the Bz–C encapsulated PS microcapsules were filtered and washed repeatedly with water followed by drying under vacuum.

Detailed characterization of the monomer and microencapsulated PS is presented in the supplementary section.

3. Results and discussion

Bz–C was prepared using solventless method, followed by its encapsulation in PS microcapsule. The effect of operating parameters, particularly stirring speed and polymer concentration on the microcapsule dimensions were investigated.

Preparation of cardanol based benzoxazine: Conventionally, the synthesis of benzoxazine monomers requires polar aprotic solvents, which have to be removed subsequently. Solvent-less synthesis of highly rigid Bz monomers results in high viscosity at higher percentage conversion which may account for incomplete conversion of reactants and poor yields. However, the low viscosity of cardanol (145 mPa s) facilitates its use as a reactive diluent and the reaction of aniline with cardanol and formaldehyde (1:1:2), post purification led to the formation of a viscous liquid (80% yield). The structure of the monomers was confirmed by ^{1}H-NMR (Fig. S1, supplementary section) and FTIR (Fig. S2, Supplementary section) spectroscopy. The ^{1}H –NMR spectra of Bz–C exhibited characteristic resonances at ~5.3 ppm (s, $ArOCH_2N$), ~4.6 ppm (s, $ArCH_2N$) suggesting conversion of hydroxyl functionalities to oxazines. The formation of oxazine ring was also substantiated by the appearance of new absorption bands at ~1250 and ~1030 cm^{-1} due to the Ar–C–O oxazine asymmetric and symmetric stretch respectively. The absence of absorption bands due to N–H stretching (3360–3442 cm^{-1}) and N–H bending (1619 cm^{-1}) in the spectra of Bz–C further suggests the absence of unreacted aniline in the Bz monomer indicating completion of reaction.

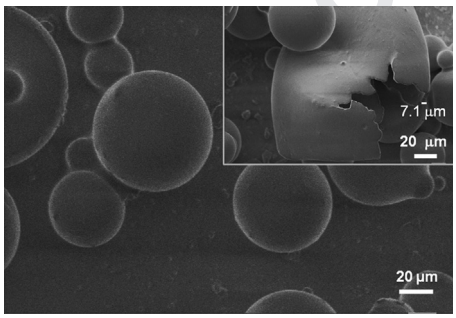


Fig. 1. SEM image of Bz–C encapsulated PS microcapsules. Inset shows the magnified image of a broken microcapsule indicating its shell wall thickness.

Cardanol based benzoxazines exhibit excellent adhesive properties, as indicated by lap shear strength (LSS) as high as 20–30 kg/cm^2 at 150 °C [3], which advocate their potential towards self-healing applications. The benzoxazine moieties underwent thermally activated ROP to form polymer networks containing > N- and –OH functionalities which result in extensive H-bonding with the matrix/surface and translate to strong adhesive properties.

Physical microencapsulation of benzoxazine: Bz–C was encapsulated in fragile PS containers by solvent evaporation technique to form pale yellow microcapsules, the SEM images of which are presented in Fig. 1. The microcapsules are perfectly spherical, and exhibit a smooth surface texture. The thickness of the shell wall, was found to be ~7 ± 1 μm, as can be seen in the SEM images of broken microcapsules (Fig. 1, Inset).

Effect of stirring speed: The effect of increasing stirring speed on the microcapsule dimensions is presented in Fig. 2. As expected, increasing the rate of stirring led to a decrease in the particle dimensions, which could be attributed to the shearing of the large oily droplets into smaller microspheres under higher shear rates [14,15].

Effect of concentration of encapsulating polymer: Core-content is one of the most important characteristics of microcapsules. In view of the similar solubilities of Bz–C and PS, conventional technique of core-content quantification proved unsuccessful and the ratio of heat of curing (DSC technique) of encapsulated benzoxazines to that of the neat benzoxazines was used to determine the same. The core content was found to increase from 19% to 38% by decreasing the PS content in the feed solution from 10% to 2% w/v, at a constant stirring speed of 500 rpm. However, the core content could not be increased further, due to the fragile nature of the shell, which appears to be incapable of cementing higher amounts of healing agent.

The TGA traces of the resultant microcapsules in air atmosphere are presented in Fig. S7. Neat PS exhibits a mass loss of ~98% at 500 °C. The thermal decomposition of PS reportedly

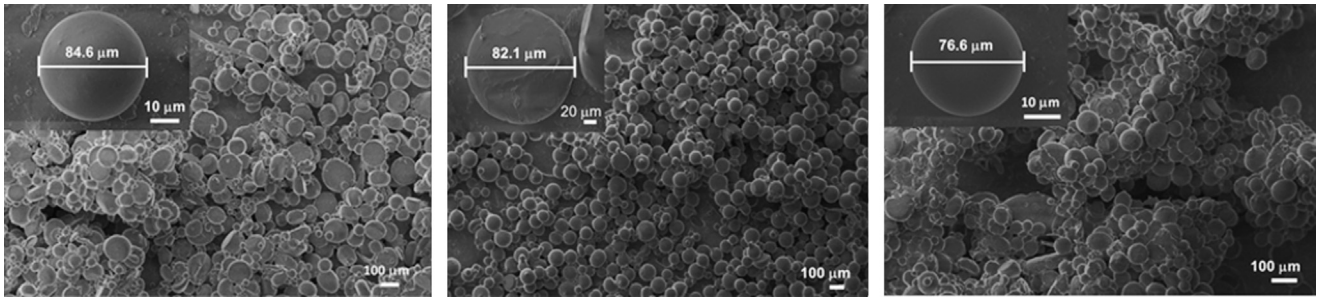


Fig. 2. SEM image of microcapsules prepared under different stirring speeds (a) 400, (b) 500, (c) 600 rpm. Inset shows the enlarged image of a single microcapsule. (Label a b and c in picture).

Please cite this article as: Sharma P, et al. Microencapsulated cardanol derived benzoxazines for self-healing applications. Mater Lett (2014), http://dx.doi.org/10.1016/j.matlet.2014.07.048

occurs in a single step via unzipping reaction, while polybenzoxazines exhibit a bimodal mass loss distribution centered at 448 and 558 °C [3]. Similar decomposition behaviour was observed in all the encapsulated microcapsules, suggesting polymerisation of the Bz–C during heating.

To demonstrate the potential of the Bz–C encapsulated microcapsules in the field of self healing, ~5 mg of microcapsules were placed between two glass slides, crushed by manual tapping and placed at ~240 °C for 30 min. Photographs, captured at different stages of experimentation are presented in the supplementary section. The glass slides were found to seal completely, and could not be separated by exposure to solvents even for extended periods (72 h).

It is to be noted that the system presently investigated necessitate elevated temperatures for curing. However, in view of possibility of structural modification by judicious choice of functional groups, low temperature self healing systems can be devised, which can open up novel opportunities in this field.

4. Conclusion

The low viscosity of cardanol facilitated its use as reactive diluent for the synthesis of Bz–C, which was subsequently encapsulated in PS by solvent evaporation technique. Microcapsules with dimensions ranging from 75 to 86 μm could be prepared by varying the stirring speed from 600 to 400 rpm. By decreasing the concentration of PS in the feed solution, it was possible to increase the core content as high as 38% (w/w).

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.matlet.2014.07.048.

References

[1] Taskin OS, Kiskan B, Yagci Y. Polybenzoxazine precursors As self-healing agents for polysulfones. Macromolecules 2013;46:8773–8.
[2] Brown EN, Kessler MR, Sottos NR, White SR. In situ poly(urea-formaldehyde) microencapsulation of dicyclopentadiene. J Microencapsul 2003;20:719–30.
[3] Lochab B, Varma I, Bijwe J. Thermal behaviour of cardanol-based benzoxazines. J Therm Anal Calorim 2010;102:769–74.
[4] Lochab B, Varma I, Bijwe J. Cardanol-based bisbenzoxazines. J Therm Anal Calorim 2012;107:661–8.
[5] Lochab B, Shukla S, Varma IK. Naturally occurring phenolic sources: monomers and polymers. RSC Adv 2014 (10.1039/C4RA00181H).
[6] Lochab B, Varma I, Bijwe J. Blends of benzoxazine monomers. J Therm Anal Calorim 2013;111:1357–64.
[7] Oie H, Sudo A, Endo T. Acceleration effect of N-allyl group on thermally induced ring-opening polymerization of 1,3-benzoxazine. J Polym Sci, Part A: Polym Chem 2010;48:5357–63.
[8] Baqar M, Agag T, Ishida H, Qutubuddin S. Methylol-functional benzoxazines as precursors for high-performance thermoset polymers: unique simultaneous addition and condensation polymerization behavior. J Polym Sci, Part A: Polym Chem 2012;50:2275–85.
[9] Tripathi M, Rahamtullah Roy PK, Kumar D, Rajagopal C. Influence of microcapsule shell material on the mechanical behavior of epoxy composites for self-healing applications. J Appl Polym Sci. 2014;10:1002.
[10] Rochmadi AP, Hasokowati W. Mechanism of microencapsulation with urea-formaldehyde polymer. Am J Appl Sci. 2010;7:739–45.
[11] Andreu R, Reina JA, Ronda JC. Carboxylic acid-containing benzoxazines as efficient catalysts in the thermal polymerization of benzoxazines. J Polym Sci, Part A: Polym Chem 2008;46:6091–101.
[12] Sánchez L, Sánchez P, Lucas A, Carmona M, Rodríguez J. Microencapsulation of PCMs with a polystyrene shell. Colloid Polym Sci. 2007;285:1377–85.
[13] Roy PK, Rawat AS, Rai PK. Synthesis, characterisation and evaluation of polydithiocarbamate resin supported on macroreticular styrene–divinylbenzene copolymer for the removal of trace and heavy metal ions. Talanta 2003;59:239–46.
[14] Chaudhary S, Parthasarathy S, Kumar D, Rajagopal C, Roy PK. Amine functionalised poly(styrene) microspheres as thermoplastic toughener for epoxy resin. Polym Compos 2014. http://dx.doi.org/10.1002/pc.22927.
[15] Roy P, Iqbal N, Kumar D, Rajagopal C. Polysiloxane-based core-shell microspheres for toughening of epoxy resins. J Polym Res 2014;21:1–9.

Please cite this article as: Sharma P, et al. Microencapsulated cardanol derived benzoxazines for self-healing applications. Mater Lett (2014), http://dx.doi.org/10.1016/j.matlet.2014.07.048

Volume 2 (2014) 505-510 ISSN 2347 - 3258
International Journal of Advance Research and Innovation

Modeling &Simulation of Nano film thickness of Gold deposited by Thermal Evaporation Process (TEP)

R. S. Mishra *, Shailendra Kumar Gaur
Department of Mechanical Engineering, Delhi Technological University, Delhi, India

Article Info
Article history:
Received 2 March 2014
Received in revised form
10 April 2014
Accepted 30 May 2014
Available online 15 June 2014

Keywords
Physical Vapor Deposition,
Modeling,
Film Thickness Simulation,
Gold Evaporation,
Thin Film Deposition

Abstract
The surface temperature is required to specifying the temperature of the evaporating gold using constant elements for turn off the refinement in the post-processing settings. The paper represents the modeling of nano scale gold film by computing the film thickness, mass deposited on the substrate and mass transfer rate with time dependent model using BDF solver. Gold is evaporated from a resistively heated evaporator source at a temperature of 2000K onto a surface held on a fixed surface. The film thickness varies between 34nm to 39 nm across the sample after 60 sec of deposition, with radial symmetry about the midpoint of the source. The film thickness as well as mass deposited at a point increases linearly with time. Since the angular distribution is of particular interest in this model, by increasing the integration resolution to a maximum value for ensuring the most accurate angular resolution when computing the flux.

1. Introduction

The nanoscale gold film thickness has found newer capabilities in different fields of science and technology such as coating glasses/mica to change their properties and many coloured optical coding for biological assays Gold nanometer thick film is being used to enhance electroluminescence and quantum efficiency in light emitting diodes. Besides signal amplification, nanometer thick gold film evoluted new types of new sensors that are capable of detecting very small amounts of analytes such as chemical vapours in the scale of few ppm. Adeleh Granmayeh Rad et al [1] discussed the usage of gold nano film is in making advanced dyes and pigments. Sometimes gold nano film has been used to dye textiles, support to supply clean energy (by solar cell) and high density data storage (flash memories and discs Gold nanoscale film quality of non-toxic and biocompatibility both in vivo and in vitro environments makes useful in biomedical applications. The unique properties of gold nanoscale structured materials supply good opportunity for coordinating biological identification events with electronic signal transduction and for making an evolution of bioelectronics devices with newer features. Gold in nanoscales have been found for potential candidates to support in photo-thermal therapy and radiotherapy. Optical and electronic properties of gold can be utilized to improve the contrast in molecular imaging for the detection of cancer initially. Gold-based technologies also help to facilitate a ultimate needle-free delivery system, a technique that used gold and permitted vaccines to be supplied through the skin producing use of the fact that small particles can go through gaps between cells while large ones cannot. Gold nanoscale-based technologies give solution to some of environmentally great issues, such as ecofriendly production methods, pollution control and water purification. No doubt gold is really one of the non-reactive metals, and it is resistant to oxidation.). Ragini Raj Singh et al [8] discussed gold is also used in Photovoltaic HgCdTe Mid wavelength Infra-red detector to make ohmic contacts to find out the passivation characterstics of CdS/HgCdTe structure using C-V measurements.

Ali Moarrefzadeh [2] discussed that Physical vapor deposition (PVD) includes a wide range of vacuum coating processes in which material is physically taken out from a source by evaporation or sputtering, transported through a vacuum or partial vacuum by the energy of the vapor particles, and condensed as a film on the surfaces of appropriately placed parts or substrates. A group of very versatile coating processes in which a material is converted to its vapor phases in a vacuum chamber and condensed onto a substrate surface as a very thin film(upto 1µm thickness). The deposition of thin film layers from the vapor phase is done through several methods.

Gold deposition depending upon the different applications as well as the suitability of the process requirement and availability of resources can be done in following ways:

Wenjun Zhou et.al [4] discussed the Chemical vapour deposition (CVD) generally uses a gas-phase precursor, often a halide or hydride of the element to be deposited. In the case of MOCVD, an organometallic gas is used. Commercial techniques often use very low pressures of precursor gas. The chamber is first evacuated to a vacuum of around 5m torr and substrate temperature fixed to 350°C. Now evaporator is resistively heated to a temperature of about 160°C to start volatilization of the precursor. The deposited precursor then thermally decomposes into pure gold by a reaction above 300°C. The thickness of the gold film can be controlled by adjusting the amount of precursor, the temperature of the deposition chamber, and the location of the samples to be coated relative to the inlet of the evaporator into the deposition chamber. Generally gold precursor is thermally stable between 150 and 300 °C. Keewah Chan [5] discussed that gold can be deposited on silicon or quartz by rf sputtering by using DC sputter coating system. The rf power can of several hundred watt, substrate temperature about 300 °C

Corresponding Author,
E-mail address: professor_rsmishra@yahoo.co.in

with a vacuum of 1mbar order or better for time depending upon the thickness required, say for nanometer thickness 1 hour.

Carl E. Larson [6] discussed that gold deposition of high purity gold onto various substrates from dimethyl-2fl-pentandionato gold (III), Me2Au (acac), by localized, laser-induced (photothermal) chemical vapor deposition. CVD of gold from dimethyl-(1,1,1-trifluoro-2,4-pentandionato) gold (III) and dimethyl-(1,1,1,5,5,5-hexafluoro-2,4-pentandionato) gold (III), Me2Au(tfac) and Me2Au(hfac) complex can be done in a stainless steel vacuum chamber evacuated by a turbomolecular/diffusion/cryo pump to a base pressure of < 10 -6 torr. A flow of argon or nitrogen carrier gas through a Pyrex vessel containing the gold complex at room temperature is used to deliver the organogold precursor into the chamber through stainless steel lines. Chamber vacuum and carrier gas flow rate could be varied independently. Deposition proceeds on a substrate placed in the chamber in the flow path of the carrier gas and heated by a copper block heater with temperature controller. Miroslav Gojo et al [7] discussed that gold can be deposited in a thermostatically controlled electode cells in a electrodeposition Cell has two openings, one for thermometer another for nitrogen inlet for purging the oxygen out of the electrolyte keeping pH value 5 to 7 and each silicon wafer piece coated and alloyed with gold as working electrode, filled with chosen electrolyte warmed upto a given temperature and purged with nitrogen. Gold deposited by electrochemical method has higher purity, lower permeability and good adhesion to the substrate. Anne-Felicie Lamic-Humblot et al[10] discussed that gold can be very easily deposited at ambient pressure and in distilled water with magnetic rod stirring , followed by thermal treatment from Tetrachloroauric acid and urea. This paper mainly deals with simulation & modeling of time dependent PVD model.

Though a lot of research work has been done on the evolution of gold film deposited on nanoscale or higher thickness by different deposition process through different sources (resistively, sputtering, E-beam, magnetron, radio frequency) in different baskets/crucibles at different temperatures and modeled mostly with direct simulation Monte Carlo. Also these methods solve in the volumes of the modeled and geometries. Number densities found to be not accurate and precise. Ineligible simulation for the accurate modeling of low pressure, low velocity gas flows in complex geometries. These studies subjected to statistical scatter. Moreover, these could not completely explained free molecular flow interface and applied dsmc computes the trajectories of large numbers of randomized particles through the system, but introduces statistical noise to the modeling process and also the method is slower. The objectives of paper is to explain the free molecular flow interface in low pressure low velocity gas flow using simulation and time dependent modeling to influence of evaporation process parameters.

2. Experimental Study of Thermal Evaporation Process

We study and analyze the physical vapor deposition techniques and equipments that are in common use in the large scale production of coatings that find uses in the optical, display, decorative, tribological, and energy-generating /saving industries. Evaporating materials are classified as dielectric compounds, metals, alloys, or mixtures. The same evaporant material can exhibit different optical, electrical, and mechanical properties depending on the deposition process.Titanium oxide is a unique example of a metal oxide compound that, depending on deposition process parameters, can be made into film layers that are: transparent, electrically conductive, chemically reactive to light and bio- agents, chemically inert, or exhibit spectrally selectively absorption. The dependent parameters are starting composition, oxidation state, and crystalline structure and packing density.

PVD techniques used generally are basically two in nature: thermal evaporation by resistively heating or by using an electron- beam heating, and sputtering, a no thermal process. Alterations and accompaniments are made to the basic thermal evaporation technique to permit different coating materials and substrate types to be included.

Process additions designed to alter the growth nano-structure or composition of the film through control of the dependent variables listed above include bombardment of the growing film by high energy inert- or / and reactive ions, substrate heating, atmosphere composition and partial pressure, rate, and vapor incidence angle. A further important variable contribution to the nucleation and self-assembling growth structure of the condensing ad atoms, that we have discussed frequently, is the condition both chemical and physical of the substrate surface.

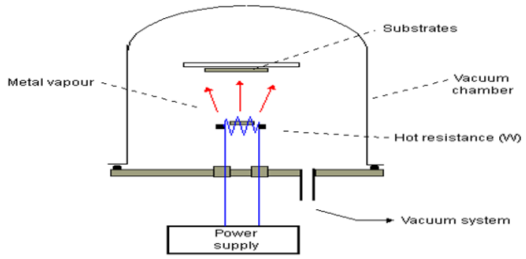


Fig: 1. Vacuum Coating Process Diagram

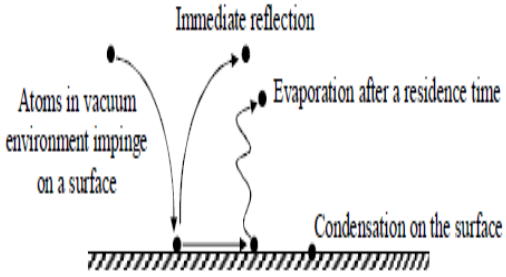


Fig: 2. Thermal Evaporation Process

Varieties of coatings can be deposited such as metals, alloys, ceramics and other inorganic compounds, and even certain polymers. Deposition can be done onto the varieties of substrates such as metals, glass, and plastics. The properties of atavistically deposited films depend strongly on:

- The material being deposited
- Substrate surface chemistry
- and morphology
- The surface preparation process
- The details of deposition process
- and the deposition Parameters
- mean free path of evaporant
- source to substrate distance
- vapor pressure
- sticking coefficient

Condensation and nucleation- Atoms that impinge on a surface in a vacuum environment may be reflected immediately, re-evaporate after a residence time, or condense on the surface shown in figure 2. Mean free path is the minimum distance between two successive collisions and is given by formula depending upon the vacuum in the chamber:

$$\lambda = \frac{k_B T_c}{\sqrt{2}\pi \cdot P \cdot d^2}$$

Where T – usually room temprature of chamber
P – Vacuum in chamber
At 300 K
d - diameter of vapor atom ~2-5 Å

$$\lambda = \frac{5x10^{-3}}{P_T\,(Torr)}\,cm$$

Based on this mean free path source to substrate distance is optimized for constant flux and better deposited film quality. Vapor pressure varies depending upon the source temperature and is different for different materials, for suitability standard plot [11] is provided from which depending upon source material vapor pressure is selected. Sticking coefficient is defined as the ratio of the condensing atoms to impinging atoms. If the atoms do not immediately react with the surface, they will have some degree of surface mobility over the surface before they condense.

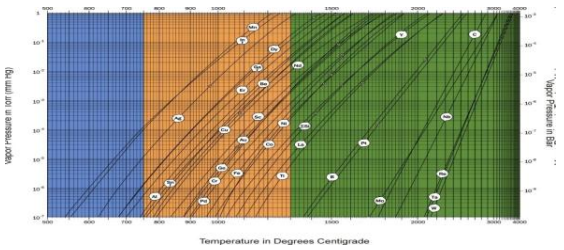


Fig: 3. Graph between Vapor Pressure of Source Vs Temperature

Re-evaporation is a function of bonding energy between the adatom and the surface, the surface temperature, and the flux of mobile adatoms. Example: The deposition of cadmium on a steel surface having a temperature greater than about 200 ˚C will result in total re-evaporation of the cadmium.

2.1 Surface Mobility

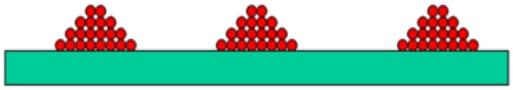
The energy of the atom, atom-surface interaction (chemical bonding) and the temperature of the surface influence the capability of an atom on a surface. The mobility on a surface can change because of variations in chemistry or crystallography. The various crystallographic plates of a surface have different surface free energies that influence the surface diffusion.

Atoms condense on a surface by losing energy. They lose energy by:

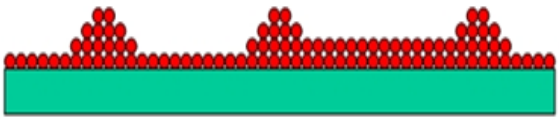
- Forming and breaking chemical bonds with the substrate atoms.
- Finding preferential nucleation sites (lattice defects, atoms steps, and impurities)
- Interacting with another diffusing surface atoms (same species)
- Colliding or reacting with adsorbed surface species.Atoms form nuclei after condensation. Homogenous nucleation is said to occur if the surface is of same material as the deposition atoms and if they are different materials, the process is called heterogeneous. In semiconductor field, heterogeneous nucleation forms heterojunctions. G.H. Gilmer et al [3] discussed that **t**hree types of nucleation mechanisms have been identified; they differ according to nature of interaction between the posited atoms and the substrate material:

Frank-Van der Merwe mechanism leading to a monolayer- by-monolayer growth (layer growth; ideal epitaxy)

Volmer-Weber (V-W) mechanism, characterized by a three-dimensional nucleation and growth (island growth)



Stranski-Krastanov (S-K) mechanism, where an altered surface layer is formed by reaction with the deposited material to generate a strained or pseuomorphic structure, followed by nucleation on this layer (Layer + island growth)



Nuclei coalescence and agglomeration the nuclei develop by collecting atoms that diffuse over the surface. Isolated nuclei grow laterally and vertically on the surface to form a continuous film and less amount of material is needed if nucleation density is high. The major growth mode of nuclei may be: -laterally over the substrate surface (wetting growth) such as gold on copper and chromium, iron on W-O surfaces, and titanium on SiO2 the nuclei may prefer to grow in a vertical mode (dewetting growth) such as nickel and copper on W-O surfaces, and gold on carbon, Al2O3, and SiO2. M.Levin et al [9] discussed that the substrate temperature also critical in thin film deposition.

2.2 PVD Film Evaluation

Deposited films are typically evaluated for visual defects, thickness, and adhesion. Visual defects such as bare spots, small voids, incorporated flakes, or debris can be observed with a stereo microscope having a magnification of 10 to 100 times. Film thickness is generally measured by one of the following methods: Thickness of film as modeled can be simulated experimently with dektek surface profilometer based on LVDT principle with diamond size of 12.5µm radius stylus. Polished metallurgical micro sections are used to microscopically observe the coating thickness on various part surfaces. This method is the most direct way to determine thickness uniformity. · Beta (high-energy electron) backscatter instruments are used to measure the film thickness nondestructively. This is an indirect method that requires calibration with a known standard; substantial errors can be made in measuring the film thickness on curved surfaces if care is not exercised. · A ball-crater instrument can be used to polish through the surface of a coating. The relationships between the diameters of the polishing ball, the maximum diameter that shows the effects of polishing, and the diameter of the substrate area that is exposed by polishing is used to calculate the thickness. Coatings that are up to 120 m-in. (3Mm) thick can be measured with an accuracy of±4 m-in. (±0.1 Mm) without difficulty on relatively smooth, flat or cylindrical surfaces. The adhesion between coating and substrate is difficult to measure directly for highly adherent films; pull tape tests capable of measuring yield strengths that are typical of metals and PVD hard coatings on metals have not been developed. In stone abrasion test, a fine sharpening stone is rubbed back and forth across the coated surface, allowing the stone particles to make grooves in the surface by nonrealistic deformation. The film is then inspected under a microscope to obtain adhesion information.

