11-13-2011

Proceedings of International Conference on Communication Network and Security

Dr. Srikanta Patnaik Chairman

Proceedings

Of

International Conference
on
Communication Network and Security

(ICCNS-2011)

13th-14th November, 2011

Editor-in-Chief

Prof. (Dr.) Srikanta Patnaik
President IRNet and Chairman, I.I.M.T., Bhubaneswar
Intersceince Campus,
At/Po.: Kantabada, Via-Janla, Dist-Khurda
Bhubaneswar, Pin:752024. Orissa, INDIA.

Organized By

IRNet
INTERSCIENCE RESEARCH NETWORK

About-ICCNS

One of the most dramatic technological developments in the era of information technology is the deployment of communication networks. This on-going revolution has raised and continues to raise fundamental challenges in the fields of science, engineering and industrial technologies. It requires new solutions, formulations and techniques for scientists and engineers from the communities of systems engineering and communications. The realization of wireless connectivity is bringing fundamental changes to telecommunications and computing and profoundly affects the way we compute, communicate, and interact. It provides fully distributed and ubiquitous mobile computing and communications, thus bringing an end to the tyranny of geography.

This conference shall provide a new forum for dissemination of knowledge on both theoretical and applied research on computer communication, network and security with an ultimate aim to bridge the gap between these coherent disciplines of knowledge. This forum accelerates interaction between the above bodies of knowledge, and fosters a unified development in the next generation communication.

Subject Coverage

The conference invites paper from all areas under the broad spectrum of Communication Network, Security and their applications. The areas are as follows:

- Hardware support, System architectures, Services and system support
- Algorithm/protocol design and analysis
- Mobile environments
- Applications
- Wireless communications and networks
- Autonomous, robotic and intelligent systems and control
- Biological/biomedical systems and control
- Internet and web based systems
- Software systems and communication systems
- Artificial intelligence and intelligent information systems
- Distributed and cooperative control systems
- Mechatronic and micromechatronic systems
- Quantum systems and nano-systems
- Economic, environmental and social system engineering
- Modelling, identification, analysis and control of time delay systems
- Networked control systems (NCS)
- Network control, e.g., admission/flow/congestion control
- Power control and mobility management of wireless networks
- Network scheduling and bandwidth allocation
- Quality of service (QoS) and quality of performance (QoP)
- Protocol-based feedback control and control-oriented communication protocols

- Informatics in control and communication
- Design and analysis of quantisers and coder/decoders
- Control with partial/intermittent/delayed information feedback
- Distributed control under power, rate and distortion constraints
- Integrated control, computing and communication systems (ICCCS)
- Applications of game theory and Markov decision theory to ICCCS
- Computationally efficient control and communication algorithms
- Fault detection, diagnostics and prognostics
- Sensors and actuators networks
- Distributed consensus, agreement, and optimisation
- Stochastic dynamic games and teams
- Evolving complex network models and design algorithms
- New features and theoretic analysis of complex network systems
- Synchronisation and control of complex dynamical networks
- Emergent behaviours and patterns on complex networks
- Complex networks and social/economic/biological systems
- Complex dynamic networks and multi-agent systems
- Case studies and application examples of any related topics
- Cellular planning for GSM, UMTS, WCDMA, cmda2000, BFWA, PMR, wifi and 802.11
- Transmission infrastructure and site location
- Channel assignment and modelling
- Frequency hop set design
- Netted radar placement
- Spectrum licensing and auctions and International spectrum management
- Spectrum trading, Revenue generation and Subscriber service pricing policy
- Application of codes
- Radio resource management and allocation
- Bluetooth scatternet formation
- Point to point link topology, propagation modelling and ray tracing
- Key performance indicators, surrogates and quality of service measurement
- Traffic modeling, Routing
- Subscriber location issues
- Billing
- Effective source code design and Software system design
- Requirements analysis and Fast data structures
- Object oriented design methodologies and Human-computer interaction
- User interface development
- High performance computing, Parallel processing
- Geographical information
- Graphics, Visualisation, simulation environments and tool kits
- Heuristics and meta-heuristic algorithm design and application