3. Experimental Observations

The island gold films were prepared using a laboratory thermal evaporation setup working at residual vacuum of (2.5 - 4) × 10-6 mbar. The deposition setup was equipped with the two-stage vacuum system based on the diffusion pump (oil based), turbomolecular or cryopump alongwith rotary vane pump or rotary screw pump (oil free). In case of oil based pumps the ultimate vacuum achieved in the system depends upon the vapor pressure of the oil used in pumps, generally Silicone based oils are used to achieve vacuum of1.0 x 10-6 order in the chamber. Sometimes high discharge Ar gas plasma substrate cleaning can be done at vacuum of 5 x 10-3 mbar, if needed. M.Levlin et al [9] discussed that annealing before evaporation affects the surface of substrate and removed the adsorbed molecules. The films were deposited on the substrates cleaned in the ultrasound bath in isopropyl alcohol and drained by a compressed air flux. During the deposition, all substrates were kept at room temperature i.e no substrate heating or cooling. Gold films were evaporated at residual vacuum from tungsten baskets at temperature of 2000˚ K. Deposition was provided up to full evaporation of material from the tungsten boat. Portions of the material for evaporation were prepared with help of the microbalance ViBRA.Value of this mass was found by calculation from the defined (nominal) film thickness using following relation for the point evaporation source:

$t=M\cos\theta/4\pi\rho r^2$

Here t is the nominal film thickness, M is the mass of the evaporated material, ρ is the material density, r is the distance between an evaporation source and a substrate, and θ is the deposition angle defined by geometry of the substrate. It should be noted, that this formula can be used only for rough estimation of an average thickness of ultra-thin films. In this case, certain assumptions about the structure of the film, the shape and size of the islands, can be made only on the basis of measurement of electrical and optical properties of this film and topographic surveillance using SEM or AFM. Optical absorption and transmittance of gold films on glass substrates were measured in wavelength range from 200 to 1100 nm using the 640 FT IR sepctrometer. F.Sharipov et al [12, 13] discussed that flow regimes are categorized quantitatively via the Knudsen number (Kn), which represents the ratio of the molecular mean free path to the flow geometry size for gases:

Table: 1. Classification of Flow Regimes

Flow type	Knudsen Number
Continuum flow	Kn<0.01
Slip flow	0.01<Kn<0.1
Transitional flow	0.1<Kn<10
Free molecular flow	Kn>10

4. Model Description

General description of the model is presented in fig. 1 where low pressure gas flow in vacuum system i.e molecular flow and it can be seen that the model inputs are ambient temperature, evaporant(gold) temperature, vapour pressure, molecular weight, density of gold, neglecting the substrate surface temperature. The Free Molecular Flow interface uses the angular coefficient method to model flows with Knudsen numbers Kn 10. This physics interface avoids solving the physics in the volumes of the modeled geometries, and requires meshing only of the surfaces. Completely diffuse scattering (total accommodation) and emission are assumed at all surfaces in the geometry, and flow is computed by integrating the flux arriving at a surface from all other surfaces in its line-of-sight. This means that the dependent variables exist only on the surfaces of the geometry, and the solution process is much faster than the dsmc method. In this model the number density and pressure of the deposited species are of limited interest and will not be computed. Furthermore, it is not subject to statistical scatter. Number densities are reconstructed using a method included in the Free Molecular Flow interface. The Molecular Flow Module is designed to offer previously unavailable simulation capabilities for the accurate modeling of low pressure, low velocity gas flows in complex geometries. It is ideal for the simulation of vacuum systems including those used in semiconductor processing, particle accelerators and mass spectrometers. Small channel applications (e.g. shale gas exploration and flow in nanoporous materials) can also be addressed. The Molecular Flow Module uses a fast angular coefficient method to simulate steady-state free molecular

flows. We can model isothermal and non-isothermal molecular flows, and automatically calculate the heat flux contribution from the gas molecules. The discrete velocity method is also included in the module for the simulation of transitional flows. Historically, flows in this regime have been modeled by the direct simulation Monte Carlo method. This computes the trajectories of large numbers of randomized particles through the system, but introduces statistical noise to the modeling process. For low velocity flows, such as those encountered in vacuum systems, the noise introduced by dsmc renders the simulations unfeasible. COMSOL uses alternative approaches: employing a discrete velocity method for transitional flows (using a Lattice Boltzmann velocity quadrature) and the angular coefficient method for molecular flows. Using the input parameters, the model computes the thickness of gold deposited, mass deposited on the surface of substrate, and mass transfer rate by using COMSOL software. The following wall conditions in terms of wall, outgassing wall, adsorption/desorption, deposition are inbuilt in COMSOL software: In adsorption/desorption boundry condition, sticking coefficient can be defined along with other condition. Gold is placed in resistively heated Tungstun boat which is having very high melting point of 3420^0 C. Substrate is one quarter of a 4□ wafer mounted on stationary support on top of tungsten boat depending upon the mean free path and Langumuire-Kundsen relation. A screen is placed to cover the substrate, if more than one source is used.

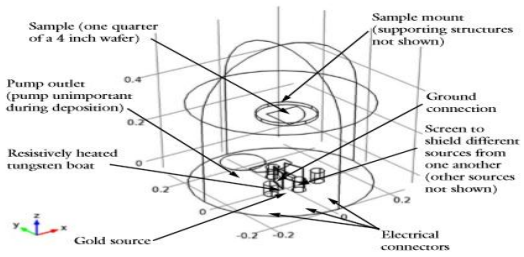


Fig: 4. Model Geometry Various Components of Evaporato

5. Result & Discussion

In table 2 various input parameter required for the model are given to compute the molecular flow in vacuum system. Generally, the ambient temperature is fixed based on the experimental setup or experimental conditions but the evaporation temperature can be varied as per the thickness or mass transfer to be deposited on the substrate. This evaporation temperature can be increased or decreased depending considering the vapor pressure, to make fast the mass transfer keeping the good adhesion with uniformity on the substrate surface.

Figure 5 shows the flux of gold molecules on the surfaces of the model. This constant flux determines the thickness of gold deposited shown in figure 6. In figure 6, it is obvious the film thickness is maximum in the centre of substrate positioned directly on vertically centre of tungsten boat/source, greater than 38.5 nm after 60 seconds of thermal evaporation process. This thickness decreases radialy outwards surface of the substrate 34.28 nm.

Table: 2. Input Parameter for the Model

Name	Expression	Value	Description
Tamb	293.15[K]	293.2 K	Ambient temperature
Tevap	2000[K]	2000 K	Evaporation temperature
pvap	50[Pa]	50.00 Pa	Vapor pressure of gold
Mn0	197[g/mol]	0.1970 kg/mol	Molecular weight of gold
Rho0	19.3 gc/m^3	1.930E4 kg/m³	Density of Gold

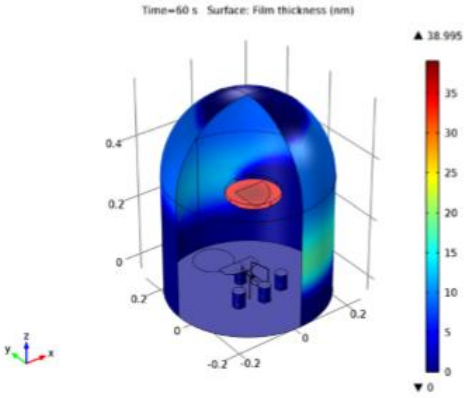


Fig: 5. Film Thickness on the Surfaces, After 60 s of Deposition

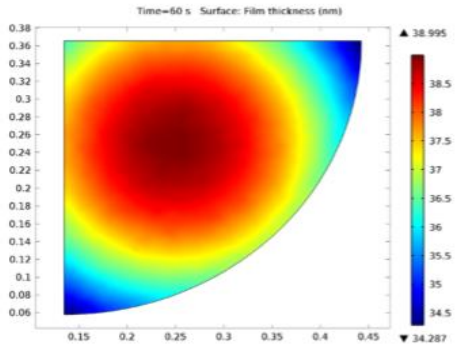


Fig: 6. Film Thickness on the Sample, After 60s of Deposition

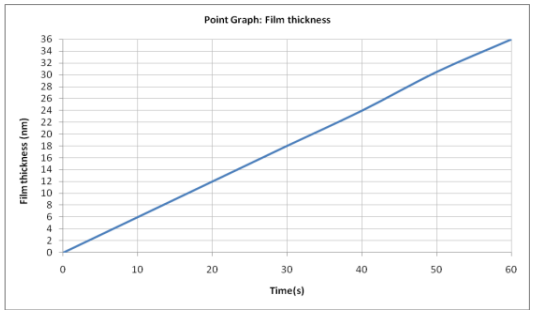


Fig: 7. Time Dependent Variation of Film Thickness

In Figure 7 based the molecular flux deposited on substrate ,graph shows the gold deposited thickness after 60 seconds from initial condition of zero second. It is clear from graph that the film thickness varies linearly with time.

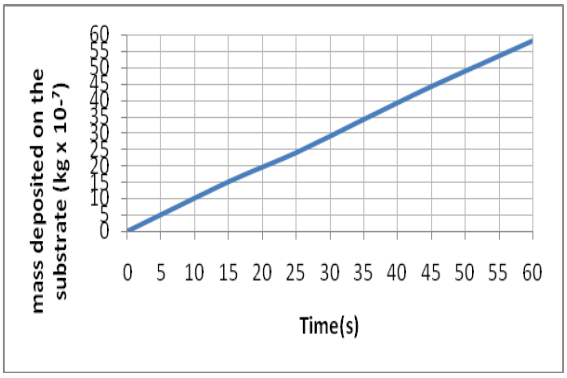


Fig: 8. Time Dependent Variation of Mass Deposited on the Substrate with Time (Sec.)

In figure 8 graph is plotted with time for mass (kg x E-7) deposited on the surface of substrate and shows that as mass deposited on substrate increases with time as in the case of figure 7 since the depositing thickness with time adds mass also on the substrate. In figure 9 mass transfer rate (kg x E-4/sec) is plotted for 60 seconds and is a straight constant abscissa parallel line. Only the flux is required to compute the deposition rate, but in this instance, since most of the computational time is used to compute the view factors, solving the time dependent problem adds little additional time to the solution process. Using a time dependent model also allows for more advanced extensions of the model, for example, re-evaporation of gold from hot surface close to the evaporative source could be included.

Certainly, loss of gold is there during thermal evaporation process for 60 seconds. It is based on the simple yield phenomenon of difference between the input gold put on boat, deposited on substrate and residual gold in the tungsten boat. This shows indirectly how much efficient is our process to achieve the desired target with a way to find out further reduce this loss by establishing the controlled process parameters.

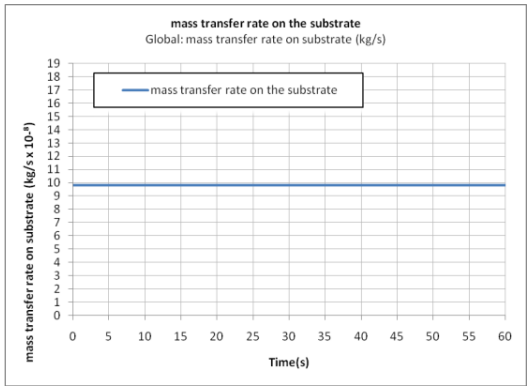


Fig: 9. Variation of Mass Transfer Rate with Time

6. Conclusion

The deposited film thickness on the substrate and the chamber wall is computed. For optimized method as well as for fast and accurate simulations, gases at low pressures cannot be modeled using conventional computational fluid dynamics tools. That is due to the fact that kinetic effects become important as the mean free path of the gas molecules becomes comparable to the length scale of the flow. Coatings of gold having good uniformity & appropriate thickness have been produced on stationary, non-rotated, flat substrates using a vacuum thermal evaporation technique. The thickness uniformity around the substrate was a sensitive function of the mean free path as well as evaporant incidence angle. Also to increase the rate of evaporation we have to raise the temperature of the source.

The following conclusions can be obtained:

i) The gold film thickness increases with time
ii) The mass deposited increases with time
iii) Thickness variation across the substrate surface follows cosine distribution due to incident angle

References

[1] Adeleh Granmayeh, et.al., Gold Nanoparticles: Synthesising, Characterizing and Reviewing Novel Application in Recent Years, Physics Procedia 22, 2011, 203 – 208
[2] Ali Moarrefzadeh, Simulation and Modeling of Physical Vapor Deposition (PVD) Process Wseas Transactions on Applied and Theoretical Mechanics, 7(2), 2012
[3] G.H. Gilmer et.a., Thin film deposition: fundamentals and modeling ,J of Computational Materials Science 12, 1998, 354-380
[4] Wenjun Zhou, Effective Permittivity of Ultrathin Chemical Vapor Deposited Gold Films on Optical Fibers at Infrared Wavelengths, J. Phys. Chem. C, 118, 2014, 670−678
[5] Keewah Chan, et.al., Formation of gold nanoparticles in silicon suboxide films prepared by plasma enhanced chemical vapour deposition J of thin solid films, 519, 2011, 4952-4957
[6] Carl E. Larson, et.al., Chemical Vapor Deposition of Gold, IBM Almaden Research Center, San Jose, California 95120, 1986
[7] Miroslav Gojo, et.al, Electrochemical Deposition of Gold in Citrate Solution Containing Thallium, Acta Chim. Slov, 55, 2008, 330–337
[8] Ragini Raj Singh, et.al, Investigation of passivation processes for HgCdTe/CdS structure for infrared application, 510, 2006, 235–240
[9] M. Levlin, et al, Evaporation of gold thin films on mica: effect of evaporation parameters, 115, 1997, 31–38
[10] Anne-Felicie Lamic-Humblot, et.al, An easy way to obtain thin film on silica glass substrate by chemical method, 539, 2013, 151–153
[11] Authur K.Burak Ucer, Vacuum evaporation, www.users.wfu.edu
[12] F.Sharipov, Universidade Federal do parana, Curitiba, 81531-900, Brazil http://fisica.ufpr.br/sharipov
[13] Yu. A. Anikin et. al(2012) Development of applied software for analysis of gas flows in vacuum devices, 86, 1770-1777

Hindawi Publishing Corporation
International Journal of Vehicular Technology
Volume 2014, Article ID 495036, 10 pages
http://dx.doi.org/10.1155/2014/495036

Research Article

A Probabilistic Analysis of Path Duration Using Routing Protocol in VANETs

Ram Shringar Rao,[1] Sanjay Kumar Soni,[2] Nanhay Singh,[1] and Omprakash Kaiwartya[3]

[1] *Ambedkar Institute of Advanced Communication Technologies & Research, Delhi 110031, India*
[2] *Delhi Technological University, Delhi 110042, India*
[3] *Jawaharlal Nehru University, New Delhi 110067, India*

Correspondence should be addressed to Ram Shringar Rao; rsrao08@yahoo.in

Received 28 February 2014; Revised 20 June 2014; Accepted 5 July 2014; Published 17 July 2014

Academic Editor: Rakesh Mishra

In recent years, various routing metrics such as throughput, end-to-end delay, packet delivery ratio, path duration, and so forth have been used to evaluate the performance of routing protocols in VANETs. Among these routing metrics, path duration is one of the most influential metrics. Highly mobile vehicles cause frequent topology change in vehicular network environment that ultimately affects the path duration. In this paper, we have derived a mathematical model to estimate path duration using border node-based most forward progress within radius (B-MFR), a position based routing protocol. The mathematical model for estimation of path duration consists of probability of finding next-hop node in forwarding region, estimation of expected number of hops, probability distribution of velocity of nodes, and link duration between each intermediate pair of nodes. The analytical results for the path duration estimation model have been obtained using MATLAB. The model for path duration estimation has been simulated in NS2. Each of the analytical results has been verified through respective simulation results. The result analysis clearly reveals that path duration increases with the increase in transmission range and node density and decreases with the increase in the number of hops in the path and velocity of the nodes.

1. Introduction

The intelligent transport system (ITS) has been working to make the road safer and efficient to cope up with increasing number of on-road vehicles day by day. The number of accidents on the roads is continuously increasing due to high growth in on-road vehicle population. The increasing number of accidents has become an issue of concern worldwide. It has made the roads vulnerable and threatening as every year millions of people are dying in accidents throughout the world. A modern network concept, VANETs, has become the hope for providing safer and well-organized transportation in near future [1].

In VANETs, routing is the process of finding optimal path for information forwarding between source and destination node. Position-based routing protocols have become one of the most investigated choices among researches due to geographic region sharing of on-road vehicles [2–7]. In position-based routing, forwarding of information is performed either through direct communication or through intermediate nodes between source and destination node. The intermediate nodes are high speed moving vehicles in VANETs that act as router during forwarding of information from source to destination. Due to ad hoc network architecture, no fixed infrastructure is considered during information forwarding in VANETs [8]. The mobility of nodes is one of the most critical factors in design, analysis, and performance evaluation of information forwarding techniques in VANETs. High mobility of nodes causes dynamic changes in the network topology that ultimately results in infrequent path failure. The communication between source and destination node is frequently interrupted due to path failure. Once a path failure occurs, a new path has to be set up for further communication. The frequent path failures degrade the performance of routing protocol and add overhead in terms of establishing a new path [9].

In VANETs, any routing path is made up of one or more links between pair of intermediate nodes. Therefore, lifetime of a link is one of the most important contributors in path duration. Path duration can be defined as the duration of time till every link of the route is active. In VANETs, the lifetime of a link is a random variable whose probability distribution depends on mobility, node density, transmission range, different traffic scenarios, and various impairments of radio communications. The links between intermediate nodes frequently break due to the high mobile nodes moving out of each other's transmission range. Therefore, estimation of path duration between source and destination decreases the chances of path breakage. The effective use of knowledge of path duration improves the performance and efficiency of routing protocols in VANETs [10].

In this paper, our main contribution is in terms of probabilistic and mathematical analysis of path duration using border-node based most forward progress within radius (B-MFR) routing protocol. For this analysis, a mathematical model to estimate path duration has been developed. B-MFR is a position-based routing protocol which selects next-hop node from the nodes belonging to the border area of the transmission range. The mathematical model for path duration estimation consists of the following modules. (1) The probability of finding at least one node in the considered forwarding region has been mathematically derived. (2) Average number of hops between source and destination has been estimated. (3) The probability distribution function for relative velocity between any two nodes has been derived. (4) Link duration between any two nodes has been obtained. Analytical results for the path duration estimation model have been obtained using MATLAB. The model for path duration estimation has been simulated in NS2 also. Each of the analytical results has been verified through respective simulation results.

The rest of the paper is organized as follows. In Section 2, the two communication modes in VANETs have been described. Section 3 presents related works in detail. In Section 4, all the mathematical formulations for the proposed path duration estimation have been presented. In Section 5, the simulation results and analysis have been discussed. Finally, we concluded the work presented in the paper in Section 6.

2. Communication Modes in VANETs

Vehicular communications in on-road traffic environments have been realized through VANETs. In this network, vehicles termed as nodes share information through wireless communication. Dedicated short range communications (DSRC) [11] technology is used for wireless communication. DSRC is an enhanced version of Wi-Fi technology specially designed for VANETs environment and this is known as wireless access in vehicular environment (WAVE). During the communication, when a sender node does not find any neighboring nodes, it forwards the information using road side units (RSUs) available along the road. However, the availability of RSUs is not strictly considered in VANETs. Thus, the communications in VANETs can be categorized in the following two modes: (1) vehicle-to-vehicle (V2V) communication and (2) vehicle-to-roadside (V2R) communication [12].

2.1. V2V Communication. V2V communication is the basic and primary aim in VANETs. It is pure ad hoc communication between two vehicles [13]. V2V communication can be through direct link or through multihop links (see Figure 1). If the destination node is present within the transmission range of the source node then the direct link is established for communication and this type of communication is known as single-hop communication. If the destination node is present outside the transmission range of the source node then the intermediate nodes are used to deliver the message up to the destination and this type of communication is known as multihop communications. V2V communication is mainly used for the safety applications such as road blockade alarm, electronic brake warning, incoming traffic warning, vehicle stability warning, lane change warning, and collision warning. This type of communication is also used for the different types of the protocol operations. To set up RSUs such as fixed infrastructure access points, internet gateways, and base station on the road side is expensive. Therefore, VANETs should use V2V communication as much as possible for communication purpose.

2.2. V2R Communication. V2R communication is the combination of ad hoc network and fixed infrastructure networks [14]. This mode of communication (as shown in Figure 1) involves on-road vehicles as well as RSUs. Only single-hop communication between a vehicle and RSU is used in V2R communication. Further, vehicle sends the message to the road side unit which broadcasts the message to all the vehicles in the neighborhood. Generally, RSUs use links of higher bandwidth for communication and broadcasting. RSUs may be placed at every one kilometer or less to enable and maintain high data rate in highly dense traffic environment.

3. Related Work

In spite of the fundamental importance of estimating the path duration of communication links in VANET, there have been some mathematical and experimental studies in MANETs. The estimation of path duration in MANETs is proposed using several theoretical and analytical models. Many research work and models are also proposed for the implementation and improvement of the VANET. Some of the related research works about the path duration have been carried out in the recent past decade.

The path duration is an important design parameter for the better performance and routing decision in VANETs. Authors in [15] show the analysis of path duration and provide different parameters related and dependent on the path duration in MANETs. Authors also present the path duration impact on the reactive routing protocols. The result shows that the path duration probability density function (*pdf*) for large number of hops can be estimated with the help of exponential distribution. The path duration depends on

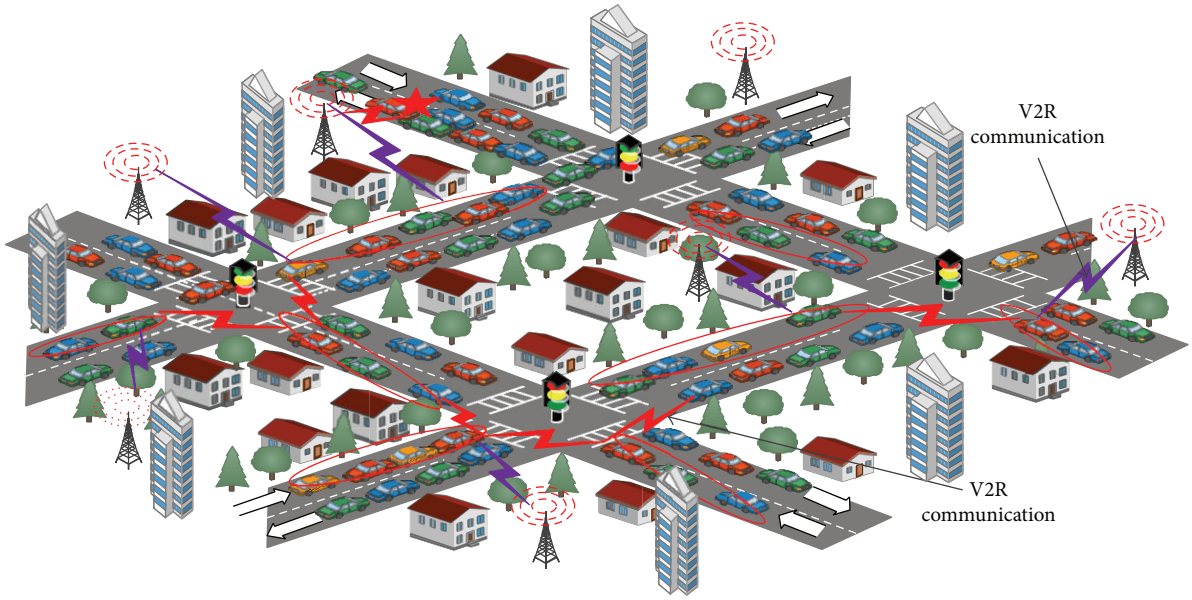


Figure 1: VANETs and its communication modes.

different parameters like relative speed of the moving nodes, transmission range, and number of hops. This result helps to enhance the performance of the reactive routing protocols used in MANETs. It also shows that the inverse of the path duration is directly proportional to the throughput of the ad hoc networks for dynamic source routing (DSR) protocol.

To maximize the path duration of the route, link duration of that path must be maximum. If there is any link breaks in the route, it means whole route will be expired. The link stability and route lifetime are directly proportional to each other. In [16], link stability and route lifetime are analyzed for the ad hoc networks. "Edge Effect" phenomenon occurring in the highly dense networks is also discussed. "Edge Effect" phenomenon is an adverse effect on the network performance when the greedy routing approach is used in dense networks. In greedy routing approach, sender node selects the border node or node cover maximum distance towards the destination. In dense network, nodes are easily available at border of the transmission range and are selected as the next-hop node for further transmission. Small movement of the border nodes outside the transmission range breaks the path and degrades the performance of the network. Therefore border node must lie within the transmission range or on the border line of the sender's transmission range to improve the routing as well as overall network performance.

Path selection is important to decide the path duration of the route in VANETs. The shortest path is not always the best path in terms of the path duration. To maximize the path duration, the shorter average link duration of nodes should be avoided over the longer average link duration of nodes. In [17], a scheme is proposed using ad hoc on demand distance vector (AODV) routing protocol to maximize the path duration and also provide a local path recovery in case of path failure with the help of the cached alternative path computing. In this schema, path information is recorded in a table with five fields: (i) destination sequence number, (ii) next-hop to the destination, (iii) hop count, (iv) inverse path duration (IPD), and (v) time stamp. The path for the transmission is chosen first on the basis of destination sequence number and then on the basis of the IPD values. If two paths tie on the basis of above two fields, then the path will be selected on the basis of hop count. The path on first rank is selected as primary path and other paths are cached as backed recovery path.

In [17], result shows that expected link duration of the path is the parameter of the exponential distribution. This exponential distribution can be used to approximate the distribution where hop count is large. In MANETs or VANETs, the greedy routing and least remaining distance (LRD) approach are used. In LRD approach, the next-hop node is the node which attempts to minimize the remaining distance between source and destination in every hop. The average progress per hop towards destination also helps to find the number of hops from source-to-destination. Authors in [18], show that the progress per hop and number of hops are related to the node density and distance of the path in greedy routing approach.