- Genetic algorithms
- Tabu search, Simulated annealing, Neighbourhood searching
- Constraint programming
- Mathematical modelling and mathematical programming formulations
- Randomisation
- Expert systems, Agents, Artificial intelligence, Neural networks and Ambient intelligence
- Emergent approaches
- Integrated mobile marketing communications
- Telematics
- Wireless advertising/wireless CRM
- Pervasive computing technologies
- Incoming and outgoing wireless links
- Efficacy of mobile communications
- Teaching mobile communication applications
- Critical success factors for mobile communication diffusion
- Metric mobile business enterprise
- Mobile communication security issues and requirements
- PDAs in health services delivery
- Interaction and integration in mobile communications
- Location management for mobile communications
- Business models for mobile communications
- Groupware, Roomware
- Mobile ad hoc networking and Nomadic communication
- Portable communications and Cross-cultural mobile communication issues
- Wireless information assurance
- Mobile and handheld devices in the classroom and Tele-learning
- Security in cellular networks (2G, 2.5G, 3G, B3G, 4G, etc.)
- Security in wireless LANs (IEEE 802.11 WLAN, WiFi, and HiperLAN/2)
- Security in wireless PANs (Bluetooth and IEEE 802.15)
- Security in wireless MANs (IEEE 802.16 and WiMAX)
- Security in sensor and ad hoc networks
- Security in mobile IP and wireless internet
- Security in integrated wireless networks and satellite networks
- Security in wired and wireless integrated networks
- Security in wireless communications and IP networks
- Security in internet and WWW
- Security in high-speed networks and peer-to-peer networks
- Security in optical systems and networks
- Security in VoIP and e-mail
- Security in domain name service
- Security in integrated networks and content-delivery networks

- Security in and communications distributed systems
- Attacks, security mechanisms, and security services
- Access control, Authentication and Authorisation
- Multicast security
- Distributed access control
- Data integrity and Data confidentiality
- Non-repudiation
- Firewall and Privacy protection
- Security specification techniques, Encryption and decryption, and Secure routing protocols
- Formal analyses and Security group communications
- Intrusion detection
- Key management, Trust establishment and Revocation of malicious parties
- Security policies, Fraudulent usage and Dependability and reliability
- Anonymity, Prevention of traffic analysis
- Secure PHY/MAC/routing protocols
- Secure location determination and Denial of service
- Network security performance evaluation
- Tradeoff analysis between performance and security
- Network forensics
- Design or analysis of security protocols and Security standards
- Energy efficiency (sleep mode, etc.)
- Applications
- Location techniques
- Routing, Medium access control (MAC)
- Coverage, Connectivity and Longevity
- Scheduling, Synchronisation and Network resource management
- Energy efficient protocols (PHY, MAC, routing, application)
- Lightweight protocols
- Fault tolerance and diagnostics
- Foundations
- Data storage and query processing
- In-network processing and aggregation
- Learning of models from data
- Mobility
- Performance analysis
- Sensor tasking and control
- Security, privacy, and data integrity
- Modelling of systems and physical environments
- Network protocols
- Simulation tools and environments
- System architectures and operating systems

EDITORIAL BOARD

Editor-in-Chief

Members of Editorial Board

Organizing Committee

Program Cochair
Prof. (Dr.) Srikanta Patnaik
President IRNet and
Chairman, I.I.M.T., Bhubaneswar
Intersceince Campus,
At/Po.: Kantabada, Via-Janla, Dist-Khurda
Bhubaneswar, Pin:752024. Orissa, INDIA

Dr. Michael R. Bartolacci
Associate Professor of Information Sciences and Technology
The Pennsylvania State University - Berks
Tulpehocken Road
P.O. Box 7009
Reading PA 19610-6009
USA
mrb24@psu.edu

Organising Committee:

Secretary IRNet
Prof. Pradeep Kumar Mallick
IIMT, Bhubaneswar
Mobile No: 08895885152

Conference Coordinator :
Mr. Bibhu Prasad Mohanty
IRNet , Bhubaneswar,India

Publication
Prof. Sushanta Kumar Panigrahi
IIMT, Bhubaneswar

Prof. Mritunjay Sharma
IIMT, Bhubaneswar
Mobile No- +91 9338499777

Post Conference Coordinator
Ujjayinee Swain
IRNet, Bhubanesawr

Head (System & Utilities)
Prof. Sanjay Sharma
IIMT, Bhubaneswar

Event Manager
Prof. Sharada Prasad Sahoo.
IIMt, Bhubaneswar

Members of IRNet:
Dr. Sukumar Mishra, Assistant Professor
Electrical Engineering Department
Indian Institutes of Technology (IIT)

Dr. Bijaya Ketan Panigrahi, Assistant Professor
Electrical Engineering Department
Indian Institutes of Technology (IIT), India

Prof. Anupam Shukla, Professor
Department of ICT
ABV-Indian Institute of Information Technology and Management
Gwalior, Madhya Pradesh

Prof. Jaya Sil,
Professor and Head Computer Science & Technology
Bengal Engineering And Science University, Shibpur
P.O. - Botanic Garden, Howrah, West Bengal.

Dr. Amit Saxena, Professor
Department of Computer Sc. and Information Technology
G G University, Bilaspur
Bilaspur PIN- 495009 C G State

Dr. T.V.Vijay Kumar
School of Computer and Systems Sciences
JNU, New Delhi Conference Secretary

Prof. Chandrasekhar Panda,
Sambalpur University, Bhubaneswar

Prof. B. K. Patnaik, ITER,
SOA University, Bhubaneswar

Mr. Rashmi Ranjan Nath
IRNet, Bhubaneswar, India

Sagarika Ray
IRNet, Bhubaneswar, India

Debashree Rath
IRNet, Bhubaneswar, India

First Impression : 2011

(c) Interscience Research Network

Proceedings of International Conference on Communication Network and Security

DISCLAIMER

The authors are solely responsible for the contents of the papers complied in this volume. The publishers or editors do not take any responsibility for the same in any manner. Errors, if any, are purely unintentional and readers are requested to communicate such errors to the editors or publishers to avoid discrepancies in future.

ISBN : 978-93-81693-02-05

Published by :

IPM PVT. LTD., Interscience Campus
At/PO.: Kantabada, Via: Janla, Dist: Khurda, Pin- 752054
Publisher's Website : www.interscience.in
E-mail: ipm.bbsr@gmail.com

Typeset & Printed by :

IPM PVT. LTD.

TABLE OF CONTENTS

Sl. No.	Topic	Page No.
	Editorial - *Prof. (Dr.) Srikanta Patnaik*	
1	**Efficient Rule Set Generation using Rough Set Theory for Classification of High Dimensional Data** - *Prasanta Gogoi, Ranjan Das, B Borah & D K Bhattacharyya*	**1 - 6**
2	**Secured Data Access in Cloud Computing** - *B. Krishna Prasad*	**7 - 12**
3	**Blind And Robust Audio Watermarking in Wavelet Domain** - *Suresh Penchala & K. Sivani*	**13 - 16**
4	**An Improved Protocol for Proxy Based Certification Authority for MANETs** - *Bibhudendu Panda & Pabitra Mohan Khilar*	**17 - 21**
5	**Development of Empirical Mode Decomposition** - *R. Samba Siva Nayak, Satish Kumar Mandava & K. Suresh Babu*	**22 - 24**
6	**Evaluation of Probabilistic Broadcasting Method in Mobile Ad Hoc Network in Different Scenario** - *Sukant Kishoro Bisoyi, Mohit Ranjan Panda & Manas Kumar Swain*	**25 - 30**

Editorial

In the recent years communication systems have acquired advanced technical mechanism which immensely contributes to the efficiency and accuracy. On the wake of information economy it is observed that communication frameworks the governance, policy and strategic behavior of the social and commercial institutions. Hence there is an urge towards information security and protection of communication ethics. As the usage of information systems have become an indispensible component of any purposeful organization, there is warfare among competitors may be a nation or a corporate. To overcome these complexities Security has been a prominent research interest among the computer professionals. As hacking and tracking has gained momentum in all spheres encompassing from a debit card database to an examination database, effective control mechanism is knocking the door of research. I must acknowledge the step taken by OECD in its communications outlook 2011 which clearly emphasizes on Communication infrastructure and it's pivotal role for competitiveness.

The advanced studies such as Compressed video Communications, Content Production technology, Computer Security and Cryptography, Distributed System Securities, Digital Audio Broadcasting have broadened the scope of information transfer and exchange. Simultaneously on the security domain the study has been widely expanded to cryptosecurity, transmission security, emission security, traffic-flow security. Even researchers are trying to manifest Signal Intelligence, NSA encryption system, Type 1,2,3,3 product to develop better security system. Secure voice over internet protocol (SVOIP) has become the defacto standard for securing voice communication, replacing the need for STU-X and STE equipment in much of the U.S. Department of Defense. USCENTCOM moved entirely to SVOIP in 2008.