To estimate the path duration of the route in VANETs, the use of the suitable routing protocol is also a critical factor. The position-based routing protocols may be the suitable routing protocols in VANETs. The routing protocols using the position information of the node in the networks are known as the position-based routing protocols. In these protocols, the next-hop node will be selected on the basis of maximum distance covered towards the destination within the sender's transmission range. Some position-based routing protocols such as border-node based most forward progress within radius routing (B-MFR) [19] and edge-node based directional routing (E-DIR) [20] have been proposed for VANET to select the best node for further transmission. These routing

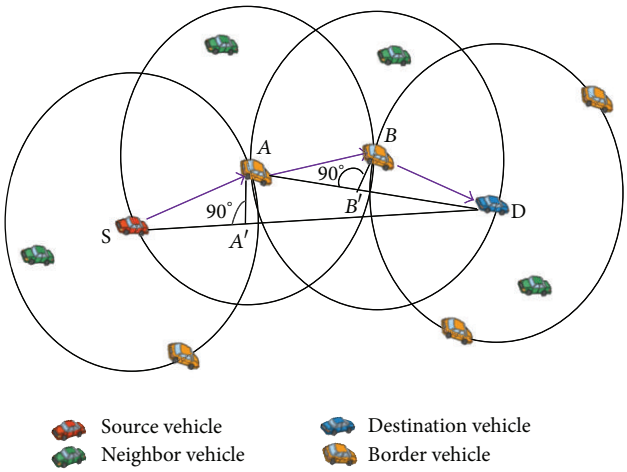


Figure 2: B-MFR forwarding method.

protocols can be used in estimation of path duration for VANET. We are introducing B-MFR for our proposed work.

B-MFR is a modified version of MFR (most forward progress within radius) routing protocol. It avoids using the interior nodes within the transmission range for packet forwarding. Therefore, in B-MFR, a packet is sent to the next-hop node, which is positioned nearer to border or on the border of the transmission range towards the destination (see Figure 2). It means the border node with the greatest progress on the line joining source and destination is chosen as a next-hop node. Thus, in B-MFR, only the border nodes present nearer to border or on the border of the sender's transmission range are selected as the next-hop nodes for further packet transmission.

In contrast, B-MFR exploits the dynamics of VANETs by choosing the node farthest from the source node to make the multihop forwarding more efficient. This routing protocol is especially important in VANETs when the node density is high and gives better performance in the dense networks like city traffic scenario. In the dense network, nodes on the border of the transmission range are easily available to forward the data. Therefore, the border-node based protocol helps to reduce the number of hops between source and destination. Moreover, B-MFR is a very useful routing concept to estimate expected distance and expected number of hops between source and destination node mathematically. These mathematical expressions can be used in estimation of path duration in VANETs.

The main drawback of the B-MFR is that the node farthest from the source cannot be sufficient selection criteria in a network with frequent topological changes. Further, B-MFR is not more suitable for sparse VANETs where smaller number of vehicles moving on the road and no or fewer road side units are present along the road. In such networks, it is important to guarantee high packet delivery rate with minimum delay and, therefore, robustness of the routing protocol is of main concern. To achieve this, we need to find the path duration by taking into consideration both the direction and speed of the vehicular nodes.

4. Proposed Work

The randomness in vehicular traffic environment motivates the need for a suitable stochastic model to determine the ability of a routing protocol in terms of successful packet delivery and to analyze the network performance. Poisson distribution has been used in the model to realize dynamic traffic environment in VANETs. We followed Poisson distribution since we are interested in finding number of nodes present in specified forwarding area given the mean density of nodes in the network area. Further, the arrival of each node is independent. We assume that the nodes in the network area are deployed in a two-dimensional space according to a spatial Poisson process. The mathematical model attempts to estimate the average path duration using a position-based routing concept. Once path duration of a path is estimated well before the path breakage then the performance of a routing protocol can be enhanced significantly.

4.1. The Routing Protocol Used. The role of routing protocol is very important in the estimation of the path duration in VANETs where the network topology frequently changes. For the proposed model, the B-MFR position-based routing protocol is used which we have explained in Section 3. B-MFR is based on greedy routing approach and typically used for long distance multihop communications. It minimizes the number of hops that the message has to travel in between source and destination. In greedy routing approach, sender node first finds the position information of its direct neighbors and then selects a node that is closest to the destination as a next-hop node for further transmission. The B-MFR method is useful to estimate the path duration as the number of hops can be decreased significantly by selecting the border node as a next-hop node.

4.2. Mathematical Models. Since we are interested in the path duration, therefore the main goal of this section is to derive a mathematical expression for path duration between two vehicles by deriving other useful mathematical expressions

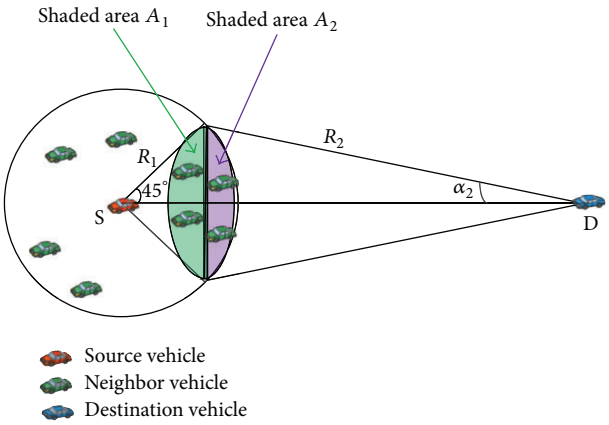


FIGURE 3: Area of shaded region ($A_s = A_1 + A_2$).

such as average number of hops and link duration. In this work, we use traditional traffic flow principle to describe the vehicular environment which will be more accurate for our path duration estimation. Vehicles are assumed to follow Poisson distributed arrivals for obtaining the probability distribution function (*pdf*).

4.2.1. Modeling Assumptions. In our proposed model, some assumptions are made which are given as follows.

(1) Nodes are equipped with GPS receiver, digital map, and sensors.

(2) No other fixed infrastructure for communication is present.

(3) Transmission range for every node is same.

(4) Average speed of the nodes in the network is constant.

(5) Link duration for the nodes moving away from the source node.

4.2.2. Mathematical Notations. The notations are used in our analysis. See Notations section.

4.2.3. Area of Border Region for Finding the Next-Hop Node. Since the node closer to the border or on the border line covers maximum distance, it may reduce the number of hop counts between source and destination. It may not be possible to find even a single node at the extreme end of the sender's transmission range. Therefore, we have considered a region around the extreme end of the transmission range towards the destination. That region is shown by shaded area in Figure 3.

In Figure 3, circle shows the transmission range of source node *S*. Shaded region of the circle ($A_1 + A_2$) can be used as the border region. Further, the shaded region can be divided into two parts showing areas A_1 and A_2. Border nodes can be chosen from the shaded area A_2, because area A_2 lies closest to border line of the sender's transmission range. Greedy routing is the appropriate method to select a next-hop node for the given shaded region.

Area of the shaded region can also be known as the area of interaction of the two circles, one with the radius R_1 and the other with the radius R_2. Area A_s of the shaded region can be calculated as

$$A_s = A_1 + A_2, \tag{1}$$

where

$$\begin{aligned} A_1 &= R_1^2 \cdot \alpha_1 - \frac{R_1^2 \cdot \sin(2\alpha_1)}{2}, \\ A_2 &= R_2^2 \cdot \alpha_2 - \frac{R_2^2 \cdot \sin(2\alpha_2)}{2}. \end{aligned} \tag{2}$$

As we can see in Figure 3, the line from source to destination (SD) is the bisector of the angle 90° ($\alpha_1 = 45°$); therefore, area of shaded region, A_s, is

$$A_s = R_1^2 \left[\frac{\pi - 2}{4}\right] + R_2^2 \left[\alpha_2 - \frac{\sin(2\alpha_2)}{2}\right]. \tag{3}$$

Thus, we can say that the shaded region is the combination of two arcs, one with radius R_1 and the other with radius R_2. The value of α_2 depends on transmission range R_1 and distance between source and destination.

4.2.4. Probability of Finding Nodes in the Shaded Region. In this section, our aim is to find the probability of at least one node in the border region to improve the performance of the network. We assume that nodes are two-dimensionally Poisson distributed over the network with node density ω. The number of nodes present in selected region can be calculated as

$$\text{Number of nodes} = \omega \times \text{Area of the region}. \tag{4}$$

If X is the random variable representing the number of nodes in the shaded region A_s, then the probability that a number of nodes x are located in the region A_s can be calculated as

$$P(X = x) = \frac{(\omega A)^x \cdot e^{-\omega A}}{x!}, \quad x = 0, 1, 2, 3, \ldots. \tag{5}$$

The probability of selecting n nodes out of x nodes is given by

$$P(Y = n) = \binom{x}{n} (p)^n (1 - p)^{x-n}. \tag{6}$$

If a node is present in the selected region, only two possibilities are there: one is selecting the node, p_s, and the other is not selecting the node, q ($= 1 - p_s$).Therefore, probability of both the cases occurring is equal; that is, $p_s = q = 1/2$. Now, probability of selecting exactly n nodes in the given shaded region is [21]

$$\begin{aligned} P(n) &= \sum_{x=n}^{\infty} \binom{x}{n} (p_s)^n (1 - p_s)^{x-n} \cdot \frac{(\omega A_s)^x}{x!} \cdot e^{-\omega A_s} \\ &= \frac{(p_s \omega A_s)^n}{n!} \cdot e^{-p_s \omega A_s}. \end{aligned} \tag{7}$$

Put the value of $p_s = 1/2$ and A_s in the above equation; then

$$P(n) = \left(\frac{\omega}{2}\left\{R_1^2\left[\frac{\pi-2}{4}\right] + R_2^2\left[\alpha_2 - \frac{\sin(2\alpha_2)}{2}\right]\right\}\right)^n \times (n!)^{-1} \cdot e^{-(\omega/2)\{R_1^2[(\pi-2)/4]+R_2^2[\alpha_2-(\sin(2\alpha_2)/2)]\}}. \quad (8)$$

Similarly, the probability for selecting at least n nodes in the shaded region is

$$P_n = 1 - \sum_{i=0}^{n-1}\left(\frac{\omega}{2}\left\{R_1^2\left[\frac{\pi-2}{4}\right] + R_2^2\left[\alpha_2 - \frac{\sin(2\alpha_2)}{2}\right]\right\}\right)^i \times (i!)^{-1} \cdot e^{-(\omega/2)\{R_1^2[(\pi-2)/4]+R_2^2[\alpha_2-(\sin(2\alpha_2)/2)]\}}. \quad (9)$$

From (9), we can easily obtain the probability, P, of having at least one node within the border region as

$$P = 1 - P(X = 0) = 1 - e^{-(\omega/2)\{R_1^2[(\pi-2)/4]+R_2^2[\alpha_2-(\sin(2\alpha_2)/2)]\}}. \quad (10)$$

4.2.5. Average Number of Hops between Source and Destination Node. Number of hops can be defined as the number of intermediate nodes in the route (source to destination). The main assumption is that each hop results in the same progress towards the destination, equal to the average distance covered by a node in one hop. Number of hops should be as low as possible. It will decrease the chances of link breakage and improve the path duration between nodes [10].

To determine the average number of hop counts, nodes within the transmission range R follow the Poisson distributed model. If destination node is present in the sender's transmission range then the probability of finding destination node is the same as the probability of finding next-hop node. We assume that Z_1 is the distance between the source and next-hop node. The probability density function (*pdf*) of the link distance, Z_1, between source and next-hop node is defined as [22]

$$f(Z_1) = 2\pi\omega Z_1 \cdot e^{-\pi\omega Z_1^2}. \quad (11)$$

Distance between two nodes which provide a link to a route can be defined as the link distance. Link distance can be increased by increasing the distance between source node and next-hop node towards border line within the transmission range [23]. The probability of one-hop count can be calculated as

$$P(1) = \int_0^{R_1} f(Z_1)\, dZ_1 = 1 - e^{-\pi\omega R_1^2}. \quad (12)$$

However, the destination node can be far away from the source node, which may be two, three, or more hop counts. If the destination node is out of transmission range, R_1, and less than $2R_1$, then at least one intermediate node is required between source and destination to transmit the packet further in the network. The probability of two-hop counts can be calculated as follows:

$$P(2) = \int_{R_1}^{2R_1} 2\pi\omega Z_1 \cdot e^{-\pi\omega Z_1^2} dZ_1 \times \left[1 - e^{-(\omega/2)\cdot A_s}\right], \quad P(2) = \left[e^{-\pi\omega R_1^2} - e^{-4\pi\omega R_1^2}\right] \times \left[1 - e^{-(\omega/2)\cdot A_s}\right]. \quad (13)$$

Similarly, the probability for three-hop count is

$$P(3) = \left[e^{-4\pi\omega R_1^2} - e^{-9\pi\omega R_1^2}\right] \times \left[1 - e^{-(\omega/2)\cdot A_s}\right]^2. \quad (14)$$

Consequently, the k-hop counts probability can be defined as

$$P(k) = \left[e^{-(k-1)^2\pi\omega R_1^2} - e^{-k^2\pi\omega R_1^2}\right] \times \left[1 - e^{-(\omega/2)\cdot A_s}\right]^{k-1}. \quad (15)$$

Now, by using (12), (13), (14), and (15), we can calculate the expected number of hops, E_H, between source and destination as follows:

$$E_H = \sum_{H=1}^{k} HP(H) = P(1) + 2P(2) + 3P(3) + \cdots + kP(k), \quad E_H = \sum_{H=1}^{k} H\left[e^{-(H-1)^2\pi\omega T^2} - e^{-H^2\pi\omega T^2}\right] \times \left[1 - e^{-(\omega/2)\{R_1^2[(\pi-2)/4]+R_2^2[\alpha_2-(\sin(2\alpha_2)/2)]\}}\right]^{H-1}. \quad (16)$$

4.2.6. Velocity of Nodes. Direction of movement and speed of a node are very essential parameters for the calculation of the path duration in case of VANETs. Link duration depends on the relative velocity of the nodes as it can increase the link distance between nodes. The relative velocity between nodes is inversely proportional to the link duration. The relative velocity of the source node and next-hop node should be known to determine the expected link duration. Let V_1 and V_2 be the velocity of source and next-hop nodes; then relative velocity V_R of the nodes can be calculated as

$$V_R = \sqrt{V_1^2 + V_2^2 - 2 \cdot V_1 \cdot V_2 \cos\theta}. \quad (17)$$

In this work, we assume that all the nodes move with constant velocity in the network; that is

$$V = V_1 = V_2. \quad (18)$$

Therefore, relative velocity is

$$V_R = V \cdot \sqrt{2(1 - \cos\theta)}. \quad (19)$$

In the above equation, θ can vary from 0 to $\pi/2$ as the next-hop node can move in the direction of destination only to maintain the communication link (link duration) between nodes. We assume that angle θ is uniformly distributed within

$(0, \pi/2)$, and *pdf* of $f_\theta(\theta)$ is $2/\pi$. Then the *pdf* of V_R, $f_{V_R}(V_R)$ can be expressed as

$$f_{V_R}(V_R) = \frac{1}{\sqrt{1-\sin^2(\theta/2)}} \cdot \frac{2}{\pi} = \sqrt{\frac{4V^2 - V_r}{V}} \cdot \frac{2}{\pi}. \tag{20}$$

4.2.7. Link Duration. Link duration is the time for which the direct link between two nodes within the transmission range is active and it is a part of the route. It is necessary that next-hop node must be present within the transmission range of the source node to maintain the communication link between source and next-hop node. In this work, as we assumed, border node will be the next-hop node for each hop between the source and destination. Since the velocity of each node in the network is constant, it means that the links between source and next-hop node will always be maintained. As we have assumed Z_1 to be the distance between source and next-hop node within radius R_1, then the expected value of Z_1 [24] can be computed as

$$E_{Z_1} = \frac{nR_1}{(n+1)}. \tag{21}$$

Therefore, link duration T can be expressed as

$$T = \frac{E_{Z_1}}{V_R} = \frac{nR_1}{V_R(n+1)}. \tag{22}$$

The *pdf* of T, $f_T(T)$ is given by

$$\begin{aligned} f_T(T) &= \int_0^V V_R \cdot f_{d_{V_R}}(V_R T, V)\, dv \\ &= \int_0^V \left[E_{Z_1}\right] \cdot \left[\frac{2}{\sqrt{4V^2 - V_R^2}} \cdot \frac{1}{\pi}\right] dV_r. \end{aligned} \tag{23}$$

4.2.8. Path Duration. The path duration is one of the key parameters which could be useful to improve the performance and throughput of a highly dynamic network such as VANET. The path duration will be helpful in the process of path selection during the transmission of packet from source to destination [25, 26]. Path duration can be derived from the *pdf* of the link duration. Let $T_1, T_2, T_3, \ldots, T_{E_H}$ denote the link duration of $1, 2, 3, \ldots, E_H$ hops, respectively. E_H is the average number of hops required to reach the destination as estimated in (16). Therefore, the path duration can be expressed as

$$T_{\text{path}} = \text{MIN}\left(T_1, T_2, T_3, \ldots, T_{E_H}\right). \tag{24}$$

By using Bayes' theorem [21], the *pdf* of T_{path} is

$$f\left(T_{\text{path}}\right) = E_H \cdot E_Z \cdot C_T^{E_H - 1}. \tag{25}$$

Here, T represents the link duration and $C_T = 1 - F_T$ is the complementary cumulative density function (cdf) of T. Therefore,

$$f\left(T_{\text{path}}\right) = E_H \cdot f_T(T) \cdot \left[1 - \int_{T=0}^{\infty} f_T(T)\, dT\right]^{E_H - 1}. \tag{26}$$

Finally, the average path duration can be estimated as

$$\begin{aligned} E_{T_{\text{path}}} &= \int_0^\theta T_{\text{path}} \cdot f\left(T_{\text{path}}\right) \cdot dT_{\text{path}} \\ E_{T_{\text{path}}} &= \int_0^\theta T_{\text{path}} \cdot E_H \cdot f_T(T) \\ &\quad \cdot \left[1 - \int_{T=0}^{\infty} f_T(T)\, dT\right]^{E_H - 1} \cdot dT_{\text{path}}. \end{aligned} \tag{27}$$

5. Simulations and Results Analysis

In this section, extensive simulations have been performed to analyze the mathematical model for estimation of path duration presented in Section 4. The impact of four parameters, namely, transmission range, number of hops, velocity of nodes, and density of nodes on path duration, have been analyzed. The simulation results of the model have been compared with analytical results obtained for the mathematical formulation of path duration estimation.

5.1. Simulation Environment. The mathematical estimation of path duration has been simulated using network simulator (NS-2.34). MOVE (mobility model generator for vehicular networks) [27] has been used to generate realistic vehicular traffic environment along with open-source microtraffic simulator, SUMO (simulation of urban mobility). The vehicular traffic scenario consists of roads, traffic lanes on roads, junctions, traffic lights at junctions, vehicles speed, probability of turning left or right of a vehicle at junctions, and so forth has been set up using road map editor and movement editor of MOVE. The trace file used in ns2 is produced following the setup procedure of MOVE.

A set of five horizontal and five vertical roads crossing each other and thus making twenty five junctions is used as simulation area. The lane width used is 5 m. The velocity range 0–60 Km/h is used for node movement and transmission range varies from 100 m to 600 m. The other basic parameters used for the simulation are packet size of 512 bytes, traffic type as CBR, wireless channel, omnidirectional antenna, 802.11p as MAC wireless standard, and 300 s simulation time. The position-based routing protocol used for the simulation is B-MFR. After setting the network and traffic flow with above discussed parameters, we conducted the simulation. The average of ten different simulation runs is taken for data record where different source and different destination are selected. MATLAB is used to obtain analytical results for the mathematical formulation of the model.

5.2. Result Analysis. The results obtained for the model have been analyzed in the following subsections. In each subsection, impact of a specific parameter on path duration has been analyzed. In each analysis, the simulation and analytical results have been discussed comparatively.

5.2.1. The Impact of Transmission Range. Figure 4 shows the impact of transmission range on path duration. Path duration

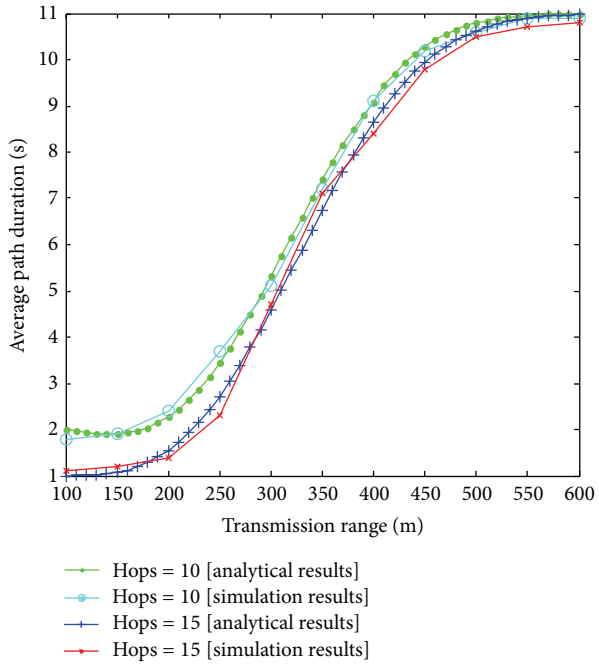


FIGURE 4: Average path duration versus transmission range.

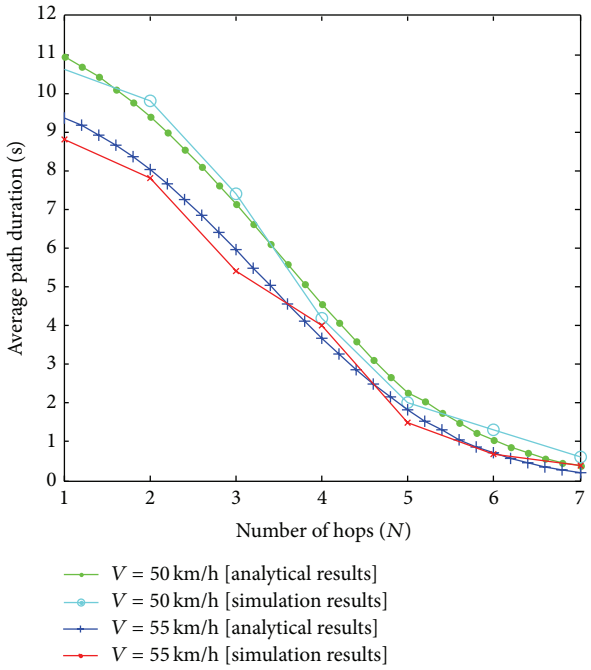


FIGURE 5: Average path duration versus number of hops.

of the routes is dependent on the transmission range of the nodes in the network. It can be clearly observed that the average path duration increases with the increase in transmission range. This can be attributed to the fact that the larger transmission range increases the probability of finding next-hop node in the border area of sender's transmission range. It means selection of border node gives better result as compared to interior node within the transmission range. Additionally, increasing the number of hops in the considered path decreases the average path duration. The simulation results are close to the analytical results that validate the model.

5.2.2. The Impact of Number of Hops. The plots in Figure 5 show the impact of number of hops on average path duration which depends on each hop of the route and it varies with number of hops for fixed transmission range. It clearly reveals that increasing the number of hops in the path decreases average path duration. The reason behind this is that increasing the number of hops in the path also increases the probability of link failure. Figure 5 also shows that the high velocity of nodes (e.g., $V = 55$ m/s) has a relatively reduced average path duration as compared to low velocity of nodes (e.g., $V = 50$ m/s). This is due to the increasing probability of link failure with higher velocity of nodes. The simulation results are close to the analytical results that verify the model.

5.2.3. The Impact of Node Density. The node density is also a critical factor for the path duration as the possibility of finding suitable next-hop node is increased with the increased number of nodes. The results depicted in Figure 6 show the impact of node density on average path duration. It clearly conveys that increasing the node density in a given area increases path duration. This is due to the increasing probability of availability of suitable next-hop node with increasing node density in the network. Moreover, the results also confirm that, for a given node density, increasing the number of hops in the path decreases average path duration. The closeness between analytical and simulation results for hops = 15 is smaller as compared to hops = 10 due to the increment in the network dynamics with increasing number of hops.

5.2.4. The Impact of Velocity of Nodes. VANETs are known for their mobility. The high velocity of nodes makes vehicles in and out from the transmission range of the source node, which causes most of the link or path breakage in the network. The results in Figure 7 show the impact of velocity of nodes on average path duration. It can be observed that increasing the velocity of nodes decreases the path duration. This can be attributed to the fact that increasing the velocity of nodes in the network increases the probability of link failure which ultimately decreases average path duration. Moreover, it also reveals that increasing the number of hops for a given velocity decreases average path duration due to the increment in the number of links in the path that again increases the probability of link failure. In the figure, simulation results are very close to the analytical results.

6. Conclusion

In this paper, we have derived a mathematical model for estimation of path duration between source and destination nodes using position-based routing concept. This model has

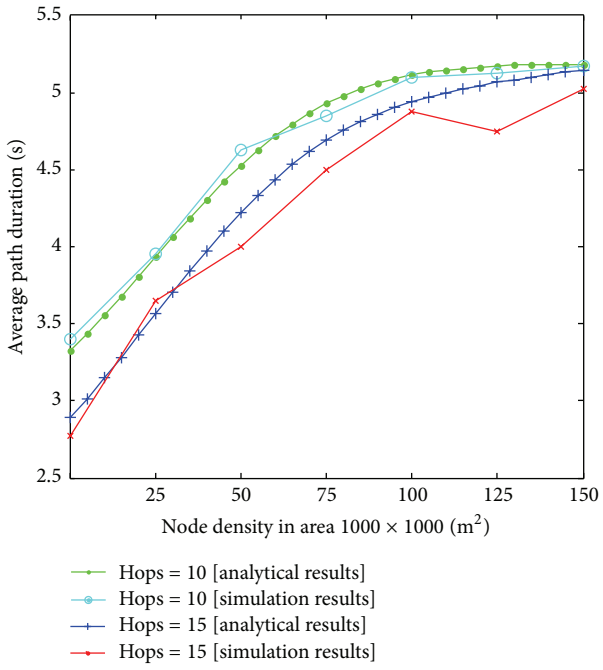


Figure 6: Average path duration versus node density.

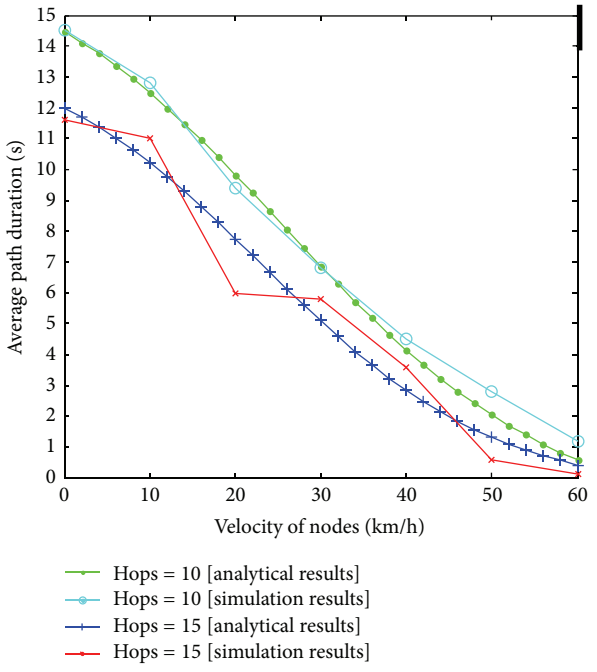


Figure 7: Average path durations versus velocity of nodes.

been verified by means of simulations and analytical results. Both the results are well approximated by the mathematical model. The mathematical model that we present is able to describe the effects of various road traffic parameters including the transmission range, number of hops, node density, and velocity of nodes on path duration. The use of the B-MFR routing protocol in this work is to find the routing path with maximum path duration. It demonstrates the selection of next-hop nodes positioned in a region around the extreme end of the transmission range towards the destination.

The mathematical analysis and simulation results reveal that high transmission range and smaller number of hops increase the path duration, but it decreases as the velocity of the nodes increases. From the above observations, it may appear that, in highly dynamic networks such as in VANET, it is very necessary to maintain the path duration between source and destination nodes. Therefore, message can be forwarded timely to reduce the large number of accidents on the road. Thus, we can say that the work in this paper helps us to improve the routing performance and decrease the number of path failures generally occurring in VANETs.

Notations and Their Descriptions

R_1: Transmission range of nodes (omnidirectional)
A: Region covered by transmission range
A_s: Selected region of the transmission range
ω: Node density
x: Number of nodes in the shaded region
n: Number of nodes selected out of x nodes
P_s: Probability of successfully selecting a node
q: Probability of not selecting a node
P_n: Probability for selecting at least n nodes in the shaded region
E_H: Expected number of hops between source and destination nodes
R_2: Distance between next-hop node and destination node
Z_1: Distance between source and next-hop node
V_R: Relative velocity between source node and next-hop node
α_1: Angle between R_1 and SD
α_2: Angle between R_2 and SD
θ: Relative angle between source and next-hop node.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] WHO, *Status Report on Road Safety*, World Health Organization (WHO), Geneva, Switzerland, 2013.

[2] T. Taleb, E. Sakhaee, A. Jamalipour, K. Hashimoto, N. Kato, and Y. Nemoto, "A stable routing protocol to support ITS services in VANET networks," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 6, pp. 3337–3347, 2007.

[3] O. Kaiwartya and S. Kumar, "Geocast routing: Recent advances and future challenges in vehicular adhoc networks," in *Proceedings of Signal Processing and Integrated Networks (SPIN '14)*, pp. 291–296, 2014.

[4] O. Kaiwartya and S. Kumar, "Cache agent based geocasting (CAG) in VANETs," International Journal of Information and Communication Technology. *In press*.

[5] O. Kaiwartya, S. Kumar, and R. Kasana, "Traffic light based time stable geocast (T-TSG) routing for urban VANETs," in *Proceedings of the 6th International Conference on Contemporary Computing (IC3 '13)*, pp. 113–117, 2013.

[6] O. Kaiwartya and S. Kumar, "GeoPSO: geocasting through particle swarm optimization in vehicular adhoc networks," in *Proceedings of Information Systems and Design of Communication (ISDOC '14)*, ACM, May 2014.

[7] O. Kaiwartya and S. Kumar, "Enhanced caching for geocast routing in vehicular Ad-Hoc networks (ECGR)," in *Proceedings of the International Conference on Advanced Computing, Networking and Informatics (ICACNI '13)*, vol. 243, pp. 213–220, Springer, 2013.

[8] J. J. Blum, A. Eskandarian, and L. J. Huffman, "Challenges of intervehicle Ad Hoc networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 5, no. 4, pp. 347–351, 2004.

[9] W. Alasmary and W. Zhuang, "Mobility impact in IEEE 802.11p infrastructure less vehicular networks," *Ad Hoc Networks*, vol. 10, no. 2, pp. 222–230, 2012.

[10] R. J. La and Y. Han, "Distribution of path durations in mobile ad hoc networks and path selection," *IEEE/ACM Transactions on Networking*, vol. 15, no. 5, pp. 993–1006, 2007.

[11] "DSRC Standards: What's New?" ITS Standards Advisory number 3, US Department of Transportation, http://www.its.dot.gov/DSRC/.

[12] R. Popescu-Zeletin, I. Radusch, and M. A. Rigani, *Vehicular-2-X Communication: State-Of-The-Art and Research in Mobile Vehicular Ad Hoc Networks*, Springer, New York, NY, USA, 2010.

[13] A. F. Molisch, F. Tufvesson, J. Karedal, and C. F. Mecklenbräuker, "A survey on vehicle-to-vehicle propagation channels," *IEEE Wireless Communications*, vol. 16, no. 6, pp. 12–22, 2009.

[14] J. Gozalvez, M. Sepulcre, and R. Bauza, "IEEE 802.11p vehicle to infrastructure communications in urban environments," *IEEE Communications Magazine*, vol. 50, no. 5, pp. 176–183, 2012.

[15] N. Sadagopan, F. Bai, B. Krishnamachari, and A. Helmy, "PATHS: analysis of PATH duration statistics and their impact on reactive MANET routing protocols," in *Proceedings of the 4th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MOBIHOC '03)*, pp. 245–256, June 2003.

[16] G. Lim, K. Shin, S. Lee, H. Yoon, and J. S. Ma, "Link stability and route lifetime in ad-hoc networks," *EURASIP Journal on Wireless Communications and Networking*, pp. 1–6, 2007.

[17] Y. Han and R. J. La, "Maximizing path durations in mobile ad-hoc networks," in *Proceedings of the 40th Annual Conference on Information Sciences and Systems (CISS '06)*, pp. 26–31, Princeton, NJ, USA, March 2006.

[18] S. De, A. Caruso, T. Chaira, and S. Chessa, "Bounds on hop distance in greedy routing approach in wireless ad hoc networks," *International Journal of Wireless and Mobile Computing*, vol. 1, no. 2, pp. 131–140, 2006.

[19] R. S. Raw and D. K. Lobiyal, "B-MFR routing protocol for vehicular ad hoc networks," in *Proceedings of the International Conference on Networking and Information Technology (ICNIT '10)*, pp. 420–423, Manila, Philippines, June 2010.

[20] R. S. Raw and D. K. Lobiyal, "E-DIR: a directional routing protocol for VANETs in a city traffic environment," *International Journal of Information and Communication Technology*, vol. 3, no. 3, pp. 242–257, 2011.

[21] S. Ukkusuri and L. Du, "Geometric connectivity of vehicular ad hoc networks: analytical characterization," *Transportation Research Part C: Emerging Technologies*, vol. 16, no. 5, pp. 615–634, 2008.

[22] S. M. Harb and J. McNair, "Analytical study of the expected number of hops in wireless Ad Hoc network," in *Wireless Algorithms, Systems, and Applications*, vol. 5258 of *Lecture Notes in Computer Science*, pp. 63–71, 2008.

[23] K. Namuduri and R. Pendse, "Analytical estimation of path duration in mobile ad hoc networks," *IEEE Sensors Journal*, vol. 12, no. 6, pp. 1828–1835, 2012.

[24] R. S. Raw and S. Das, "Performance analysis of P-GEDIR protocol for vehicular ad hoc network in urban traffic environments," *Wireless Personal Communications*, vol. 68, no. 1, pp. 65–78, 2013.

[25] H. Ghafoor, K. Insoo, and G. Nasirud-Din, "Neighboring and connectivity-aware routing in VANETs," *The Scientific World Journal*, vol. 2014, Article ID 789247, 10 pages, 2014.

[26] C. Chen, Y. Jin, Q. Pei, and N. Zhang, "A connectivity-aware intersection-based routing in VANETs," *EURASIP Journal on Wireless Communications and Networking*, vol. 2014, article 42, 2014.

[27] F. K. Karnadi, Z. H. Mo, and K. Lan, "Rapid generation of realistic mobility models for VANET," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC '07)*, pp. 2508–2513, IEEE, March 2007.

Polym. Bull.
DOI 10.1007/s00289-014-1217-z

ORIGINAL PAPER

Rubber toughening of unsaturated polyester with core–shell poly(siloxane)-epoxy microspheres

Prasun Kumar Roy · Nahid Iqbal · Devendra Kumar · Chitra Rajagopal

Received: 27 January 2014 / Revised: 18 May 2014 / Accepted: 8 July 2014

Abstract In this paper, we have explored the potential of core–shell poly(siloxane)-epoxy microspheres towards improving the dynamic properties of a conventional unsaturated polyester (USP). Micro-sized poly(siloxane) beads were prepared by suspension polymerisation route, where the particle dimensions could be tailored by varying the operating parameters, particularly stirring speed and feed concentration. The core was subsequently coated with epoxy to form an external layer compatible with the USP resin, with an aim to aid its homogeneous dispersion in the thermosetting resin. Toughened USP composites containing varying amounts of both coated and uncoated microspheres (3–10 % w/w) were prepared by curing under ambient conditions, and their mechanical properties were evaluated under both quasi-static and dynamic loading conditions. The introduction of epoxy-coated poly(siloxane) led to a proportional decrease in the tensile strength and modulus, which were found to compare well with the predictions based on Halpin–Tsai and Lewis–Nielson empirical models. Significant improvements in the impact strength of USP could be achieved, and under optimised conditions, 101 % increase in the impact strength was observed, which corroborated with significant increase in mean critical stress intensity factor (76 %). Morphological investigations of the fractured surface revealed the presence of characteristic features which were used to establish the underlying routes responsible for the toughened nature of the USP composites.

Electronic supplementary material The online version of this article (doi:10.1007/s00289-014-1217-z) contains supplementary material, which is available to authorized users.

P. K. Roy (✉) · C. Rajagopal
Centre for Fire, Explosive and Environment Safety, Timarpur, Delhi 110054, India
e-mail: pk_roy2000@yahoo.com; pkroy@cfees.drdo.in

N. Iqbal · D. Kumar
Department of Applied Chemistry and Polymer Technology, Delhi Technological University, Delhi 110042, India

Published online: 17 July 2014

Keywords USP · Core–shell · Toughening

Introduction

Unsaturated thermosetting polyester (UPR) (USP) is one of the most widely used fibre-reinforced composite matrix material; its popularity arising from its excellent mechanical properties, chemical resistance, low cost, ease of processing and excellent fibre wettability [1]. Conventional UPRs (USPs) are primarily solutions of unsaturated polyesters with a diluent, usually styrene, and the curing process results in the linking of polyester chains through poly(styrene) linkages at the unsaturated sites, thereby forming an amorphous three-dimensional molecular structure. However, this crosslinked network is inherently brittle, which restricts its usage in highly demanding applications.

Introduction of suitable materials into USP has long been practiced to increase the impact strength of the base resin. In this context, the most common approach includes blending with liquid reactive rubbers and diluents [2–6]. Curing of the resin in presence of liquid rubber leads to its precipitation, thereby forming phase-separated blends with higher impact strength. Unfortunately, blends with liquid rubbers suffer from specific disadvantages including lowering of the elastic modulus and the plasticising action of the rubber can result in significant reduction of the glass transition temperature. These issues can be overcome by using thermoplastics as toughening agents, but these modifiers are relatively difficult to process [7–10].

Lately, the potential of nanofillers has also been explored [11], but the extent of toughening is much lesser [12]. It is to be noted that the properties of all the above-mentioned blends are governed by the thermodynamically and kinetically controlled process of phase separation, which becomes increasingly difficult to control in fast curing compositions. Hence, developing techniques of controlling the blend morphology has attracted the attention of researchers worldwide. An alternative route involves inclusion of small quantities of preformed rubbers within the matrix [2, 13, 14], which permits control over the resultant morphology, without decreasing the glass transition temperature of the base polymer [15, 16]. However, due to the incompatibility between the preformed rubbers and matrix, proper dispersion is an issue, which can be resolved by coating these with an additional layer leading to the formation of core–shell rubbers (CSR). Surprisingly, most of the relevant literature in this field is confined to particles, where the elastomeric core is obtained from organic rubbers, which are not heat resistant, thereby limiting its use to less demanding applications [16–18]. In this context, the use of a thermally stable siloxane elastomeric core offers a potential solution [16, 19, 20].

The dimensions of elastomeric filler are one of the most important parameters which affect its toughening ability. The effect of particle size on the fracture properties of epoxy has been extensively investigated, which indicates that the particles should be large enough to allow their deformation energy to be higher than their interfacial bonding to the matrix to permit cavitation [15, 21, 22]. It has been reported that rubbers possessing diameter $\sim$100–200 μm are the most efficient

impact modifiers, which have reportedly led to over a tenfold improvement in the mechanical properties [2]. At present, elastomeric microspheres are synthesised by emulsion polymerisation technique, but very few papers describe the synthesis of such materials in the open literature [18, 23, 24].

Interestingly, elastomeric particles of these dimensions can easily be obtained by suspension polymerisation route, and the particle dimensions can be easily controlled by varying parameters like stirring speed and concentration of the feed solution. The aim of this work was to toughen unsaturated polyester resins by blending with preformed siloxane elastomeric microspheres. Our approach was to synthesise the core PDMS microspheres by a simple suspension polymerisation technique followed by coating with a layer of epoxy, which would lead to its apt dispersion within the thermoset. The mechanical properties and fracture energy of these CSR-toughened USP were determined and the results were compared with well-established analytical models. Post-mortem analysis of the fracture surface morphology of the toughened USP composites was performed to identify the underlying toughening mechanisms.

Experimental

Materials and methods

Silicone resin (Elastosil M4644) and the platinum-based hardener were obtained from Wacker, Germany. Isophthalic acid-based unsaturated polyester resin (USP RPL211) with a styrene content of 32 % was kindly provided by Revex (Delhi, India). Methyl ethyl ketone peroxide and cobalt octoate (6 %) were used as the catalyst and accelerator, respectively. PVA and chloroform (CDH) were used without further purification.

Preparation of elastomeric microspheres

The preparation of the core–shell PDMS-epoxy microspheres has been described in our previous paper [25]. In brief, the core PDMS microsphere was prepared by suspension polymerisation process, for which a chloroform-diluted feed solution of vinyl-terminated siloxane macromonomer (30–60 % w/v) was reacted with requisite amounts of platinum-based hardener. The curing reaction was performed in reaction vessel under inert atmosphere where the feed was introduced through a hypodermic syringe into an aqueous PVA solution (1.5 % w/v), maintained at 45 °C under continuous stirring. The polymerisation reaction was allowed to continue for 8 h under varying stirring speeds (500–700 rpm), after which the reaction mixture was cooled and filtered off. The extent of conversion was measured gravimetrically as the ratio of mass of microspheres obtained to the amount of macromonomer used for its preparation.

The obtained microspheres were subsequently coated with epoxy in a separate step. For this purpose, a mixture of PDMS microspheres (15 g) and epoxy resin (7 g) was ultrasonicated for 15 min prior to addition of stoichiometric amounts of

hardener. The mixture was diluted with small amounts of chloroform (5 mL), and the slurry was slowly introduced into a reactor containing aqueous PVA solution being stirred under the conditions mentioned previously.

Preparation of composites

Rubber-toughened polyester composites containing varying amounts of coated and uncoated microspheres (3–10 % w/w) were prepared by room temperature curing process. The microspheres were first sieved through a 60–80 mesh to obtain a particle size range 250–177 μm which was added to the unsaturated polyester. The suspension was degassed under vacuum and transferred to greased silicone moulds, where the curing reaction was allowed to proceed for 24 h at 30 °C in the presence of hardener (10 % w/w). Neat specimens were also prepared under similar conditions for comparison purposes, which will be referred to as UP in the subsequent text. Composites have been designated as UP followed by its concentration and by the type of microsphere used for its preparation, i.e. 'R' denotes uncoated PDMS microsphere and 'ECR' denotes epoxy-coated core–shell microspheres with PDMS core. For e.g. UP5R refers to cured unsaturated polyester containing 5 % PDMS, and UP5ECR refers to the composition containing 5 % epoxy-coated PDMS.

Characterisation

The effect of epoxy coating on the microsphere dimensions was determined by a particle size analyser (DIPA 2000, Donner). FTIR spectra of samples were recorded in the wavelength range 4,000–600 cm^{-1} using a Thermo Fisher FTIR (NICOLET 8700) analyser with an attenuated total reflectance (ATR) crystal accessory. The changes in the thermal properties of the samples were investigated using DSC (TA instruments, Q 20 module) under nitrogen atmosphere. Approximately 4–6 mg of the sample was placed in a 40-μL aluminium cap without pin and sealed with a lid. After erasing the thermal history of samples, they were subjected to a heating programme from −60 to 250 °C at 10 °C min^{-1}. Thermal degradation behaviour was investigated using Perkin Elmer Diamond STG-DTA-DSC under N_2 atmosphere in the temperature range 50–800 °C. A heating rate of 10 °C/min and a sample mass of 5.0 ± 0.5 mg were used for each experiment. The quasi-static mechanical properties were determined as per ASTM method D638 using a Universal Testing System (International equipments) at ambient temperature. The dog-bone-shaped specimens used for tensile testing were 165 mm long, 3 mm thick, and 13 mm wide along the centre of the casting for epoxy resin. The samples were subjected to a cross-head speed of 50 mm min^{-1}. The notched Charpy impact strength of the specimens (notch radius = 45°) was determined as per ASTM D 256 using an impact strength testing machine (International Equipments, India). Notched flexural testing of the samples was performed under three-point, single-edge, notch-bending mode using the same instrument. For this purpose, specimens with requisite dimensions (127 mm length × 12.5 mm width × 3.5 mm thickness and 3 mm notch) were prepared and the samples were subjected to a deformation

rate of 2 mm min^{-1} while maintaining a span length of 60 mm. The mode I critical stress intensity factor (K_{IC}) of the samples was determined under plain strain conditions as per the following equation [26]:

$$K_{Ic} = \frac{3 \times P \times L \times a^{1/2}}{2 \times B \times w^2} Y\left(\frac{a}{w}\right), \tag{1}$$

where, P, L and B refer to the load at break, span length and sample thickness, respectively. The geometry factor, $Y\left(\frac{a}{w}\right)$, is calculated as per the formula below, where a is the notch length and w is the sample width.

$$Y\left(\frac{a}{w}\right) = 1.93 - 3.07 \times \left(\frac{a}{w}\right) + 14.53 \times \left(\frac{a}{w}\right)^2 - 25.11 \times \left(\frac{a}{w}\right)^3 + 25.8 \times \left(\frac{a}{w}\right)^4. \tag{2}$$

The K_{IC} was used to estimate the fracture energy (G_{Ic}), which was calculated using the following equation

$$G_{IC} = \frac{K_{IC}^2(1 - v^2)}{E}, \tag{3}$$

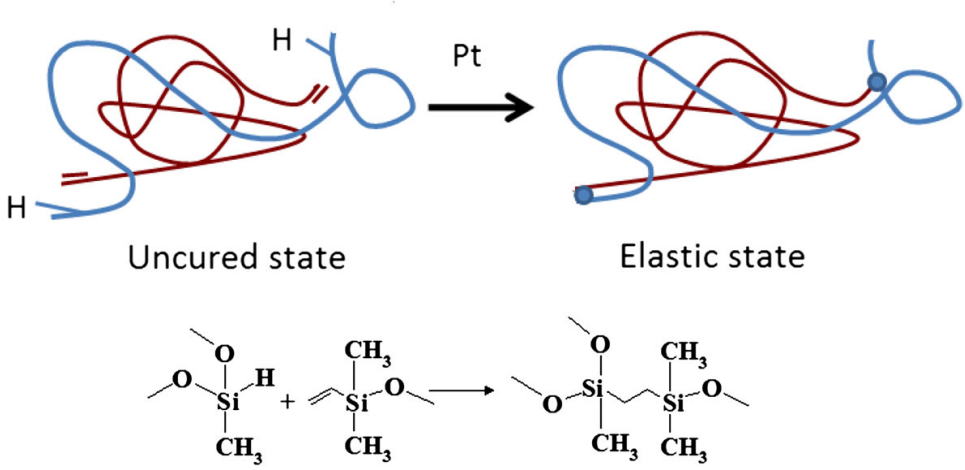
where E is the flexural modulus of the polymer, and v is the Poisson's ratio of polyester resin (0.4) [27]. For each composition, at least five identical specimens were tested and the average results along with the standard deviation have been reported.

The surface morphology of samples was studied using a scanning electron microscope (Zeiss EVO MA15) under an acceleration voltage of 20 kV. Samples were mounted on aluminium stubs and sputter coated with gold and palladium (10 nm) using a sputter coater (Quorum-SC7620) operating at 10–12 mA for 120 s. The core–shell structure of the ECR was also confirmed using energy dispersion analyser (EDS). For this purpose, a representative ECR was cut with a sharp razor blade and carefully mounted on a stub, followed by elemental composition determination in both core as well as in the shell region.

Results and discussion

Preparation of core poly(dimethylsiloxane)microspheres

The suspension curing of vinyl-terminated methyl hydrosiloxane dimethylsiloxane at 45 °C in the presence of a hydrosilylation catalyst led to the formation of smooth PDMS beads, as per the reaction scheme presented in Scheme 1 [25]. The effect of increasing feed concentration on the particle size distribution is presented in Fig. 1, which revealed that the distribution shifts towards larger-sized microspheres with increasing polymerisable content in the dispersed oily droplets.



Scheme 1 Schematic illustrating the hydrosilylation of vinyl-terminated silicone macromonomer using platinum catalyst

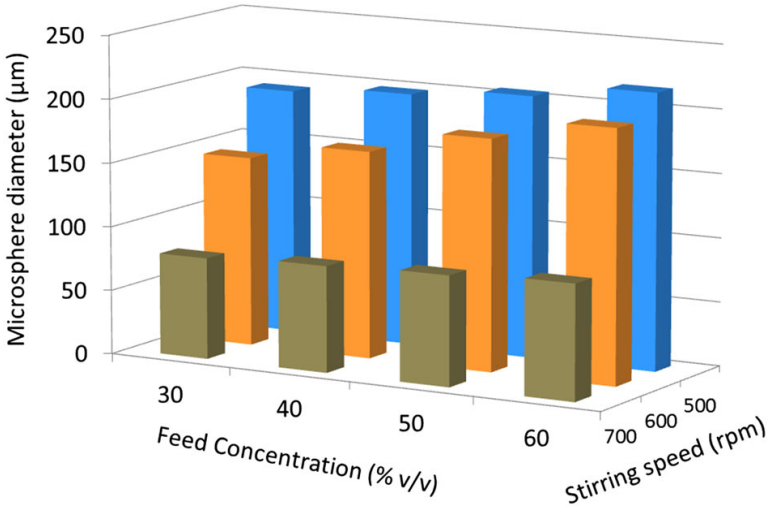


Fig. 1 Effect of feed concentration and stirring speed on PDMS microsphere diameter

Core–shell PDMS-epoxy microspheres

The siloxane microspheres were coated with a layer of cycloaliphatic epoxy in a separate step to improve its compatibility with the thermosetting polyester resin. The detailed characterisation of core–shell PDMS-epoxy microspheres is reported in our previous paper [25]. Coating of the siloxane core with epoxy led to an increase in the particle dimensions. Particle size analysis of the microspheres after coating was used to determine the coating thickness, which was found to be ~20 µm for an internal core of 198 µm (Figure S1). Interestingly, in comparison to neat PDMS, the surface of the ECR was found to be rather rough. The elemental composition of the coated and uncoated PDMS microspheres was determined by EDX, and the average value over a specific square area within a microsphere is presented in Fig. 2. As can be seen, the amount of silicon present in the shell is

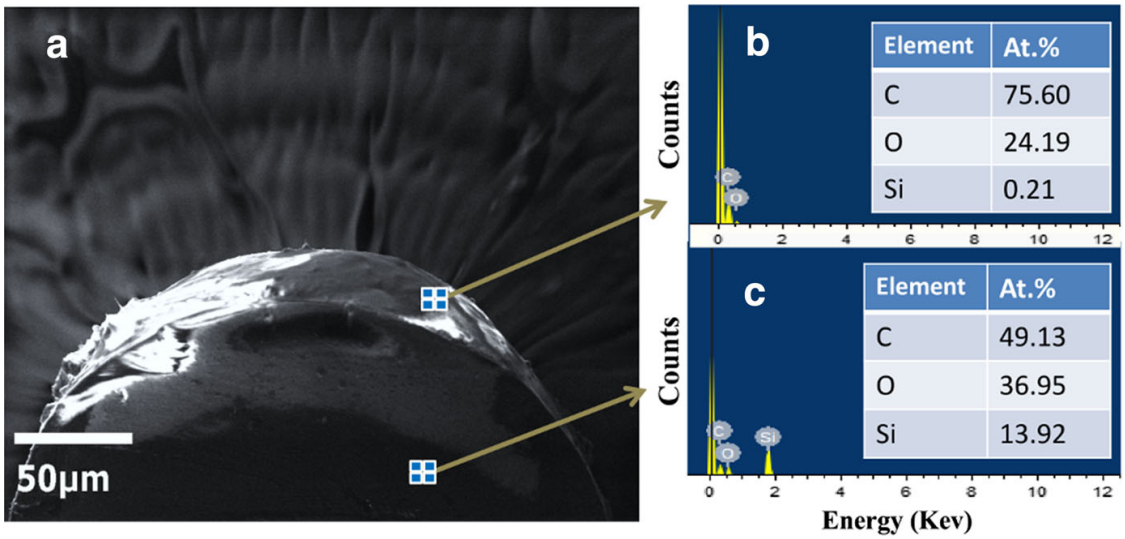


Fig. 2 Cut cross-section of single-coated microspheres **a** SEM image, and EDX analysis **b** shell **c** core

practically negligible in comparison to that in the core, which confirms the formation of a core–shell structure.