The conference is designed to stimulate the young minds including Research Scholars, Academicians, and Practitioners to contribute their ideas, thoughts and nobility in these two integrated disciplines. Even a fraction of active participation deeply influences the magnanimity of this international event. I must acknowledge your response to this conference. I ought to convey that this conference is only a little step towards knowledge, network and relationship.

I congratulate the participants for getting selected at this conference. I extend heart full thanks to members of faculty from different institutions, research scholars, delegates, IRNet Family members, members of the technical and organizing committee. Above all I note the salutation towards the almighty.

Editor-in-Chief

Prof. (Dr.) Srikanta Patnaik
President IRNet and
Chairman, I.I.M.T., Bhubaneswar

Efficient Rule Set Generation using Rough Set Theory for Classification of High Dimensional Data

Prasanta Gogoi, Ranjan Das, B Borah & D K Bhattacharyya

Department of Computer Science and Engineering, Tezpur University, Napaam, Tezpur, India 784028
E-mail: {prasant, ranjan, bgb, dkb }@tezu.ernet.in

***Abstract* -** In this paper, a rough set theory (RST) based approach is proposed to mine concise rules from inconsistent data. The approach deals with inconsistent data. At first, it computes the lower and upper approximation for each concept, then adopts a learning from an algorithm to build concise classification rules for each concept satisfying the given classification accuracy. Lower and upper approximation estimation is designed for the implementation, which substantially reduce the computational complexity of the algorithm. UCI ML Repository datasets are used to test and validate the proposed approach. We have also used our approach on network intrusion dataset captured using our local network from network flow. The results show that our approach produces effective and minimal rules and provide satisfactory accuracy over several real life datasets.

Keywords- *Rough Set; Inconsistency; Minimal; Redundant; Intrusion Data; PSC*

I. INTRODUCTION

The rules are the prescribed standards on the basis of which decisions are made for specific purpose. The rule is a statement that establishes a principle or standard, and serves as a norm for guiding or mandating action or conduct. The rule can be a conditional statement that tells the system how to react to a particular situation. In data mining, the rule generation was first introduced by Agrawal et. al. in 1993 in terms of Market-basket analysis [1] as association rules. In association rules, the rule generation is based on the concept of frequent pattern mining for the discovery of interesting associations and correlations among itemset. Afterwards, methods were developed for classification rule mining [2]. The rule-based methods can be found in different applications of decision making and prediction like in the domain of medical research [4], in the areas of economics and finance [5] and in network security [6].

The cost of developing and maintaining rule sets is an important issue for the rule based systems. Based on the literature survey, it has been observed that three types of rule generation techniques are commonly found, viz., frequent association rule mining [7], rare association rule mining [8], and multi-objective rule mining [9].need to create these components, incorporating the applicable criteria that follow.

We observe that the rules generated by the above three approaches often are incapable of

- handling inconsistency in the database
- generating minimal rule set
- generating non-redundant rule set

In different real life or synthetic dataset, inconsistency is a common problem. Inconsistency is caused by the existence of indiscernbility relation in decision table. A data set is represented as a table, where each row represents an object or record. Every column represents an attribute that can be measured for each object. This table is a decision table. Attributes are of two categories: condition and decision. The indiscernbility relation occurs in a decision table if in objects of equivalent condition attributes, decision attributes are different. Consider a decision table with objects p_1, p_2, p_3 in *Table I*. Condition attributes are *A* and *C*, and decision attribute is *D*. The attributes of objects p_1 and p_3 are equivalent whereas their decision attributes are different. Here, objects p_1 and p_3 are indiscernible and the decision table has inconsistency. None of the previously mentioned techniques can provide any means to generate classification rules in these situations. In view of the above mentioned limitations, rough set theory (RST) is introduced for classification rule generation on inconsistent dataset. RST was first introduced by Pawlak [11] in the year 1982. RST is especially well suited to deal with inconsistencies [10]. One of the major advantages of RST is that it does not require any additional information on the data such as probability distribution or grade membership and it is capable of handling inconsistency.

TABLE I. INCONSISTENT DATASET

Objects	**Condition Attributes**		**Decision Attributes**
	A	**C**	**D**
P_1	*low*	*high*	*yes*
P_2	*low*	*low*	*no*
P_3	*low*	*high*	*no*

Followings are our contributions in this paper

- A method is proposed to find indiscernibility relation in data set to find inconsistency
- Determination of lower and upper approximation for inconsistent data.
- Minimized and non-redundant rule generation by using lower and upper approximation for classification.