Coating of PDMS core with epoxy also resulted in a change in the FTIR spectra (Figure S2, Supplementary information). Characteristic absorption peaks at 802 and 1,258 cm^{-1} due to the $(CH_3)_2SiO$ group vibration and broad absorption at $\sim$1,000–1,130 cm^{-1}, due to the Si–O–Si stretching could be observed in the spectra of the core PDMS, while epoxy-coated microspheres exhibited additional absorption bands at 1,100 cm^{-1} due to the asymmetrical stretching vibration of polyether. Appearance of absorption in the region 1,250–1,360 cm^{-1} due to ν_s (C–N) and in 1,580–1,650 cm^{-1} due to ν_s (N–H) further confirms coating of epoxy on the PDMS surface. Coating the siloxane microspheres with epoxy resulted in a change in the thermal degradation behaviour and two-step degradation was observed in ECR (Figure S3, Supplementary information). The first step, due to pyrolytic decomposition of epoxy occurs at much lower temperatures (T_{onset} $\sim$250 °C) followed by a subsequent pyrolytic step resulting from the degradation of the core PDMS silicone at $\sim$400 °C leaving behind a residual char content of $\sim$65 %, consistent with previous studies.

The effect of coating on the chemical resistance was measured by gravimetric studies which reveal that PDMS microspheres were chemically inert, with the mass loss being practically negligible when placed in contact with common solvents like methanol, DMF, THF, chloroform, water and toluene. However, in the coated microspheres, the surface epoxy layer turned brittle on extended exposure to methanol, chloroform and toluene where a mass loss of 15 $\pm$ 4 % was observed.

Rubber-toughened unsaturated polyester composites

USP Composites containing both PDMS microsphere and core–shell microspheres were prepared by dispersing them within the unsaturated resin by ultrasonication followed by room temperature curing in the presence of cobalt octoate and MEKP.

The glass transition temperatures (T_g) of the pristine cured unsaturated polyester resin were 67.3 °C. As expected, the introduction of the elastomeric microspheres was found to have no significant reduction in T_g values of the base polymer, and the traces appear to overlap (Fig. 3). This is in contrast with the results reported on blends with organic liquid rubbers where significant reduction in T_g is observed. Our studies clearly highlight the benefit of employing preformed elastomeric microspheres as impact modifiers as the second phase remains phase separated, in view of which their plasticizing action is practically negligible. The TG trace of USP and representative composites containing PDMS and core–shell PDMS-epoxy is presented in Figure S4. It was observed that the introduction of elastomeric fillers do not affect the thermal properties of the base resin and all the compositions can be used in service till 300 °C, irrespective of the type of filler employed, Fig. 4.

Figure 5 presents the variation in mechanical properties of USP containing both neat PDMS as well as ECR microspheres. As expected, introduction of elastomeric microspheres led to decrease in the tensile strength, the extent of which was found to be proportional to the amount of loading while the tensile strain remained largely unaffected. Previous studies have revealed that silicones resulted in less severe loss in mechanical properties, in comparison to other acrylate-based elastomers, and this is one of the primary reasons for advocating their potential in high-end applications [28].

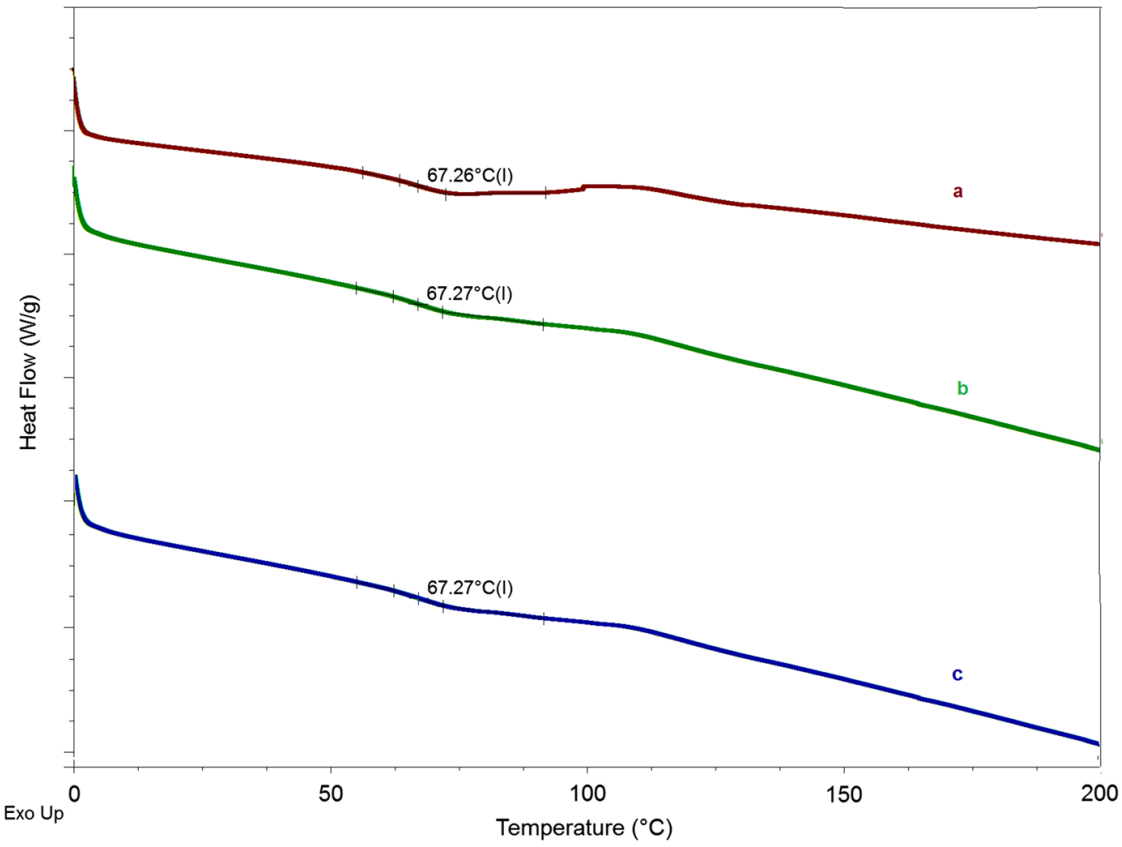


Fig. 3 DSC trace of *a* USP, *b* UP10R and *c* UP10ECR

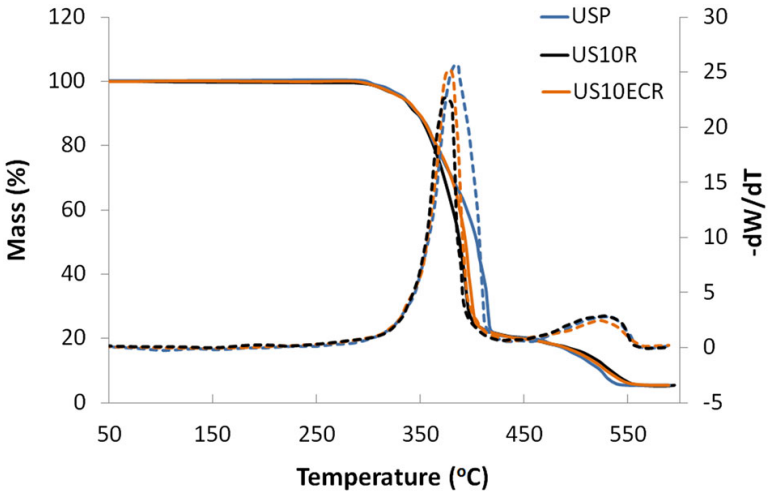


Fig. 4 TG–DTG traces of cured unsaturated polyester and its composites with neat and core–shell rubber

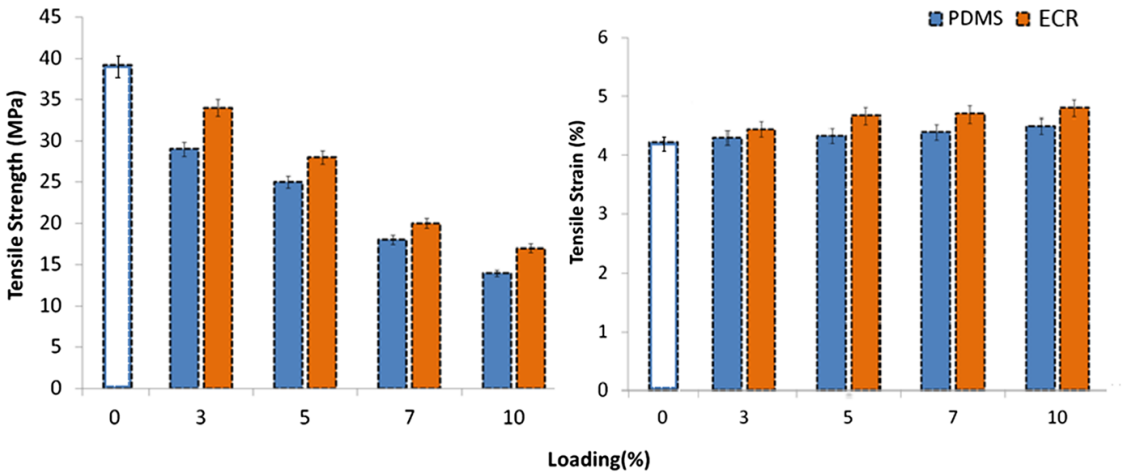


Fig. 5 Effect of rubber loading on the tensile strength and strain of USP composites

The values of the experimentally obtained moduli were compared with two existing theoretical models, namely Halpin–Tsai and the Lewis–Nielsen models, which are more representative in predicting the modulus of composites as a function of the modulus of neat USP, E_m, and that of silicone rubber filler, E_f [29, 30].

According to the Halpin–Tsai model [34], the modulus is predicted as

$$E = \frac{1 + \zeta \eta V_f}{1 - \eta V_f} E_m \tag{4}$$

where ζ is the shape factor, V_f is the volume fraction of the microsphere, and

$$\eta = \frac{\left(\frac{E_f}{E_m} - 1\right)}{\left(\frac{E_f}{E_m} + \zeta\right)}. \tag{5}$$

The shape factor in the Halpin–Tsai model is a function of the aspect ratio of the filler particles, which is usually calculated as $\zeta = 2w/t$, where w and t are the length and thickness of the reinforcement, respectively [31]. In view of the spherical nature

of the microspheres used in the present work, a shape factor of $\zeta = 2$ was used for the prediction of the modulus. The value of moduli (E_f) of polysiloxane elastomer has been assumed to be 2.5 MPa, based on the previous studies [32].

The basic Lewis–Nielsen model [33], takes into account the effect of the adhesion between the polymer and the fillers by employing the following equation to predict the modulus of the composite:

$$E = \frac{1 + (k_E - 1)\beta V_f}{(1 - \beta\mu V_f)} E_m, \tag{6}$$

where, k_E is the generalised Einstein coefficient and β and μ are constants. The constant β depends on the ratio of E_f and E_m and is given by:

$$\beta = \frac{\left(\frac{E_f}{E_m} - 1\right)}{\left(\frac{E_f}{E_m} + (k_E - 1)\right)}. \tag{7}$$

The constant μ being dependant on the maximum volume fraction of the filler (V_{max}) by:

$$\mu = 1 + \frac{(1 - V_f)}{(V_{max})}[V_{max}V_f + (1 - V_{max})(1 - V_f)]. \tag{8}$$

The values of V_{max} for a range of particle shapes and types of packing have been tabulated by Nielsen and Landel [34]. The morphological studies indicated that all the microspheres were randomly dispersed, therefore, a figure of 0.632 for V_{max} was employed for the present studies, which is representative of random close-packed and non-agglomerated spheres [34]. The value of the generalised Einstein coefficient, k_E, reportedly varies with the Poisson's ratio of the polymeric matrix and the degree of the adhesion of the polymer to the particles. Since the present investigation deals with USP resin ($v = 0.4$), where no slippage at the interface between the polymeric matrix and the particles was evidenced from the morphological investigations, a value of $k_E = 2.167$ was used [34, 35].

The experimentally observed values were compared with the predictions based on the above-mentioned models, the results of which are presented in Fig. 6. Excellent agreement can be seen, with the modulus of composites containing ECR microspheres being relatively higher. However, for both the cases, the experimental data were found to generally lie between the Halpin–Tsai and Lewis–Nielsen predictions, where the Halpin–Tsai model yields the upper bound and the Lewis–Nielsen model gives the lower bound figure.

Interestingly, the introduction of elastomeric microspheres led to a remarkable improvement in the impact strength of USP, as is evident from data presented in Fig. 7. The Charpy impact strength increased from 8.6 J m^{-1} (unfilled resin) to 16.1 J m^{-1} on addition of 5 % w/w ECR (~101 % increase). In comparison, the improvement in impact strength was less pronounced when uncoated PDMS was used as the filler. Under similar loadings of ECR, the impact strength of epoxy has been reported to increase from 22.5 J m^{-1} (unfilled resin) to 55.8 J m^{-1} [25]. In

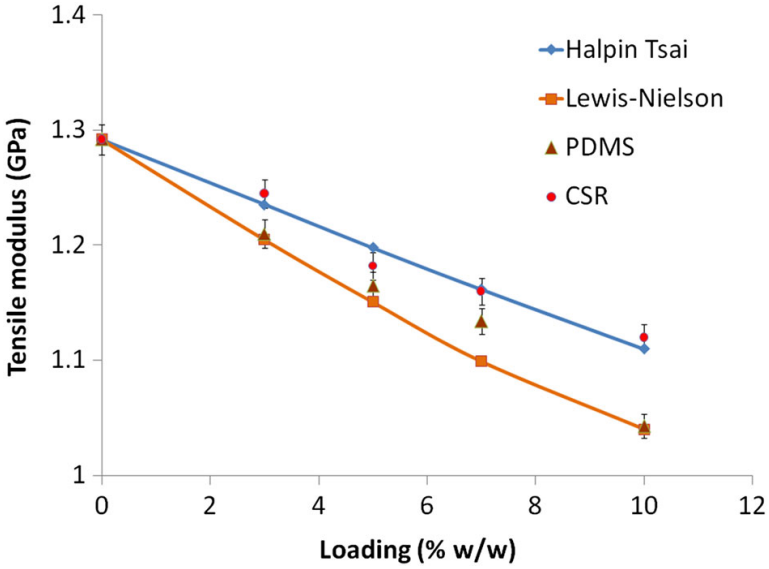


Fig. 6 Comparison of experimental modulus of toughened compositions with Halpin–Tsai and Lewis–Nielson models

comparison, introduction of halloysite nanotubes (3 %) has been reported to increase the impact strength by 29 % [36].

Flexural three-point bending tests were also performed on specimens, and as expected, the composites exhibited increased flexural strain in comparison to the pristine USP, which can be attributed to the high extensibility of the elastomeric domain present in the composites. The ECR containing samples could be flexed to a higher extent, which in turn reflected in larger values of critical stress intensity factor (K_{IC}). A comparison of K_{IC} and fracture energy (G_{IC}) as a function of composition type is presented in Fig. 8.

It can be seen that the critical stress intensity factor (K_{IC}) of the USP increased with the addition of the elastomeric microspheres, from 0.41 MPa $m^{1/2}$ for the unmodified USP to 0.72 MPa $m^{1/2}$ for composites containing 5 % w/w of ECR, which corresponds to an increase of 76 %. The mean fracture energy of the unmodified USP (110 J/m^2) increased substantially on addition of elastomeric microspheres to a maximum value of 400 J/m^2 for composition containing 5 % w/w of the ECR microspheres, which is a significant improvement ($\sim$263 %) in fracture energy as compared to the neat resin. Under similar loadings of ECR, the critical stress intensity factor (K_{IC}) of epoxy was found to increase from 2.17 MPa $m^{1/2}$ (unfilled resin) to 2.81 MPa $m^{1/2}$ [25], which was associated with an increase of 70 % in the mean fracture energy. Previously, introduction of 4.5 % (v/v) Al_2O_3 nanoparticles reportedly led to $\sim$100 % increase in the fracture toughness of unsaturated polyester [11].

Toughening mechanisms

The morphology of the fractured surface was studied by SEM imaging and the secondary electron images are presented in Fig. 9. As can be seen, the fractured surface of USP is smooth and featureless and exhibits no sign of plastic deformation

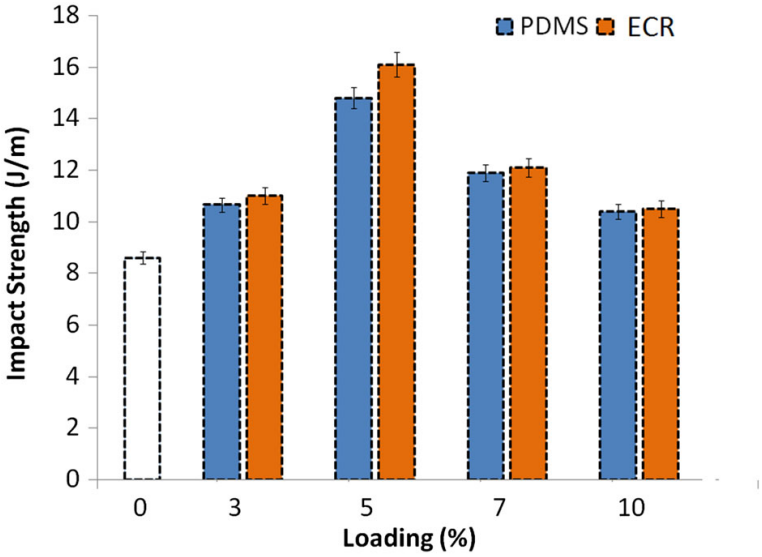


Fig. 7 Impact strength of composites prepared in the presence of neat PDMS and core–shell rubbers

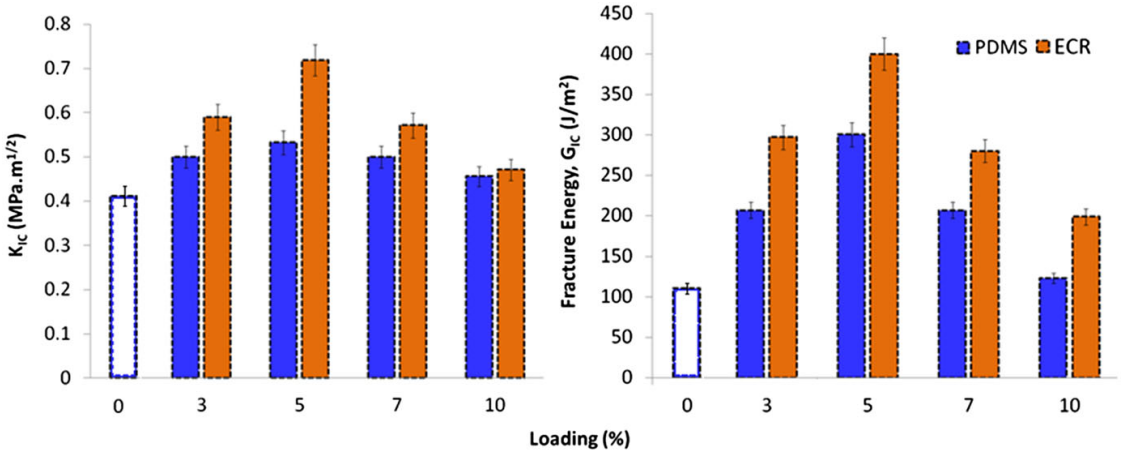


Fig. 8 Increase in the critical stress intensity factor, K_{IC}, and fracture energies, G_{IC}, of USP due to introduction of elastomeric microspheres

(Fig. 9a), which is typical behaviour of a brittle thermosetting polymer [28]. In comparison, the surface of the fractured composite is rather rough and shows the presence of holes.

The most important criteria for the elastomeric microspheres to exhibit their full potential as impact modifiers are their existence in the matrix as in a well-separated phase. It appears that the surface layer of epoxy in the ECR is compatible with the unsaturated polyester resin, which leads to its homogenous dispersion as compared to its uncoated analogue, and the same is evident from the SEM images. However, with increased loadings, agglomeration of these microspheres was observed which was responsible for the decrease in the mechanical properties. The foremost contributing route behind the improved toughenability of silicone–USP composites appears to be "rubber cavitation" [12, 32, 37]. During the tensile loading, the plastic region surrounding the microspheres tend to dilate, and the silicone microspheres, in

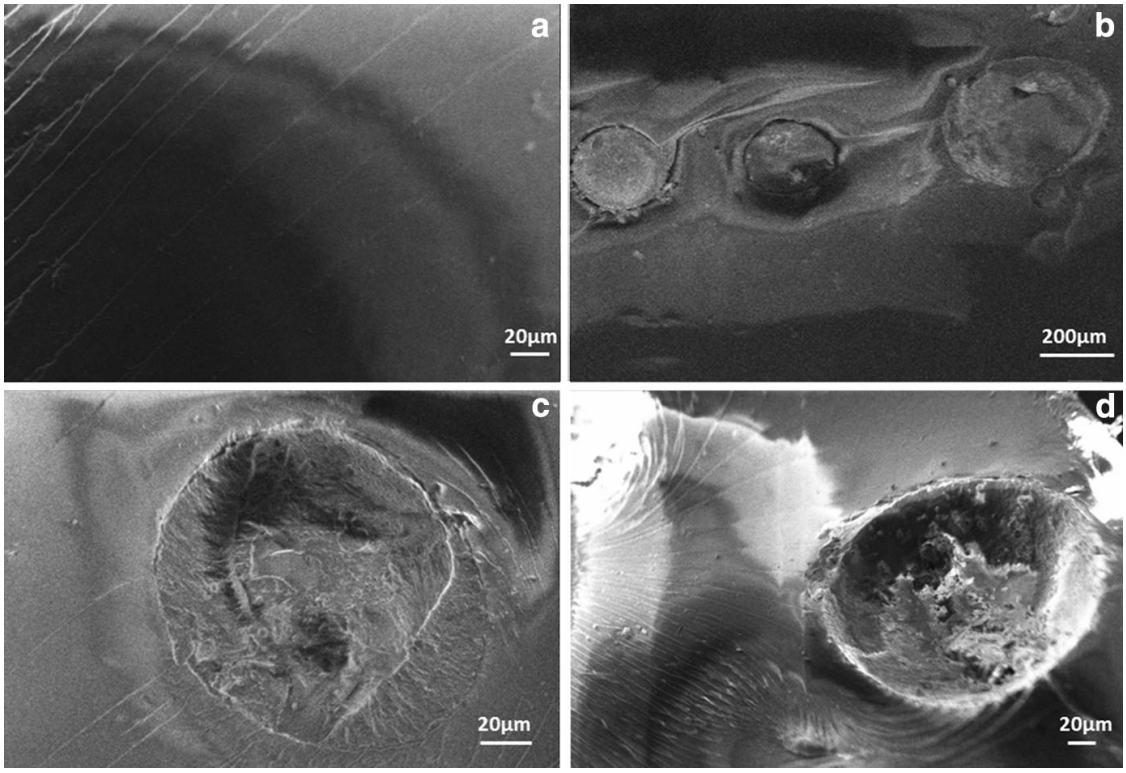


Fig. 9 SEM image of the fractured surface of **a** neat USP, **b** USP composite containing 5 % PDMS **c**–**d** USP composite containing 5 % CSR

view of their large extensibility ($v = 0.499$), cavitate from within [38]. The particle cavitation basically relieves the plane strain constraint from the surrounding matrix thereby allowing plastic deformation within the matrix [39]. This mechanism can be used explaining the presence of voids with dimensions much larger than that of the microspheres which were introduced during the curing process.

Another mechanism which can be used to explain the toughening in these toughened composites is that of particle-yielding-induced shear banding [40]. The siloxane-based microsphere filler tends to yield, which produces significant stress concentration, which in turn initiates shear banding in the matrix. In addition to these, other features indicative of crack path deflection and microcracking were also observed on the rough texture of the fracture surface [41]. As a result of these phenomena, the surface area of the crack increases and the mode I character of the crack opening is reduced, thereby resulting in increased energy for crack propagation.

Conclusion

The potential of core–shell rubber towards improving the dynamic properties of unsaturated polyester has been explored. The core elastomeric beads of PDMS microspheres were prepared by suspension polymerisation process, which involved the addition reaction of silicone macromonomer in the presence of platinum-based

hydrosilylation catalyst. The microspheres were subsequently coated with an epoxy layer in a separate step to prepare core–shell rubbers, to improve their compatibility with the USP resin in the subsequent step of composite preparation. Both uncoated as well as epoxy-coated PDMS microspheres were dispersed in USP matrix to prepare toughened composites (3–10 % w/w). The effect of addition of these fillers on mechanical properties (both quasi-static and dynamic) was studied. The presence of a superficial epoxy coating on the elastomeric core seems to play an important role in dispersion of the microspheres, as evident from the improved mechanical properties in the epoxy-coated rubber (ECR)-loaded compositions. Impact strength and fracture energy increased by $\sim$101 and 76 %, respectively, on addition of 5 % w/w ECR. The reduction in the experimentally observed tensile modulus was compared with two analytical models which revealed excellent agreement. Spherical cavities were observed in the fracture surface of composites which substantiate the role of rubber cavitation as the primary toughening mechanism in microsphere-toughened USP composites.

Acknowledgments The authors gratefully acknowledge Director, Centre for Fire, Explosive and Environment Safety, Delhi, India for providing logistic support to perform this work. The authors also thank Dipesh Gupta, Revex Plasticizers Private Limited, Delhi, India, for supplying the unsaturated polyester and Madhurima Baral for performing mechanical tests.