The remainder of this paper is organized as follows. In the next section, we present related work on rule generation. In Section 3, we have given the proposed method of rule generation. Experimental results are presented in Section 4. In Section 5, we outline the conclusion and future work.

II. RELATED WORKS

A. *Preliminaries of Rough Set*

Rough set was proposed to classify imprecise and incomplete information. There have been contributions on applying rough sets theory (RST) to rule discovery. In [19], RST was used on clusters to determine rules for association explanations. Adetunmbi [20] used rough sets to data that contain the minimal subset of attributes associated with a class label for classification. RST can help to determine whether there is redundant information in the data to gather the essential data needed for applications. The RST based rule generation approach can be able to generate minimal and non-redundant rule set in inconsistent data.

B. *Rough Set*

RST is an approach to vagueness. It is an extension of the classical set theory, for use when representing vagueness i.e. imprecision. Rough set is expressed by a boundary region of a set [11]. The basic concept of the RST is the notion of approximation space, which is an ordered pair *I = (U, R)*, where

- *I*: information system
- *U*: nonempty set of objects, called universe
- *R:* equivalence relation on *U*, called indiscernibility relation. If $x, y \in U$ and *xRy* then *x* and *y* are indistinguishable in *I.*

Each equivalence class induced by *R*, is called an elementary set in *A* and represented as *U/R*. A definable set in *I* is any finite union of elementary sets in *I*. For $x \in U$, let $[X]_R$ denote the equivalence class of *R*, containing *x*. For each $X \subseteq U$, *X* is characterized in *I* by a pair of sets- its lower and upper approximation in *I*, defined respectively as:

$$\underline{R}X = \{x \in U \mid [X]_R \subseteq X\} \qquad (1)$$
$$\overline{R}X = \{x \in U \mid [X]_R \cap X \neq \varnothing\}$$

A rough set in *I* is all subsets of *U* having the same lower and upper approximations. Reduct and core are two related concepts in RST.

Reduct: A reduct [11] is a set of attributes that preserves partition. It means that a reduct is the minimal subset of attributes that enables the same classification of elements of the universe as the whole set of attributes.

In order to express the idea of reduct, let $B \subseteq A$ and $a \in B$ in an information system *I = (U, A)* where *U* is the universe of objects, A is set of attributes, and *R(B)* is a binary relation.

- *a* is dispensable in *B* if *R(B) = R(B −{a})*; otherwise *a* is indispensable in *B*
- Set *B* is independent if all its attributes are indispensable.
- $B' \subseteq B$ is a reduct of *B* if B' is independent and $R(B') = R(B)$

The attributes other than the reduct are redundant attributes. The removal of redundant attributes cannot deteriorate the classification. Usually, there are several reducts in a dataset.

Core: The core [11] is the set of all indispensable attributes, i.e., it is the intersection of all reducts.

Let *Red(B)* is the set of all reducts of *B* in an information system *I = (U, A)* where $B \subseteq A$ then the core of *B* is defined as

$$core(B) = \bigcap \mathrm{Re}\,d(B) \qquad (2)$$

The core is included in every reduct, i.e., each element of the core belongs to some reduct. Thus, the core is the most important subset of attributes, for none of its elements can be removed without affecting the classification.

C. HCRI Algorithm

It is a heuristic algorithm for mining concise rules from inconsistent data (HCRI [3]). This method is based on the variable precision rough set model. It deals with inconsistent data to mine concise rules. It first computes the reduct for each concept, then computes the reduct for each object. It adopts a heuristic method to build concise classification rules for each concept. To compute the equivalence classes, it uses two hash functions, which substantially reduce the complexity to *O(n)*, $n = |U|$. The hash functions compute the cardinality of lower approximation. The input to the method is a set of inconsistent objects *U* and the output is a set of concise rules satisfying a given classification accuracy.

D. LEM2 Algorithm

LEM2 [21] (Learning by Example Module, Version - 2) is a machine learning algorithm based on rough set theory. The usual task of LEM2 algorithm is to learn a discriminate rule set, i.e., to learn the smallest set of minimal rules, describing the concept. This algorithm can generate both certain and possible rules from a decision table with attributes being numerical as well categorical. LEM2 needs discretization for numerical attributes.

For inconsistent data, LEM2 induces two sets of rules: certain rule set and possible rule set. The first set is computed from lower approximations of concepts, the second one from upper approximations. It is assumed that the rule set will be used automatically by a classification component. Nevertheless, induced rules are available and comprehensible by the user. Thus, it is possible to use rules manually, like in other systems.