References

1. Bauer RS, Stewart SL, Stenzenberger HD (2000) Composite materials, thermoset polymer-matrix. In: Kirk-Othmer E (ed) Encyclopedia of chemical technology. Wiley, USA. doi:10.1002/0471238961.2008051802012105.a01
2. Ratna D, Banthia A (2004) Rubber toughened epoxy. Macromol Res 12(1):11–21. doi:10.1007/bf03218989
3. Bhuvaneswary MG, Thachil ET (2008) Blends of natural rubber with unsaturated polyester resin. Int J Polym Mater Polym Biomater 57(6):543–554. doi:10.1080/00914030701818272
4. Chuayjuljit S, Siridamrong P, Pimpan V (2004) Grafting of natural rubber for preparation of natural rubber/unsaturated polyester resin miscible blends. J App Polym Sci 94(4):1496–1503. doi:10.1002/app.21064
5. Chaudhary S, Parthasarathy S, Kumar D, Rajagopal C, Roy PK (2014) Poly(ethyleneterephthalate) glycolysates as effective toughening agents for epoxy resin. J App Polym Sci 131(4). doi:10.1002/app.39941
6. Chaudhary S, Parthasarathy S, Kumar D, Rajagopal C, Roy PK (2014) Graft-interpenetrating polymer networks of epoxy with polyurethanes derived from poly(ethyleneterephthalate) waste. J App Polym Sci 131:40490. doi:10.1002/app.40490
7. Hodgkin JH, Simon GP, Varley RJ (1998) Thermoplastic toughening of epoxy resins: a critical review. Polym Adv Tech 9(1):3–10. doi:10.1002/(sici)1099-1581(199801)9:1<3:aid-pat727>3.0.co;2-i
8. Yıldız E, Özarslan Ö, İnan TY, Kuyulu A, Güngör A (2007) Toughening of epoxy resins by amine terminated poly(arylene ether ketone)s having pendant tertiary butyl groups. Polym Bull 58(3):503–511. doi:10.1007/s00289-006-0689-x
9. Chaudhary S, Parthasarathy S, kumar D, Rajagopal C, Roy PK (2014) Simple toughening of epoxy thermosets by preformed thermoplastics. Soc Plast Eng Plast Res Online. doi:10.2417/spepro.005409
10. Chaudhary S, Parthasarathy S, Kumar D, Rajagopal C, Roy PK (2014) Amine functionalised poly(styrene) microspheres as thermoplastic toughener for epoxy resin. Polym Comp. doi:10.1002/pc.22927
11. Zhang M, Singh RP (2004) Mechanical reinforcement of unsaturated polyester by AL2O3 nano-particles. Mater Lett 58(3–4):408–412. doi:10.1016/S0167-577X(03)00512-3

12. Roy PK, Ullas AV, Chaudhary S, Mangla V, Sharma P, Kumar D, Rajagopal C (2013) Effect of SBA-15 on the energy absorption characteristics of epoxy resin for blast mitigation applications. Iran Polym J 22(9):709–719. doi:10.1007/s13726-013-0169-8
13. Hayes BS, Seferis JC (2001) Modification of thermosetting resins and composites through preformed polymer particles: A review. Polym Comp 22(4):451–467. doi:10.1002/pc.10551
14. Levita G, Marchetti A, Lazzeri A (1991) Toughness of epoxies modified by preformed acrylic rubber particles. Makromol Chem Macromol Symp 41(1):179–194. doi:10.1002/masy.19910410115
15. Pearson RA, Yee AF (1991) Influence of particle size and particle size distribution on toughening mechanisms in rubber-modified epoxies. J Mater Sci 26(14):3828–3844. doi:10.1007/bf01184979
16. Maazouz A, Sautereau H, Gérard JF (1994) Toughening of epoxy networks using pre-formed core-shell particles or reactive rubbers. Polym Bull 33(1):67–74. doi:10.1007/bf00313475
17. Lovell PA (1995) An overview of the preparation and use of emulsion polymer particles for the toughening of plastics. Macromol Symposia 92(1):71–81. doi:10.1002/masy.19950920108
18. Bécu L, Sautereau H, Maazouz A, Gérard JF, Pabon M, Pichot C (1995) Synthesis and structure–property relationships of acrylic core–shell particle-toughened epoxy networks. Polym Adv Tech 6(5):316–325. doi:10.1002/pat.1995.220060510
19. Hsieh TH, Kinloch AJ, Masania K, Sohn Lee J, Taylor AC, Sprenger S (2010) The toughness of epoxy polymers and fibre composites modified with rubber microparticles and silica nanoparticles. J Mater Sci 45(5):1193–1210. doi:10.1007/s10853-009-4064-9
20. Murias P, Maciejewski H, Galina H (2012) Epoxy resins modified with reactive low molecular weight siloxanes. Eur Polym J 48(4):769–773. doi:10.1016/j.eurpolymj.2012.01.009
21. Kunz-Douglass S, Beaumont PWR, Ashby MF (1980) A model for the toughness of epoxy-rubber particulate composites. J Mater Sci 15(5):1109–1123. doi:10.1007/bf00551799
22. Dompas D, Groeninckx G (1994) Toughening behaviour of rubber-modified thermoplastic polymers involving very small rubber particles: 1. A criterion for internal rubber cavitation. Polymer 35(22):4743–4749. doi:10.1016/0032-3861(94)90727-7
23. Lin M, Chu F, Guyot A, Putaux J-L, Bourgeat-Lami E (2005) Silicone–polyacrylate composite latex particles. Particles formation and film properties. Polymer 46(4):1331–1337. doi:10.1016/j.polymer.2004.11.063
24. Lin K-F, Shieh Y-D (1998) Core–shell particles to toughen epoxy resins. I. Preparation and characterization of core–shell particles. J App Polym Sci 69(10):2069–2078. doi:10.1002/(sici)1097-4628(19980906)69:10<2069::aid-app18>3.0.co;2-x
25. Roy P, Iqbal N, Kumar D, Rajagopal C (2014) Polysiloxane-based core-shell microspheres for toughening of epoxy resins. J Polym Res 21(1):1–9. doi:10.1007/s10965-013-0348-5
26. Knott JF (1976) Fundamentals of fracture mechanics. Butterworths, London
27. De La Caba K, Guerrero P, Gavaldá J, Mondragon I (1999) Fracture behavior–morphology relationships in an unsaturated polyester resin modified with a liquid oligomer. J Polym Sci Part B Polym Phys 37(14):1677–1685. doi:10.1002/(sici)1099-0488(19990715)37:14<1677:aid-polb12>3.0.co;2-f
28. Voo R, Mariatti M, Sim LC (2012) Flexibility improvement of epoxy nanocomposites thin films using various flexibilizing additives. Compos Part B Eng 43(8):3037–3043. doi:10.1016/j.compositesb.2012.05.032
29. Giannakopoulos G, Masania K, Taylor AC (2011) Toughening of epoxy using core–shell particles. J Mater Sci 46(2):327–338. doi:10.1007/s10853-010-4816-6
30. Johnsen BB, Kinloch AJ, Mohammed RD, Taylor AC, Sprenger S (2007) Toughening mechanisms of nanoparticle-modified epoxy polymers. Polymer 48(2):530–541. doi:10.1016/j.polymer.2006.11.038
31. Affdl JCH, Kardos JL (1976) The Halpin-Tsai equations: a review. Polym Eng Sci 16(5):344–352. doi:10.1002/pen.760160512
32. Chen J, Kinloch AJ, Sprenger S, Taylor AC (2013) The mechanical properties and toughening mechanisms of an epoxy polymer modified with polysiloxane-based core-shell particles. Polymer 54(16):4276–4289. doi:10.1016/j.polymer.2013.06.009
33. McGee S, McGullough RL (1981) Combining rules for predicting the thermoelastic properties of particulate filled polymers, polymers, polyblends, and foams. Polym Comp 2(4):149–161. doi:10.1002/pc.750020403
34. Nielsen LE, Landel RF (1994) Mechanical properties of polymers and composites. Marcel Dekker, Boca Raton
35. Kinloch AJ (1987) Adhesion and adhesives: science and technology. Chapman & Hall, London
36. Albdiry MT, Ku H, Yousif BF (2013) Impact fracture behaviour of silane-treated halloysite nanotubes-reinforced unsaturated polyester. Eng Fail Anal. doi:10.1016/j.engfailanal.2013.06.027

37. Pucciariello R, Villani V, Bianchi N, Braglia R (1991) Impact behaviour and morphology of rubber-modified epoxy resins. Die Angewandte Makromolekulare Chemie 187(1):75–86. doi:10.1002/apmc.1991.051870108
38. Bagheri R, Pearson RA (1996) Role of particle cavitation in rubber-toughened epoxies: 1. Microvoid toughening. Polymer 37(20):4529–4538. doi:10.1016/0032-3861(96)00295-9
39. Bagheri R, Pearson RA (2000) Role of particle cavitation in rubber-toughened epoxies: II. Inter-particle distance. Polymer 41(1):269–276. doi:10.1016/S0032-3861(99)00126-3
40. Chen T, Jan Y (1991) Toughening mechanism for a rubber-toughened epoxy resin with rubber/matrix interfacial modification. J Mater Sci 26(21):5848–5858. doi:10.1007/bf01130124
41. Pearson RA, Yee AF (1993) Toughening mechanisms in thermoplastic-modified epoxies: 1. Modification using poly(phenylene oxide). Polymer 34(17):3658–3670. doi:10.1016/0032-3861(93)90051-B

Study of diffuse phase transition in $Pb(Cd_{1/3}Nb_{2/3})O_3$ compound

M. Pastor, Nawnit Kumar, Bineet Kumar, A. Panwar, K. Biswas, A.C. Pandey

PII: S0925-8388(14)01366-8
DOI: http://dx.doi.org/10.1016/j.jallcom.2014.06.032
Reference: JALCOM 31448

To appear in: *Journal of Alloys and Compounds*

Received Date: 2 July 2013
Revised Date: 5 June 2014
Accepted Date: 5 June 2014

Please cite this article as: M. Pastor, N. Kumar, B. Kumar, A. Panwar, K. Biswas, A.C. Pandey, Study of diffuse phase transition in $Pb(Cd_{1/3}Nb_{2/3})O_3$ compound, *Journal of Alloys and Compounds* (2014), doi: http://dx.doi.org/10.1016/j.jallcom.2014.06.032

Study of diffuse phase transition in $Pb(Cd_{1/3}Nb_{2/3})O_3$ compound.

M. Pastor[1,6a], Nawnit Kumar[2], Bineet Kumar[3], A. Panwar[4], K. Biswas[5], and A.C. Pandey[1,6]

[1]Nanotechnology Application Centre, University of Allahabad, Allahabad U.P-211002 India

[2]Department of Physics and Meteorology, Indian Institute of Technology, Kharagpur, (W.B.) -721302, India

[3]Department of Physics, Acharya Narendra Dev College,

University of Delhi, New Delhi 110019 India

[4] Department of Applied Physics, Delhi Technological University, Delhi -110042, Delhi.

[5]Department of Metallurgical & Materials Engineering, Indian Institute of Technology, Kharagpur, (W.B.) -721302, India

[6]Bundelkhand University, Jhansi (U.P.) 284002, India

Abstract

Present paper contains study of phase evolution, dielectric properties and complex impedance spectroscopy of lead based ferroelectric ceramic $Pb(Cd_{1/3}Nb_{2/3})O_3$ synthesized by double step solid state reaction route. X-ray diffraction pattern shows that the material has single cubic phase. Temperature dependence of relative dielectric constant exhibits diffuse phase transition in the present compound. Frequency dependence of the real (Z') and imaginary (Z'') parts of the electric impedance spectroscopy reveal contribution due to grain (bulk) resistance in a wide temperature range and indicate that the relaxation mechanism is less temperature dependent and strongly influenced byoxygen vacancies and charge carrier. The conductivity spectrum has been observed to follow the Jonscher's universal power law. The existence of week ferroelectricity is confirmed from hysteresis loop study.

Keywords: ceramics; impedance spectroscopy; X-ray diffraction; dielectric properties; ferroelectricity.

[1a]Corresponding author: mukul.ptr@gmail.com

Tel: +91-3222-283244, Fax: +91-3222282220

1. *Introduction*

Lead based ferroelectric oxides with a general formula $Pb(B'_{1/3}B''_{1/3})O_3$ (where B' = Mg, Zn, Ni, etc.; B"=Nb and Ta) have drawn attention because of their attractive relaxor characteristics or diffuse phase transition (DPT) useful for a wide range of industrial application such as electrical, optical, electromechanical devices [1-4]. These materials exhibit excellent physical properties (i.e., high dielectric constant, frequency dependent broad maximum as a function of temperature, large electrostriction and small hysteresis loop) for device applications [3-6]. It is commonly believed that the diffuse phase transition behaviour of the materials is due to the presence of two or more cations of different ionic radii and valencies on the octahedral B-site of the perovskite ABO_3 (A=mono or divalent and B=trivalent to tetravalent cations). The change in physical properties are associated with ionic radii, charge distribution and competition between B'-site and B"-site ordering. Many of the materials of the perovskite family do not give stable and single perovskite phase, and/or solid solutions with normal ferroelectric behaviour. They provide high dielectric constant and low dissipation factor near morphotropic phase boundary [7-9]. Most common problem in the synthesis of lead-based single-phase ceramics is the strong inclination of this type of material to form detrimental secondary pyrochlore (Py) phase(s). The presence of the secondary phases reduces the dielectric, electrostrictive and other properties of the materials [7, 8]. Though various attempts have been made to solve this problem using various synthesis routes, columbite synthesis method has been found very effective to increase the percentage of perovskite to the higher possible level (i.e., to obtain single phase material) [10]. In this methodology, divalent oxides (B'O) /carbonate ($B'CO_3$) and Nb_2O_5 are taken as precursor compounds to obtain another $B'Nb_2O_6$ precursor, and in sequence, it reacts with a stoichiometric amount of PbO to obtain $Pb(B'_{1/3}Nb_{2/3})O_3$ powder with dominant perovskite phase. The main advantages of columbite method are: (a) requires the short time and low temperature heat treatment, (b) better homogeneity and (c) uniform cation distribution. To obtain pure perovskite compound $Pb(B'_{1/3}Nb_{2/3})O_3$, some essential steps have been followed. PbO volatilization or deficiency is a main source of pyrochlore phase of $Pb(B'_{1/3}Nb_{2/3})O_3$ compounds. Though several synthesis techniques have been tried to avoid the formation of undesirable pyrochlore phase [11-13], addition of small amount (5-10%) of PbO (to compensate for PbO volatilization) was found effective [14-16]. The role of excess amount of PbO in the formation of pure in $Pb(B'_{1/3}Nb_{2/3})O_3$ is still not clear. Only indirect effects of PbO excess have been reported on the dielectrics and micro-structural properties in PMN bulk ceramics [13-15]. The present work is mainly aimed at study of ferroelectric/relaxor properties in $Pb(Cd_{1/3}Nb_{2/3})O_3$ compound

synthesized without evolution of any secondary phase by double step solid state reaction route (b) change in morphology of substitution of Cd at the A site. No attempt has been made to study the Cd containing Pb compound (i.e., $Pb(Cd_{1/3}Nb_{2/3})O_3$) to our knowledge. Therefore we have carried out the perovskite phase, dielectric and electrical properties of the $Pb(Cd_{1/3}Nb_{2/3})O_3$ compound. Presented compound has been characterized by XRD, SEM, temperature dependence of dielectric study and impedance spectroscopy, which is used to study the structural and electrical properties of the compound, respectively.

2. *Experimental procedure*

The polycrystalline sample of $Pb(Cd_{1/3}Nb_{2/3})O_3$ was prepared by two step solid state reaction route or columbite method. First, columbite structure was formed by pre-reacting Nb_2O_5 (99.9 % Loba Chemie Pvt. Ltd., India) with $CdCO_3$ (99.9% pure, M/s B. D. H. Chemicals U. K) followed by calcination at 1100°C for 6 h. Calcined powder was then characterized by an X-ray diffraction technique (XRD) using X-ray diffractometer (Philips PW1710, target Co-K_α, $\lambda = 0.17905$ nm) to ensure the formation of columbite structure. This material was then reacted with PbO (with an excess of PbO 4 wt.%)(99.5 % Loba Chemie Pvt. Ltd., India) in a alumina crucible at 950°C for 6 h to get the pyrochlore free desired compound and then again characterized by X-ray diffraction to make sure the formation of perovskite phase. The fine and homogeneous powder of the above compound (with PVA binder) was pressed into cylindrical pellets (10 mm diameter and 1-2 mm thickness) under a uniaxial pressure of 300 MPa using a hydraulic press. The compacted pellets were fired first at 500°C to remove the binder and then sintered at 1000°C for 6 h to get highly dense (≥95 %) samples. Microstructural evolution was investigated using Carl Zeiss Supra 40 field emission scanning electron microscope (FESEM) The X-ray powder diffraction patterns of the first phase precursor and final phase material were taken at room temperature. The dielectric and impedance spectroscopic data were recorded on sintered pellets as a function of temperature from room temperature to 200°C using HIOKI 3532 LCR Hi TESTER, impedance analyzer. The studied frequency range and perturbation voltage were 100 Hz- 1 MHz and 1.2 V, respectively. The hysteresis loop of the poled sample was obtained using workstation of loop tracer (M/S Radiant Technology Inc., USA). Piezoelectric coefficient d_{33} is measured using Piezo-Meter (PM35 Take-control, UK) at 110 Hz (pressure 1 N m^{-2}) at room temperature of the poled sample. For the piezoelectric and P–E hysteresis loop measurements, the sample was poled at 110 °C temperature in a silicon oil bath under the dc field of 20 kV cm^{-1} for 1 h using APLAB high voltage dc power supply (model 7342P).

3. Results and discussions

3.1 Structural properties

Room temperature XRD pattern of 1100°C calcined $CdNb_2O_6$ powder confirm the formation of pure orthorhombic columbite phase [17]. Figure 1 shows the fitted pattern of X-ray diffraction (XRD) data of the final phase calcined compound. Observed peaks are indexed by using x-ray interpretation and analysis computer software X'pert High Score Plus. Best agreement between the observed and fitted profile suggests that the prepared material stabilized in cubic single phase and well matched with PDF no. 04-012-6370 (space group Pm-3m (2271)).

FESEM image is shown in figure 1 (b) which indicates dense microstructure. Grain distribution is inhomogeneous with 1.4 µm average grain size.

Density of the sintered pellet has been calculated by Archmedes' principle and found to be ~97%.

3.2 Dielectric properties

Temperature dependence of relative dielectric constant *(ε_r)* at different frequencies is shown in figure 2 (a). The variation of tangent loss *(tanδ)* with temperature is also shown in figure 2 (b). Both ε_r and *tanδ* show a broad maximum at ferro–paraelectric phase transition temperature *(T_m)*. The broadening of this dielectric peak increases with increasing frequency. Further, it is observed that the maximum value of ε_r. *(i.e., ε_{max})* decreases with increase in frequency. Dielectric constant increases with rise in temperature up to its maximum value *(ε_{max})* at the T_m, and then decreases with further increase in temperature. T_m shifts toward higher temperature side at higher frequencies (1, 5, 10, 25, 50, 100, 1000 kHz) (figure 2). The values of ε_{max} at T_m (60, 70, 74, 76, 80, 84 and 90°C) are 5466, 4532, 4023, 3767, 3552, 3417 and 3230 at 1, 5, 10, 25, 50, 100, 1000 kHz , respectively. The value of *tanδ* also increases with rise in temperature up to T_m and then decreases. The magnitude of *tanδ* also decreases with increasing frequencies and T_m shifts to higher ones. Broadening in dielectric peaks and small shift in T_m (i.e., ε_r and *tanδ* Vs temperature, shown in figure 2(a) and (b)) towards higher frequencies, shows diffuse phase transition. The rate of change of *tanδ* with temperature is relatively small (i.e., *tanδ* is almost temperature independent) at lower temperatures as compared to higher temperatures. At higher temperatures, *tanδ* increases sharply and then decreases with rise in temperature. It is clear from figure 2 (a) and (b) that frequency-dependent shift in the temperature (T_m) corresponding to dielectric peak in ε_r and *tanδ* vs *temperature* graph is not very significant . This type of behavior is observed due to presence of

significant amount of oxygen vacancies and charge carriers generated inside the material during sintering process. Temperature dependence of dielectric (ε_r and $tan\delta$) result is strongly influenced under the singly ionized oxygen vacancies due to the transition of Nb^{5+} to Nb^{4+}. Diffuse phase transition (peak broadening) exhibits more than one type of relaxation process involved the compound. Different types of relaxations due to cation disordering are mainly cause of peak broadening. This process can be explained with compositional fluctuation which leads growth of polar micro or nano-regions with different Curie temperature. Since, in ABO_3 type system, A B site cation disordering promotes the short range ordering, this is responsible for the formation of micro or nano-polar regions [18]. In present compound, cation disordering at Cd/Nb site manipulate the peak at Curie temperature. Presence of oxygen vacancy/defects/charge carriers also plays an important role to break the long range ordering which is ultimately responsible for diffuse phase transition. Therefore relaxation processes in the present compound may be the consequence of either B-site cation disordering or oxygen vacancy present in the system or due to both. It is also observed that oxygen vacancy plays a dominant role in development of ferroic domains rather than B site cation disordering which is the cause of peak diffuseness [19]. Formation of oxygen vacancy is not an intrinsic property of the system, but it has an important role in oxygen vacancy conductivity relaxation process. Strong dielectric anomaly at low frequencies provides that this phenomenon is predominant for low- frequency relaxation process [20]. Transition due to oxygen vacancies and charge carriers cover-ups the ferroelectric phase transition in the present material. In this consequence there is no obvious frequency-dependent change in the temperature (T_m) corresponding to dielectric peak and it is also noticeable that the temperatures (T_m) corresponding to ε peak and loss peak appears almost the same at a fixed frequency.

For a classic ferroelectric in a certain temperature, relative dielectric permittivity ε_r follows the Curie-Weiss law

$$1/\varepsilon = (T-T_c)/C' \qquad (1)$$

where T_c is the Curie temperature and C' is the Curie-Weiss constant.. Figure 3 shows the *inverse relative dielectric permittivity* ($1/\varepsilon_r$) vs. temperature plot at 50 kHz. Data obtained from experimental measurement are fitted in Eq. 1. The fitting parameters are C=6.54×10^5 K (error=1.91×10^{-8}) and T_c=-70.93°C. The dielectric constant has good agreement with the Curie-Weiss law until the temperature reduces to a certain value, where the deviation from the Curie-Weiss law occurs. This deviation can be defined by ΔT_m as follows:

$$\Delta T_m = T_{dev} - T_m, \qquad (2)$$

The inverse dielectric constant starts to deviate from Curie-Weiss law at 76°C, known as burn or deviation temperature (T_{dev}). At 5 kHz, ΔTm =6.0°C.

3.3 Impedance studies

Variation of real part of impedance Z' with frequency is plotted at different temperatures in figure 4(a). The figure suggests that the value of resistance is temperature independent. At lower frequencies, Z' is almost constant and then decreases monotonically with increasing frequency. At higher frequencies, the value of Z' is almost constant which shows frequency independent behavior of resistance. The higher value of Z' at lower frequency and low temperature clearly suggests the higher value of polarization. It also suggests that the resistive grain boundaries become conductive. It is observed that the grain boundaries do not relax even at very high frequencies and higher temperatures. The frequency variation of imaginary part of impedance Z'' at different temperatures is shown in figure 4(b). The broadness of peak increases with rise in temperature and there is no shift in peak position. As temperature increases, significant broadening of the Z'' peaks suggests the existence of temperature dependent relaxation process in the system. The relaxation process may be due to the presence of immobile charge of different species at different temperature. The figure 4(b) also indicates the spreading of the relaxation time. This would imply that the relaxation process is temperature dependent, and there is apparently no single relaxation time. Therefore, the peaks occurring at different temperatures have their own discrete relaxation times. This behavior is apparently associated with space charge polarization. Temperature and frequency independent Z'' peak shift clearly shows that the resistance of the bulk material is almost constant with temperature. Figure 5 (a) shows the complex impedance spectrum (i.e., Nyquist plot *(i.e Z'' Vs Z'))* at different temperatures. It is interesting to note that individual contribution comes from grain component and is represented by single semicircle, which is corresponding to grain body contribution [21, 22]. Depressed semicircular arcs show relatively less-uniform grain distributions and variation of areas under semicircular arc being very small (figure5 (a)), reveals the variation of resistance is less temperature dependent. Figure 5(b) shows the scaling behavior of imaginary part of impedance *(i.e. Z''/Z''_{max} vs ω/ω_{max})*. The perfect overlapping of the curves at different temperatures into a single master curve indicates that there is no contribution of impedance dynamic processes at higher temperatures. This may be due to the migration of ions that takes place via hopping mechanism.

The frequency dependence of $M'(\omega)$ is shown in figure 6(a). Initially at low frequencies, the values of real part of electric modulus are constant and then start increasing with frequency in entire temperature range. Figure 6(b) shows the frequency dispersion behaviour of $M''(\omega)$ at different temperatures. It is seen that $M''(\omega)$ exhibits a symmetric transition between 50 kHz – 500 kHz (fig.5 (b)). Since, the behaviour of $M''(\omega)$ is less temperature dependent therefore almost single relaxation phenomena occurs in the material below and above the Curie temperature. There is no shift in relaxation frequency with increase in temperature. In high frequency region the M'' tends to M''_{∞}, which indicates the mobility of charge ions from short range to long range. Figure 7(a) shows the complex master modulus spectra at different temperature. It can be observed from the spectrum that bulk properties of the compound come from the grain capacitance of the materials below and above Curie temperature. Very little dispersion in the spectra for different temperature indicates that the compound has single relaxation process. Figure 7(b) show the scaling behaviour of the M'' by M''_{max} versus $log\ (\omega/\omega_{max})$ for each temperature. All curves tend to collapse into one within the experimental error. The overlapping of all the curves into one indicates that the dynamical processes are less temperature dependent.

The frequency variation of ac conductivity (σ_{ac}) at different temperatures is shown in figure 8. The behavior of σ_{ac} with frequency at room temperature exhibits both low and high frequency dispersion phenomena. This obeys the Jonscher's power law [23];

$$\sigma(\omega) = \sigma_{dc} + A\omega^{n} \qquad (2)$$

where n is the temperature dependent frequency exponent $(0 < n < 1)$, A is a thermally activated quantity, hence electrical conduction of the materials is a thermally activated process. According to Jonscher [22], the origin of the frequency dependence of conductivity is based on motion of mobile charge carriers. When a mobile charge carrier hops to a new site from its original position, it remains in a state of displacement between two potential energy minima. After a sufficiently long time, the defect could relax until the two minima of lattice potential energy coincide with the lattice site of the material. Also, the conduction behavior of the materials obeys the power law $\sigma(\omega) \propto .\ \omega^{n}$ with a slope of σ_{ac} curve governed by n in the low temperature region. According to Funke [24], the value of n has a physical meaning, $n \bullet 1$ for hopping motion involving a translational motion, whereas for $n>1$, the motion involves localized hopping without the species leaving the neighborhood. In the present sample, values of n lie in the range of 0.74-0.84 in the rage from −50°C to 200°C (given in table1). The frequency at which change in slope takes place is known as hopping frequency (ω_p), which is temperature

dependent. A small peak shift of ω_c is being observed with rise in temperature. The calculated value of hopping frequency (ω_p) is tabulated in table 1, at different temperatures.