The LEM2 algorithm is a single local covering approach. It yields single minimal discriminate description, which means, learning the smallest set of minimal rules for every concept. The local coverings are constructed from minimal complex. A formal definition of minimal complex and local covering is reported next from [11].

Definition: Minimal complex and Local covering Let *B* be a nonempty lower or upper approximation of a concept represented by a decision-value pair *(d,w)*. The set *T* is a minimal complex of *B* if and only if *B* depends on *T* and no proper subset T' of *T* exists such that *B* depends on T'. Let $\Im$ be a collection of non-empty set of attribute-value pairs for equivalence class *[T]* of *T*. Then $\Im$ is the local covering of *B* iff

- each member *T* of $\Im$ is a minimal complex of *B*
- $\bigcup_{T\in\Im}[T] = B$ and
- $\Im$ is minimal i.e., $\Im$ has the smallest possible number of members.

LEM2 algorithm is found suitable in rule generation for inconsistent data. In the next section, we propose a LEM2 based technique for rule generation.

III. PROPOSED WORK

We have proposed an effective rule generation technique using RST based on LEM2 algorithm. The proposed method can be found to be significant especially for those datasets having inconsistencies. Our method starts with the inconsistency checking for each concept in the dataset. If it finds the inconsistency, then computes the upper approximation and the lower approximation. To compute the inconsistency and to find the upper and lower approximations it introduces the following method to support the LEM2 based rule generation technique.

A. Evaluation of Upper & Lower Approximation

For computation of the upper and lower approximation for each concept of dataset, it executes the steps given below.

Algorithm: *Compute CLU*

1. ***Identify the set of concepts*** Γ.
2. ***Take an arbitrary object*** **c** ***from a concept*** $C_i \in \Gamma$ ***and make a comparison of each attribute-value pair with all the objects*** c' ***of another concept*** $C'_j \in \Gamma$. If attribute-value pairs of object *c* and object c' are matching, then inconsistency occurs with respect to the concepts C_i and C'_j .
    a. Search any other inconsistent pair of objects of the concepts (C_i , C'_j).
    b. Generate U_{approx} by taking union of the set of objects *{* C_i *}* of concept $C_i \in \Gamma$ with the inconsistent pairs of objects of concept (C_i , C'_j) i.e., $U_{approx} = \{C_i\} \bigcup \{C'_j\}$.
    c. Generate L_{approx} by subtracting C'_j from C_i i.e., $L_{approx} = \{C_i\} - \{C'_j\}$.
3. Otherwise objects are consistent.

B. *Proposed Method*

The proposed rule generation approach is based on LEM2 algorithm. LEM2 is a single local covering approach and it yields single minimal discriminate description. In LEM2, the user may or may not consider any attribute priority. However, in contrary to LEM2, the proposed rule generation approach considers

- the attribute priority, and
- extracts the output of the method to *Compute CLU*

Let β is upper or lower approximation of a concept or a concept itself and *B* is a members of β. $\Im$ is a single local covering for the set β, i.e., it yields the smallest set of minimum rules for the entire set β. *G* is a temporary storage of *B*. *T* is a set of attribute-value pairs. *t* is a member of *T*, i.e. $t \in T$. *[t]* is a equivalence class of *t*, i.e., the set of all objects which have the attribute-value pair *t*. *T(G)* is a set of attribute-value pairs which are present in objects of *G*, i.e., $T(G) := \{t \mid [t] \cap G \neq Null\}$. *[T −{t}]* is a set of objects which have the attribute-value pairs other than *t*. *S* is a member of $\Im$ other than *T*, i.e., $S \in \Im - \{T\}$.

<u>*Procedure*</u>

```
input : a set $\beta$
output: a single, local covering $\Im$ of set $\beta$;

begin {Procedure}
  while($\beta \neq$ Null )
  begin
    for each concept,
      if found inconsistency then
        $L_{approx}$ and $U_{approx}$ will be the member of $\beta$
      else
        the concept will be the member of $\beta$.
    for each $B \in \beta$ do
    begin
    $G:=B;$
    $\Im := Null;$
    while ($G \neq Null$)
    begin
      $T := Null;$
      $T(G) := \{t \mid [t] \cap G \neq Null\};$
      while ($T=Null$) or ($[T] \not\subseteq B$)
        begin
      Select an attribute-value pair $t \in T(G)$ with the highest attribute priority, if a tie occurs, select a $t \in T(G)$ such that $|t \cap G|$ is maximum;
      if another tie occurs, select a $t \in T(G)$ with the smallest cardinality of $[t]$; if further tie occurs, select the first pair;
      $T := T \cup \{t\};$
      $G := [t] \cap G;$
      $T(G) := \{t \mid [t] \cap G \neq Null\};$
      $T(G) := T(G) - T;$
        end{while}
        for each $t$ in $T$ do
        if $[T - \{t\}] \subseteq B$ then $T := T - \{t\};$
          $\Im := \Im \cup \{T\};$
          $G := B - \bigcup_{T \in \Im} [T];$
        end{while}
    end{while}
    for $T$ in $\Im$ do
    if $\bigcup_{S \in \Im - \{T\}} [S] = B$ then $\Im := \Im - \{T\};$
end{Procedure}
```

C. *Complexity Analysis*

Let *n* be the total number of objects in our sample dataset. Now, in order to verify the inconsistency, we have to compare each individual object with every other objects present in our dataset. So, the complexity of the computation of upper and lower approximation is $O(n^2)$. LEM2 has the complexity of *O(nm)* where *n* is the number of objects and *m* is the number of attributes. The complexity of our proposed algorithm is $O(n^2) + O(nm)$.

Our algorithm expects that the sample dataset to have inconsistency. The inconsistencies may arise in only some of the concepts (not all). So, our algorithm initially compute the U_{approx} and L_{approx} for those concepts only. The concepts which do not have inconsistency will fed to the program without finding upper and lower approximation. The upper and lower approximation will be computed before execution of the main procedure.

IV. EXPERIMENTAL RESULTS

All the necessary experiments were carried out in a workstation having the configuration of Intel core 2 Quad @2.4GHz, 2 GB RAM, 160GB HDD. The procedures are executed in Linux environment with *C* compiler.

The accuracy of each experiment was measured based on percentage of successful classification *(PSC)* [20] on the evaluated dataset, where

$$PSC = \frac{\text{No. of Correctly Classified Instances}}{\text{No. of Instances in Dataset}} \times 100 \qquad (3)$$

A. Results on UCI Dataset

The proposed method was tested on several real life datasets from UCI Machine Learning Repository dataset [22] and also the one given in [10]. We implemented our proposed method using *C*. The results of the experiments are reported in *Table 2*. It can be observed from the table that it performs consistently well for categorical dataset. Since, the method has been specially designed for handling inconsistency in the dataset, it expects the occurrence of at least some inconsistencies in the dataset. Another important advantage of the method is its input order independency. As can be observed from *Table II* that for the UCI Machine Learning Repository datasets like mushroom, glass identification, breast cancer etc, the algorithm has been able to generate rules which classify with more than *90%* accuracy. The example of the generated rules for different UCI dataset are reported in *Table III*. An interesting observation is that, number of rule generation is not dependent on the number of instances in the dataset. For example, as can be seen from the *Table II* that Mushroom [22] dataset has the maximum number of instances, i.e., 8124, however, the number of rule generated (as can be seen from *Table II*) for this dataset is not maximum. But with the increase in the dimensionality, the number of rules generated also increases as given in *Table II*, as it is evident for the case of Soyabean-Small [22] dataset as shown in *Table II*.

B. Result on Real Life Network Intrusion Dataset

The proposed method was also evaluated using our own dataset that include various type of features extracted based on net-flow data captured using our local network. Using some of the existing attack tools, we generated a group of attacks against a local network server and collected the produced traffic as known attack traffic. The existing attacks are generated using tools found in [23].

A flow is a unidirectional series of IP (internet protocol) packets passing through an observation point in the network during a certain time interval. All packets belonging to a particular flow have a set of common properties. *NetFlow* is a network protocol based collection of summarized IP traffic information. We used open source collecting tool *nfdump* for receiving the exported flow records from network devices. For gathering the normal traffic, we recorded samples of the usual traffic of the network within 4 weeks period. Thus, we did collection of 1,48,712 net-flow records of 16 attack types and normal records. The extracted net-flow level features are reported in *Table IV*.

The results of the net-flow intrusion dataset is given in *Table V*. The detection performance of the method over net-flow intrusion dataset is well. The *PSC* of net-flow intrusion dataset, in case of normal class, is found as *99.94%* whereas for all-attacks class, it is *96.21%*. Examples of the generated rules for net-flow intrusion dataset are given in *Table VI*.