3.4 Hysteresis loop study

Hysteresis areas as a function of the field amplitude E and frequency f or dynamic hysteresis has key role to play in many applications on the basis of signal amplitude and frequency, as well as in ferroelectric random access memories [25-28]. When a ferroelectric sample is submitted to a periodically varying external electrical field E of frequency f and amplitude E_0, a loop-like hysteresis is observed. The hysteresis area represents the energy dissipation. Hysteresis profiles at constant frequencies (20Hz) with variable electrical field are shown in figure 9 (a). The saturated (P_s) and remnant (P_r) polarization values were determined from figure 9(a) and are given in table 2. Hysteresis loop area and remnant polarization increases with an increase in field. In figure 9(a), the loop does not saturate and appears as an ellipse with sharp ends. The nature of hysteresis loop indicates the lossy behaviour of the compound [29]. When f is held unchanged (say f=20 Hz), area under the hysteresis curve, P_r and P_s are raised with applied field as shown in figure 9 (a). The contributions to the apparent ferroelectric loop possibly arise from the leakage, especially in the mechanical activation assisted samples where the loop essentially represents the integrated leakage current. In the presence of conducting phases to suppress the leakage, low temperature measurements are required.

Hysteresis profiles at different frequencies (20-100Hz) with constant electrical field are shown in figure 9(b). Area of the loop as well as the value of remnant polarization (P_r) initially decrease and then get steady as frequency increases.

At low frequencies oxygen vacancies and other charge carriers have prominent role in electrical properties. The effect of oxygen vacancies and charge carriers is also noticeable in hysteresis plot given in the manuscript. At low frequencies, hysteresis loop is not symmetric that is due to the role of oxygen vacancies and charge carrier.

Piezoelectric measurement of poled sample has been taken at room temperature. The obtained value of piezoelectric constant found to be d_{33} =25 pCN^{-1}.

Conclusions

The single phase polycrystalline sample of $Pb(Cd_{1/3}Nb_{2/3})O_3$ was prepared by solid state reaction route. A dielectric anomaly of $Pb(Cd_{1/3}Nb_{2/3})O_3$ suggests the possible existence of diffuse phase transition due to oxygen vacancies and charge carriers.

Detailed studies of real part of impedance Z' at lower frequencies suggest that various polarization mechanisms play the critical role in this system. Imaginary part of impedance Z" shows distributed dielectric relaxation phenomena at high temperatures, which is almost temperature-independent. Nyquist plot reveals that electrical conduction is mainly due to the bulk resistance. The electrical modulus analysis suggests polydispersive nature of the compound.

The ac conductivity spectrum was found to obey Jonscher's universal power law, because of hopping mechanism of conduction. The electrical conduction properties of this material appear to be strongly dependent on frequency and analysis confirms that electrical relaxation and conduction arise out of similar type of charge carriers. Dominant oxygen vacancies relaxation cover-ups relaxation due to ferroelectric property therefore it is difficult to find out the change in Impedance and AC conductivity near transition temperature range Low P_s and P_r values obtained from hysteresis loop study week ferroelectric type nature of the compound.

Acknowledgements

Dr. Mukul Pastor is thankful to UGC Dr D.S. Kothari fellowship (Grant no. F.4-2/2006(BSR)/13-360/2010(BSR)) for financial support. We are also grateful to Prof. R.N.P. Choudhary, Department of Physics and Meteorology, Indian Institute of Technology, Kharagpur and Prof. I. Manna, Department of Metallurgical and Materials Engineering, Indian Institute of Technology, Kharagpur for their kind suggestions and support in concern research.

References

[1] Chen SY, Wang CM, Cheng, SY (1997) The effect of pyrochlore phase on formation mechanism and electrical properties of perovskite PZMN relaxors. Mater Chem. Phys 49: 70–77.

[2] Kuwata J, Uchino K, Nomura S (1979) Diffuse Phase transitions in Lead Zinc Niobate. Ferroelctrics 22: 863-867.

[3] Randall CA, Bhalla AS, Shrout TR, Cross LE (1990) Classification and consequences of complex lead perovskite ferroelectrics with regard to B-site cation order. J Mater Res 5: 829-834.

[4] Randall CA, Bhalla AS (1990) Nanostructural-Property Relations in Complex Lead Perovskites. Jpn J Appl Phys 29: 327-333.

[5] Choi SW, Shrout TR, Jang SJ, Bhalla AS (1989) Dielectric and pyroelectric properties in the $Pb(Mg_{1/3}Nb_{2/3})O_3$–$PbTiO_3$ system. Ferroelectrics 100: 29–38.

[6] Shrout TR, Halliyal A (1987) Preparation of lead-based ferroelectric relaxors for capacitors. Am Ceram Soc Bull 66: 704-711.

[7] Lejeune M, Boilot JP (1986) Optimization of dielectric properties of lead magnesium niobate ceramics. Am Ceram Soc Bull 64: 679–682.

[8] Kang DH, Yoon KH (1988) Dielectric properties due to excess PbO and MgO in lead magnesium niobate ceramics. Ferroelectrics 87: 255–264.

[9] Halliyal A, Kumar U, Newnham RE, Cross LE (1987) Stabilization of the perovskite phase and dielectric properties of ceramics in the $Pb(Zn_{1/3}Nb_{2/3})O_3$–$BaTiO_3$ system. Am Ceram Soc Bull 66: 671–676.

[10] Swartz SL, Shrout TR (1982) Fabrication of perovskite lead magnesium niobate. Mater Res Bull 17: 1245-1250.

[11] Joy PA, Sridhar K (1997) Formation of lead magnesium niobate perovskite from niobate precursors having varying magnesium content. J Am Ceram Soc 80: 770-772.

[12] Cavalheiro AA, Zaghete MA, Paiva-Santos CO, Varela JA, Longo E (1999) Influence of synthesis and processing parameters of the columbite precursor on the amount of perovskite PMN. Mater Res 2: 255–260.

[13] Jang HM, Lee KM, Lee MH (1994) Stabilization of perovskite phase and dielectric properties of Pb(Zn, $Mg)_{1/3}Nb_{2/3}O_3$-$PbTiO_3$ ceramics prepared by excess constituent oxides. J Mater Res 9: 2634-2644.

[14] Villegas M, Fernández JF, Moure C, Durán P (1994) Preparation, microstructural development and dielectric properties of $Pb(Mg_{1/3}Nb_{2/3})O_3$-$Pb(Ti_xZr_{1-x})O_3$ multilayer ceramic capacitors. J. Mater. Sci. 29: 4999-5004.

[15] Wang HC, Schulze WA (1990) The role of excess magnesium oxide or lead oxide in determining the microstructure and properties of lead magnesium niobate. J Am Ceram Soc 73 (4): 825–832.

[16] Gupta SM, Kulkarni AR (1995) Role of excess PbO on the microstructure and dielectric properties of lead magnesium niobate. J Mater Res 10: 953–961.

[17] Pastor Mukul, Biswas Koushik (2013) Synthesis and electrical characterization of $Ba(Cd_{1/3}Nb_{2/3})O_3$ ferroelectric compound. Mater Chem Phys 139: 634-639.

[18] Setter N, Cross LE (1980) The role of B-site cation disorder in diffuse phase transition behavior of perovskite ferroelectrics. J Appl Phys 51:4356

[19] Mihailova B, Gospodinov M, Güttler B, Petrova D, Stosch R, Bismayer U (2007) Ferroic nanoclusters in relaxors: the effect of oxygen vacancies. J Phys Condens Matter 19(24):246220

[20] Kang BS, Choi SK, Park CH (2003) Diffuse dielectric anomaly in perovskite-type ferroelectric oxides in the temperature range of 400–700°C. J Appl Phys 94: 1904

[21] James AR, Srinivas K (1999) Low temperature fabrication and impedance spectroscopy of PMN-PT ceramics. Mater Res Bull 34: 1301-1310.

[22] Wu Y, Forbess MJ, Seraji S, Limmer SJ, Chou TP, Cao G (2001) Impedance study of $SrBi_2Ta_2O_9$ and $SrBi_2(Ta_{0.9}V_{0.1})_2O_9$ ferroelectrics. Mater Sci Engg B 86: 70-78.

[23] Jonscher AK (1977) The 'universal' dielectric response. Nature 276: 673-679.

[24] Funke K (1993) Jump relaxation in solid electrolytes. Prog Solid State Chem 22: 111-195.

[25] Sahoo PS, Panigrahi A, Patri SK, Choudhary RNP (2009) Ferroelectric phase transition in $Ba_4SrSmTi_3V_7O_{30}$ ceramics. Mater Lett 63: 864–866.

[26] Jaffe B, Cook WR, Jaffe H (1971) Piezoelectric Ceramics. Academic, New York.

[27] Scott JF, Alexe M, Zakharov ND, Pignolet A, Curran C, Hesse D (1998) Nano-phase SBT family ferroelectric memories. Integer Ferroelec 21: 1-14.

[28] Uchino K (2000) Ferroelectric Devices. Marcel Dekker, New York.

[29] Waser R, Böttger U, Grossmann M (2004) in Ferroelectric Random Access Memories, edited by Ishiwara H, Okuyama M, Arimoto Y, Springer-Verlag, Berlin, Heidelberg.

Figure captions

Figure 1. (a) X-ray diffraction pattern of $Pb(Cd_{1/3}Nb_{2/3})O_3$ on calcined powder at room temperature and (b) FESEM micrograph of sintered pellet.

Figure 2. (a) Temperature dependence of ε_r and (b) Variation of $tan\ \delta$ with temperature, at various frequencies

Figure 3. (a) Plots of 1/ε versus T at 5khz.

Figure 4. Frequency dependence of (a) real part of impedance *(Z')* and (b) imaginary part of impedance *(Z'')*, at different temperatures

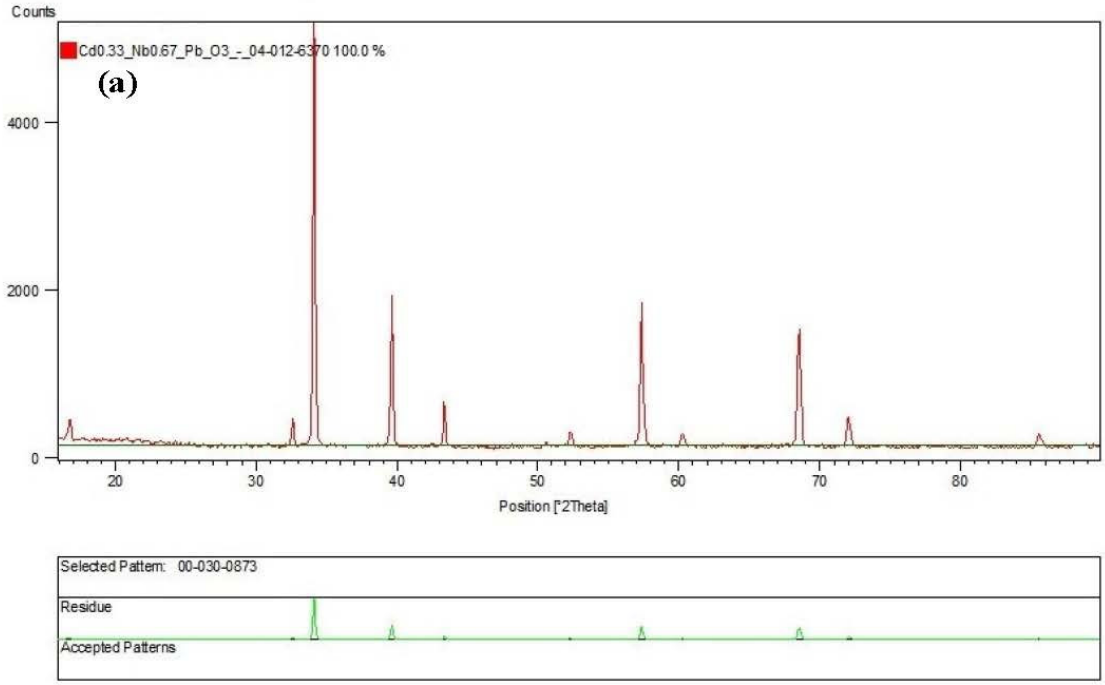
Figure 5. (a) Complex *Z'* vs. *Z''* (Nyquist) plots (b) Scaling behavior of *Z''* at various temperatures

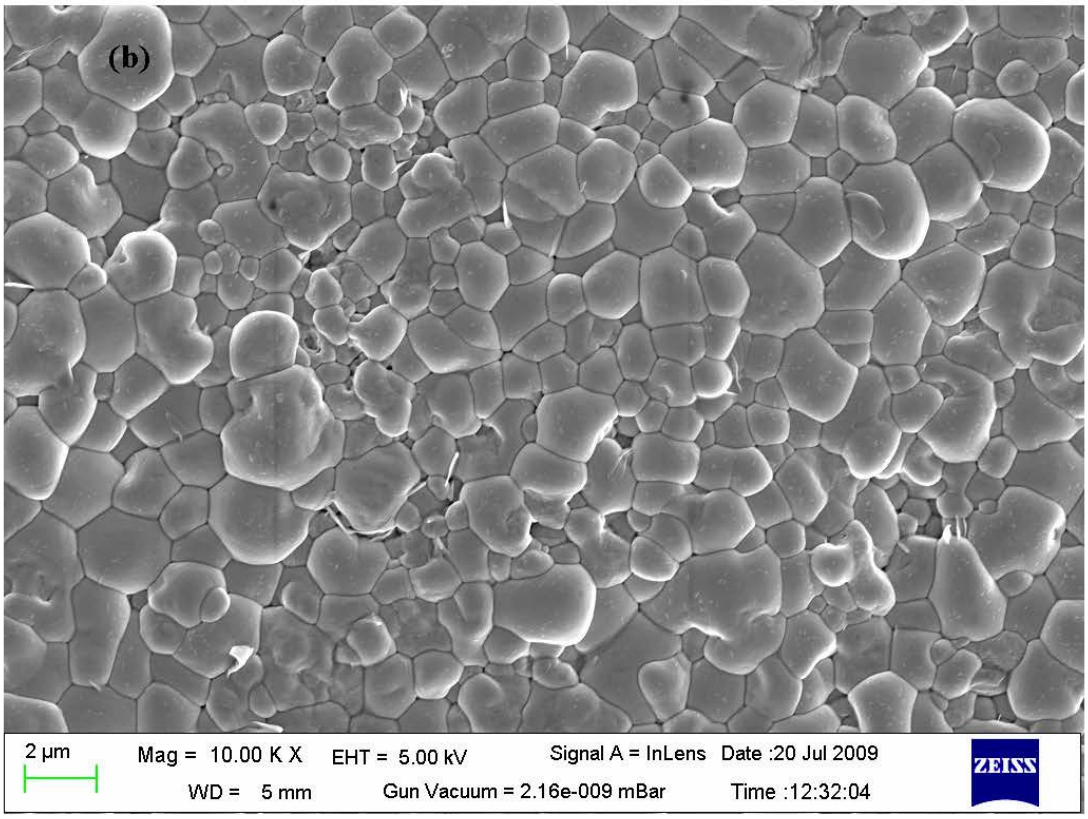
Figure 6. Frequency dependence of (a) real part of modulus *(M')* and (b) imaginary part of modulus *(M'')*, at different temperatures

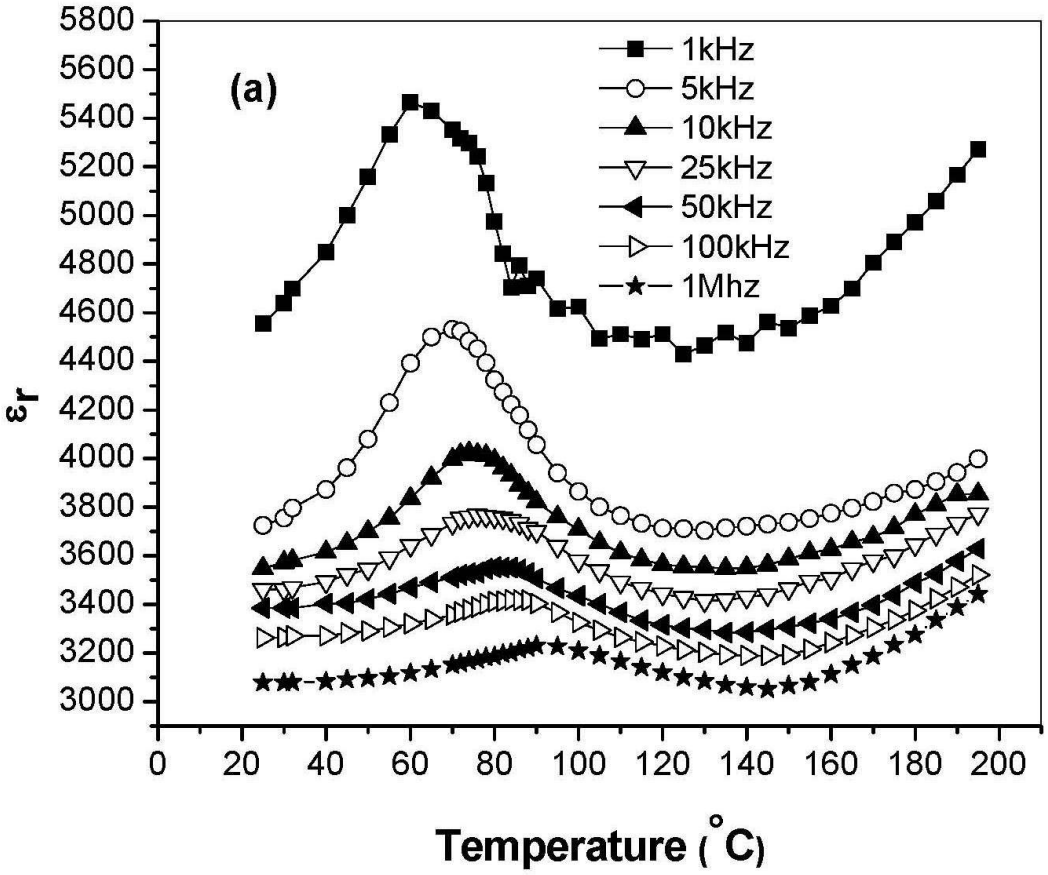
Figure 7. (a) Complex *M'* vs. *M''* plots (b) Scaling behavior of *M''* at various temperatures

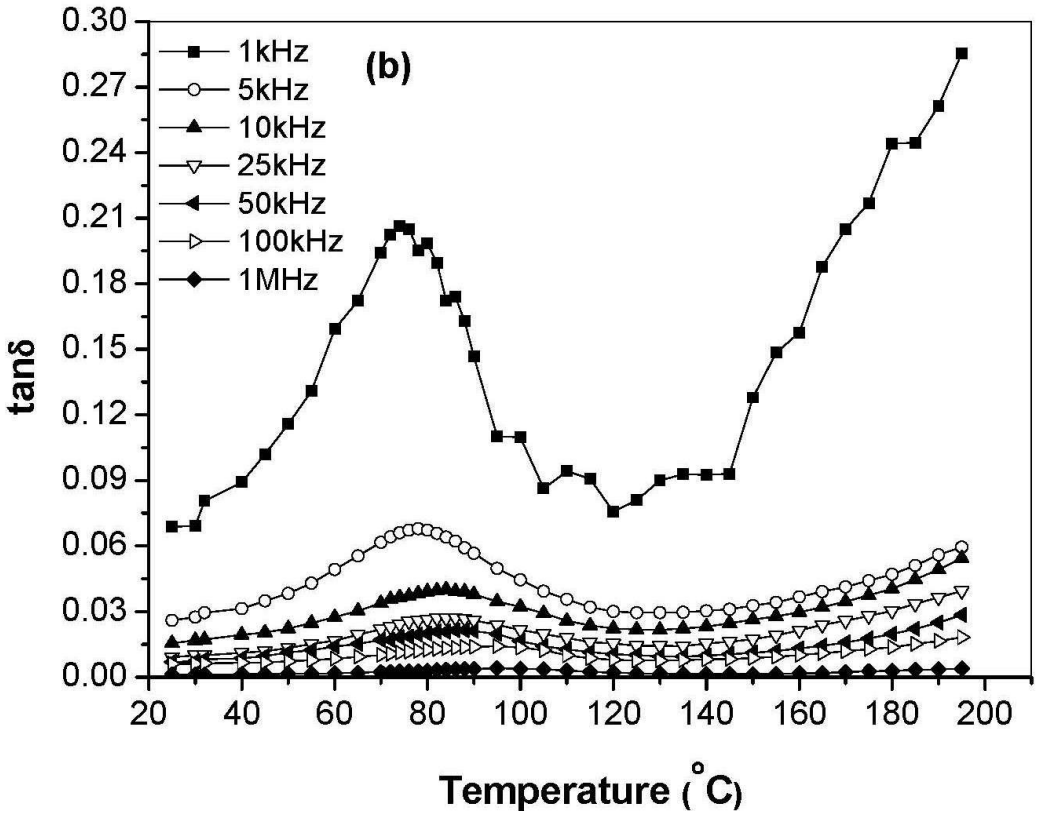
Figure 8. Frequency dependence of ac conductivity at different temperatures

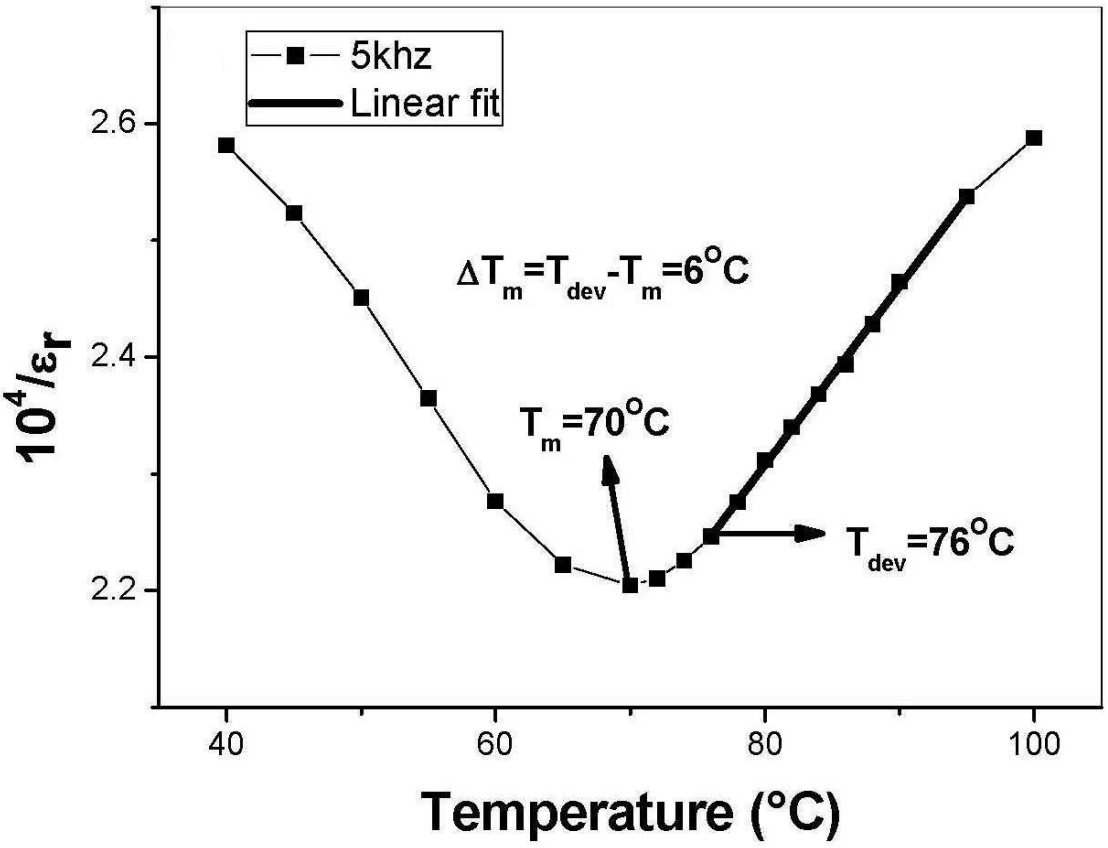
Figure 9. Hysteresis loops variation (a) with field E_0 at fixed frequency $f = 20\ Hz$ (b) with frequency at fixed field E=2500 V/cm.

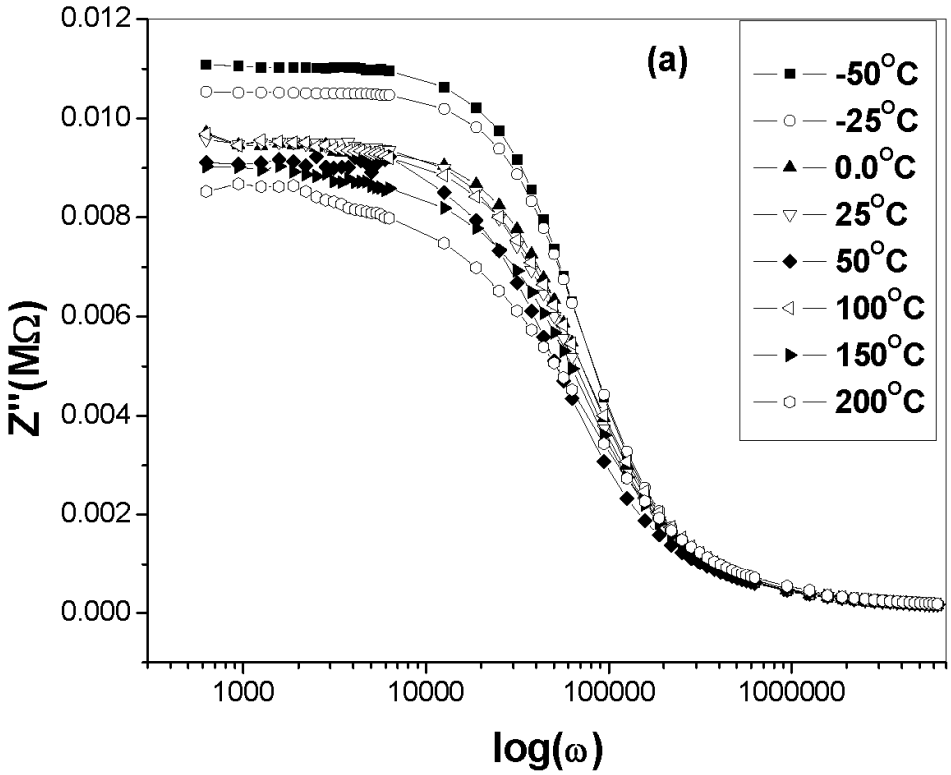


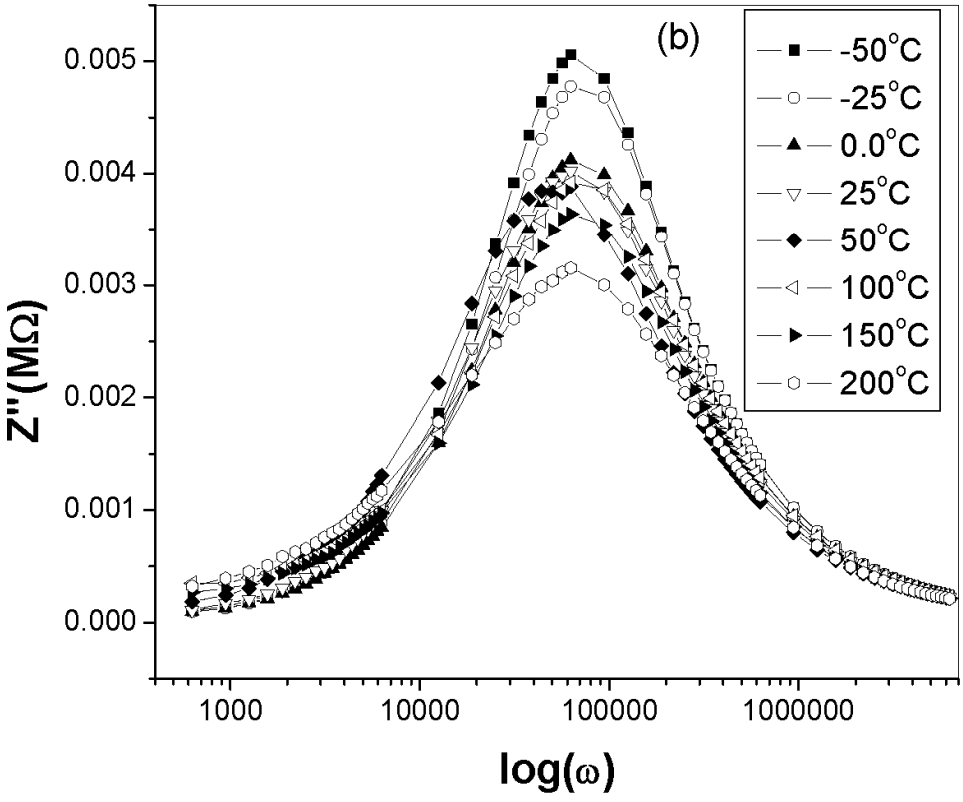


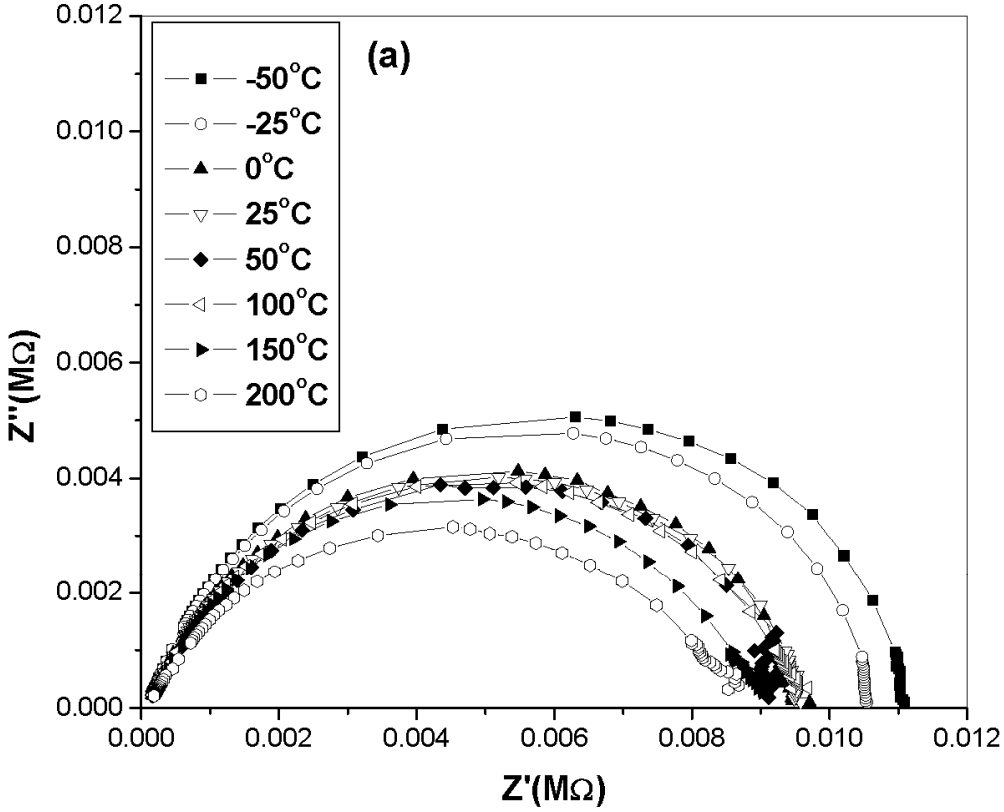


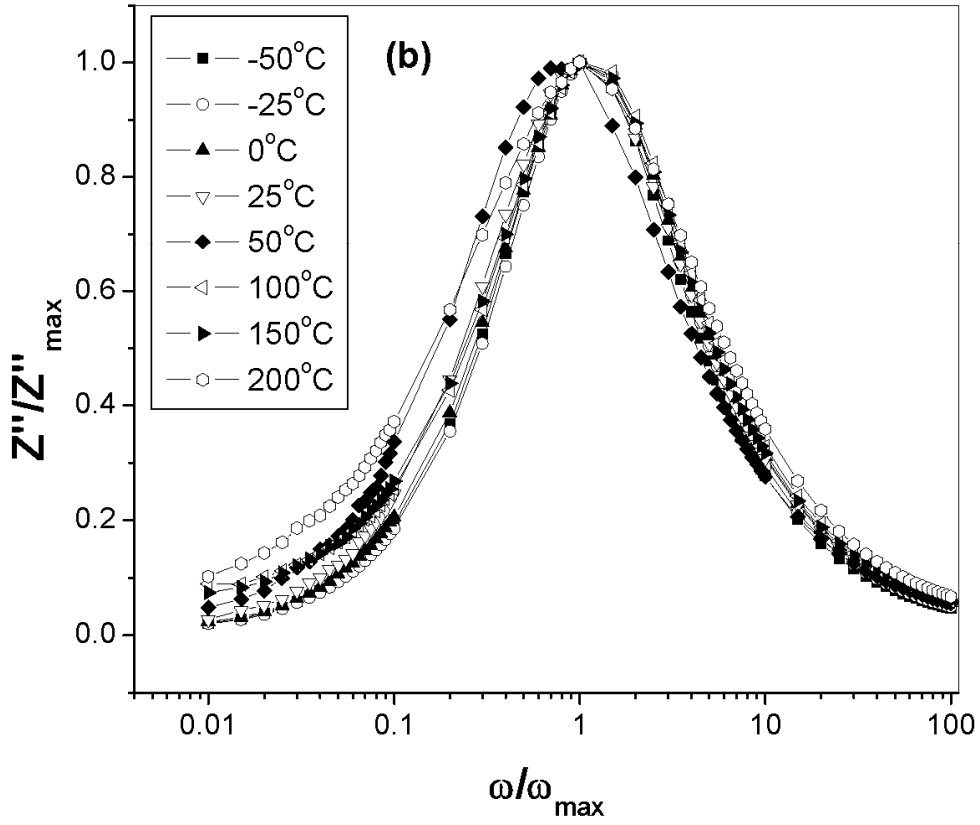


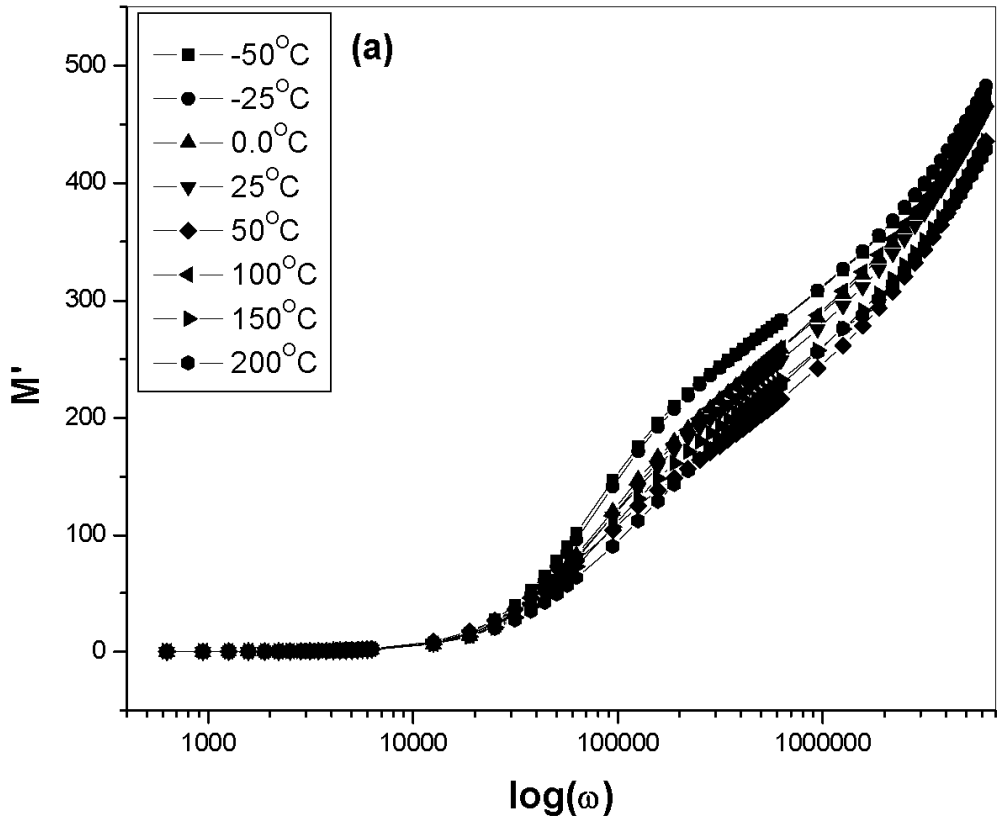


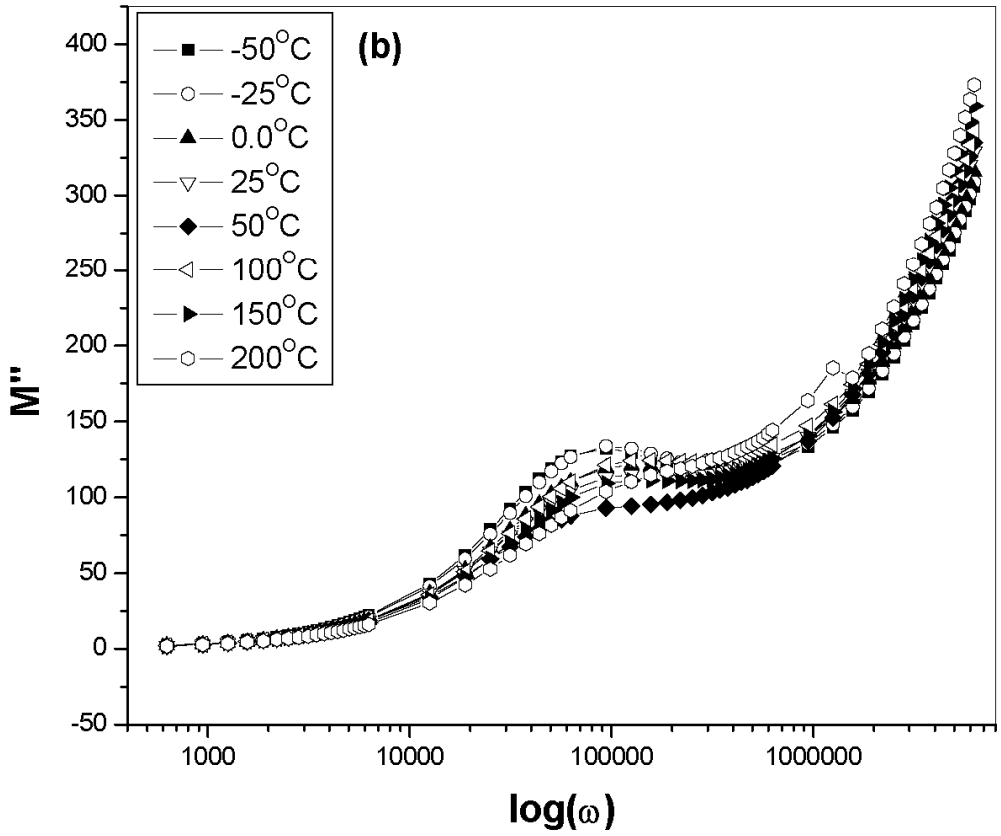


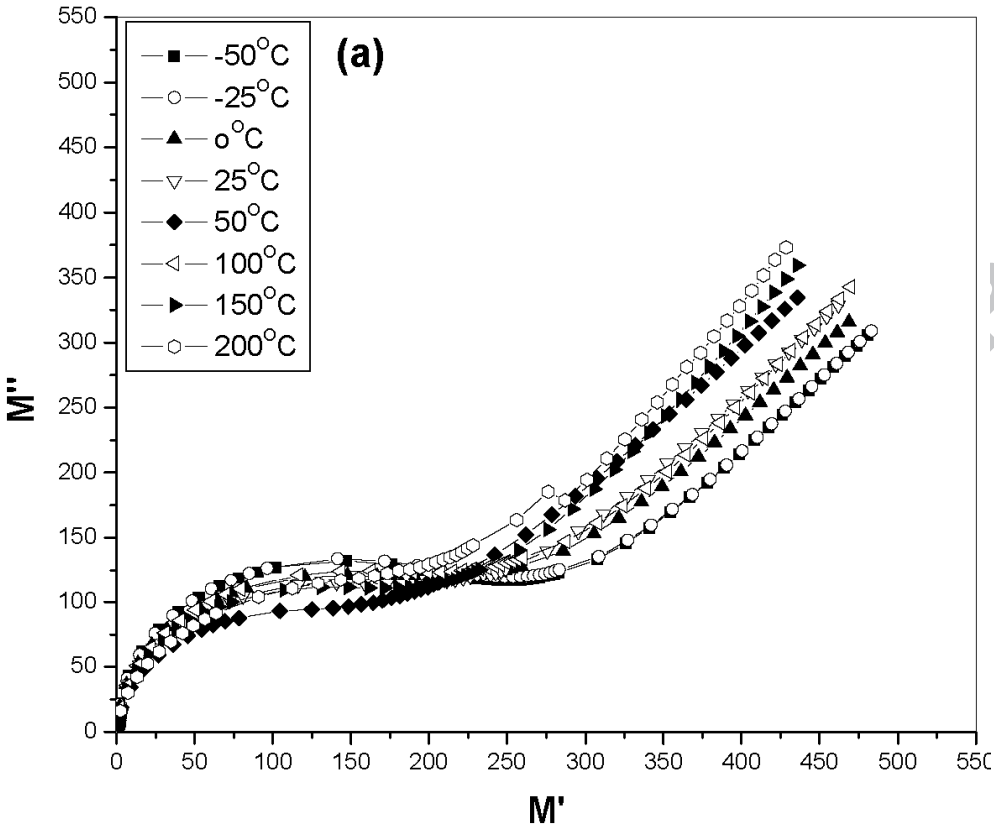


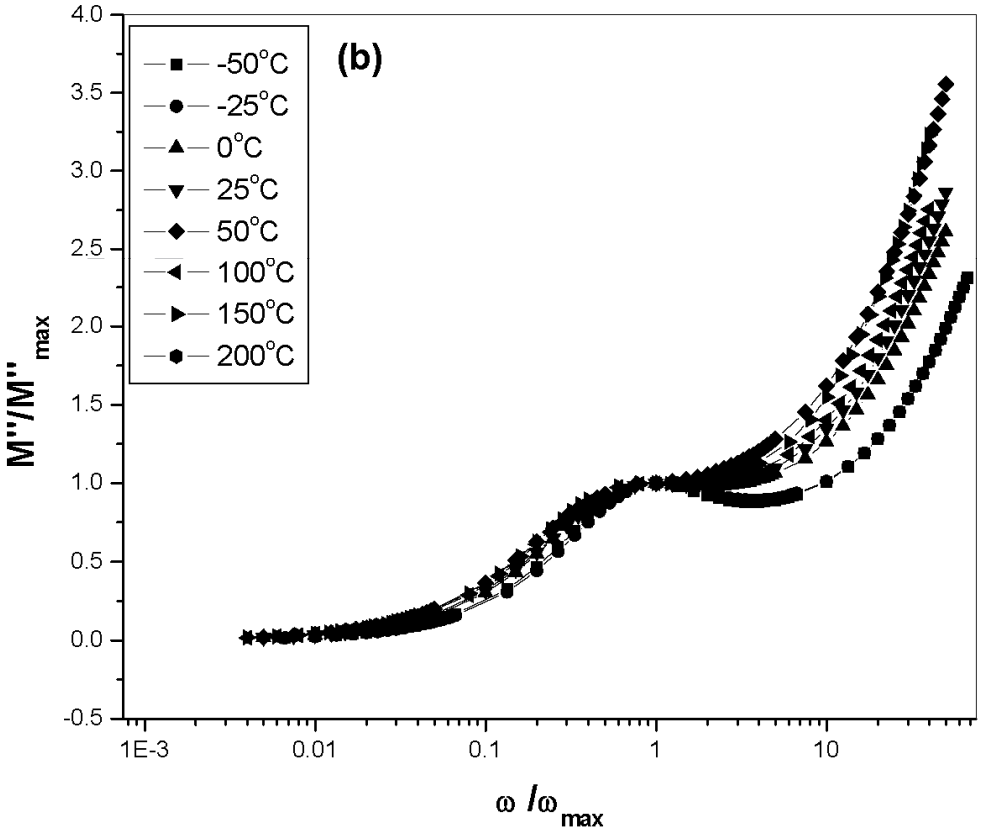


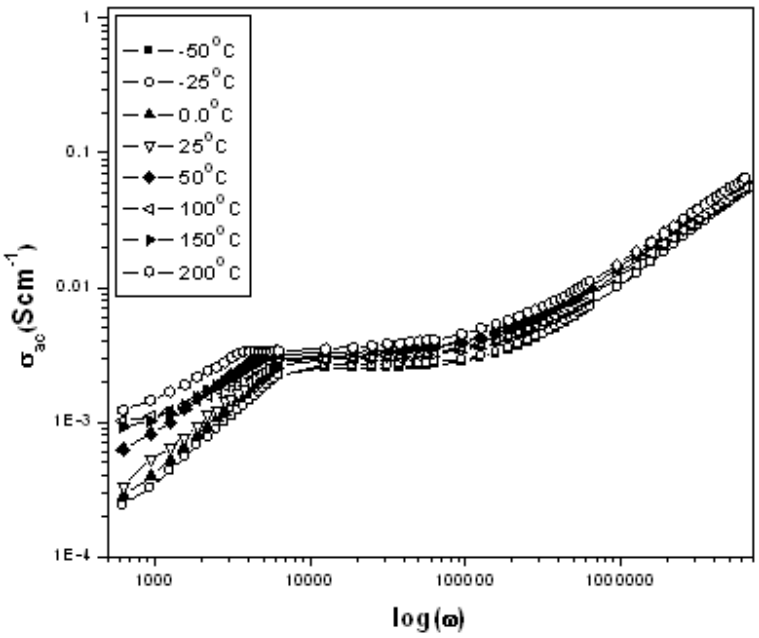


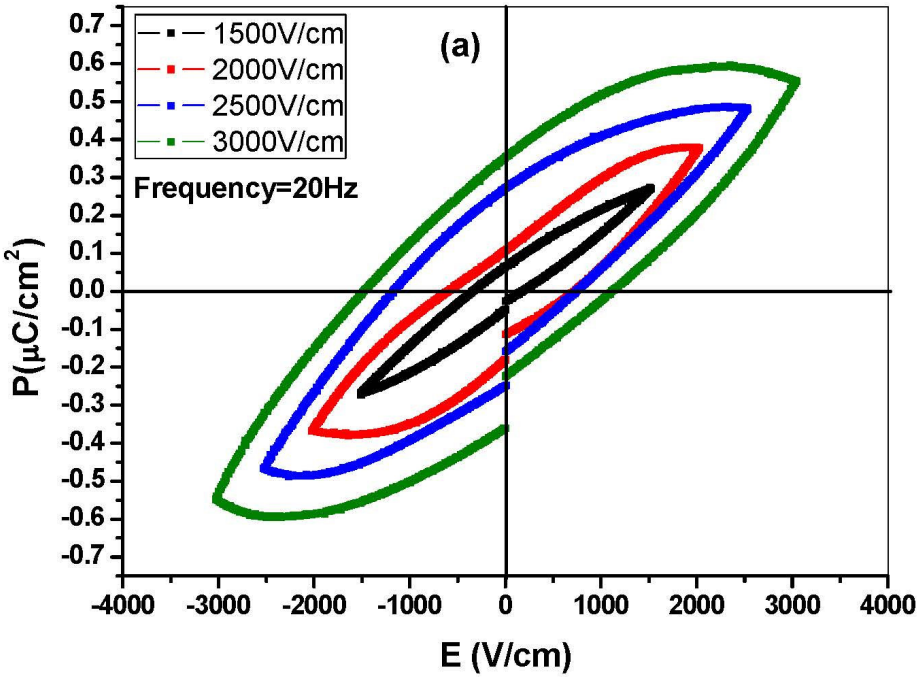












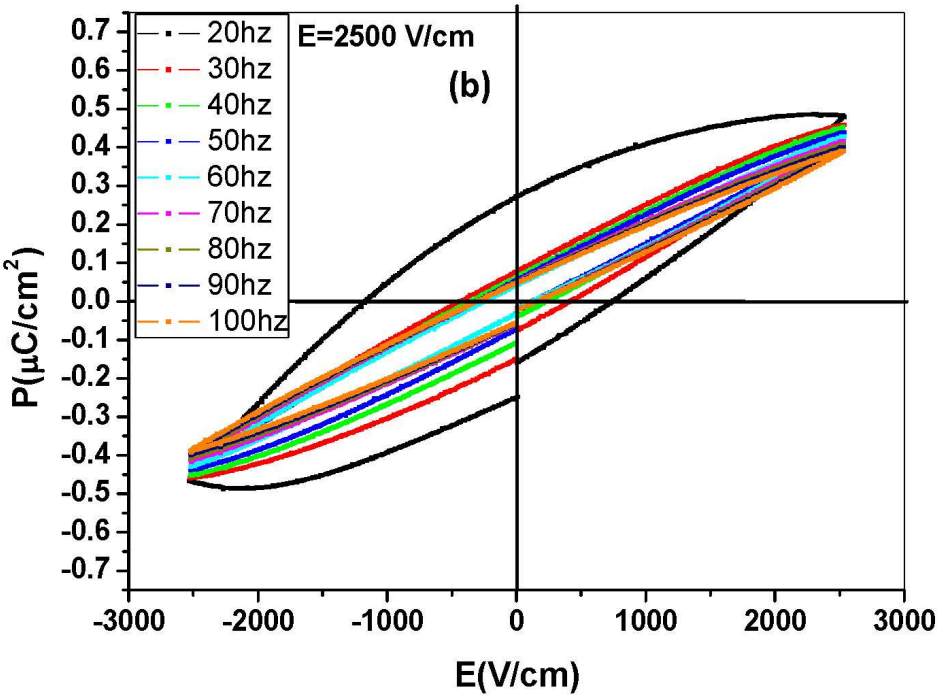


Table1. Hopping frequency (ω_p) and Jonscher's power (n) calculated from ac conductivity plot, at various temperatures.

Temperature(°C)	Hopping frequency(ω_p)	Jonscher's power (n)
-50	3768	0.78
-25	5338	0.83
0.0	5652	0.84
25	5024	0.77
50	5966	0.74
100	6280	0.78
150	6280	0.74
200	6280	0.78

Table.2 Calculated values of spontaneous (P_s) and remnant (P_r) polarization at fixed frequency with various field and fixed field with various frequency.

Fixed frequency (20Hz)			Fixed Field=2.5kV/cm		
Field (E_0 in V/cm)	P_{max} ($\mu C/cm^2$)	P_r ($\mu C/cm^2$)	Freq. (f in Hz)	P_{max} ($\mu C/cm^2$)	P_r ($\mu C/cm^2$)
1500	0.272	0.067	20	0.487	0.272
2000	0.380	0.107	30	0.459	0.081
2500	0.487	0.272	40	0.449	0.063
3000	0.595	0.354	50	0.438	0.055
			60	0.428	0.051
			70	0.416	0.049
			80	0.408	0.048
			90	0.400	0.047
			100	0.391	0.046

Highlights

- $Pb(Cd_{1/3}Nb_{2/3})O_3$ is very new and still has not been reported.
- Characterized by XRD, FESEM, temperature dependence of dielectric study, impedance spectroscopy and hysteresis loop study.
- Temperature dependence of relative dielectric constant exhibits diffuse phase transition due to oxygen vacancies and charge carriers in the compounds and the value of dielectric permittivity is quite high ~4532 at frequency 5 kHz.
- Frequency dependence of the real (Z') and imaginary (Z'') part of the electric impedance spectroscopy revealed that material has grain (bulk) resistance contribution.