V. CONCLUSION AND FUTURE WORKS

This work proposes a classification rule generation method based on LEM2 algorithm. The proposed method is typically employable in those datasets which have inconsistencies. The method has been found to exhibit satisfactory performance whenever the dataset contains inconsistencies at least for some concepts. We have tested our rule generation method on several, real life UCI machine learning repository datasets for the classification and the results have been found satisfactory. The experimental results discussed in the earlier section demonstrate the effectiveness of the proposed method.

The method covers only the local covering option. For every concept, it generates a minimum, non-redundant set of classification rules. However, the method is silent to address the generation of minimum, non-redundant classification rule set collectively over the whole dataset, that is global covering. There are scopes to consider the global covering option as well. It might have yield better results if we go for a mixed approach that local as well as global coverings. We are working towards LEM2 algorithm based minimal rule generation for other network intrusion datasets.

REFERENCES

[1] Srikant, R., Vu, Q., and Agrawal, R. (1997) Mining association rules with item constraints. Proc. of the 3rd International Conference on Knowledge Discovery and Data Mining, California USA, August, pp. 67-73. AAAI Press.

[2] Han, J. and Kamber, M. (2001) Data Mining: Concepts and Techniques. Morgan Kaufmann Publishers, 500 Sansome Street, Suite 400, San Francisco, CA 94111.

[3] Sai, Y., Nie, P., Xu, R., and Huang, J. (2006) A rough set approach to mining concise rules from inconsistent data. Proc of IEEE GrC 2006, Atlanta USA, May 10-12, pp. 333-336. IEEE.

[4] J.W., J. P. and J.W., R. B. (2002) Rule generation and model selection used for medical diagnosis. Journal of Intelligent & Fuzzy Systems: Applications in Engineering and Technology - Challenges for future intelligent systems in biomedicine, 12(1), 69-78.

[5] Grosan, C. and Abraham, A. (2006) Stock market modeling using genetic programming ensembles. Genetic Systems Programming: Theory and Experiences, 13, 133-148.

[6] Vollmer, T., Foss, J. A., and Manic, M. (2011) Autonomous rule creation for intrusion detection. Proc. of SCCI 2011, Paris, France, Apr.11-15, 2011, pp. 1-8. IEEE.

[7] Agrawal, R., Imielinski, T., and Swami, A. (1993) Mining association rules between sets of items in large databases. Proc of 1993 ACM SIGMOD, Washington, DC, USA, May 25-28, pp. 207-216. ACM, New York, NY, USA.

[8] Kiran, R. U. and Reddy, P. K. (2010) Mining rare association rules in the datasets with widely varying items' frequencies. Lecture Notes in Computer Science, 5981/2010, 49-62.

[9] Ghosh, A. and Nath, B. T. (2004) Multi-objective rule mining using genetic algorithms. Information Sciences: an International Journal, 163, 123-133.

[10] Slowinski, R. (1992) In Intellegent Decision Support: Handbook of Applications and Advances of the Rough Set Theory. Kluwer Academic Publishers Norwell, MA, USA.

[11] Pawlak, Z., Grazymala-Busse, J. W., Slowinski, R., and Ziarko, W. (1995) Rough sets. Communications of the ACM, 38, 88-95.

[12] Fernandez, M. C., Menasalvas, E., scar Marban, Pena, J. M., and Millan, S. (2001) Minimal decision rules based on the apriori algorithm. International Journal of Applied Mathematics & Computer Science, 11, 691-704.

[13] Grzymala-Busse, J. W. (1997) A new version of the rule induction system lers. Fundamenta Informaticae, 31, 27-39.

[14] Shichao, Z. and Xindong, W. (2011) Fundamentals of association rules in data mining and knowledge discovery. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 1, 97-116.

[15] Agrawal, R. and Srikant, R. (1994) Fast algorithms for mining association rules in large databases. In Bocca, J. B., Jarke, M., and Zaniolo, C. (eds.), Proc of VLDB'94, Santiago de Chile, Chile, September 12-15, pp. 487-499. Morgan Kaufmann.

[16] Lin, D. and Kedem, Z. (1998) Pincer-search: A new algorithm for discovering the maximum frequent set. Proc of EDBT '98, Valencia, Spain, March 23-27, pp. 105-119. Springer-Verlag London, UK.

[17] Han, J., Pei, J., and Yin, Y. (2000) Mining frequent patterns without candidate generation. Proc of ACM SIGMOD '00, NY, USA, May 14-19, pp. 1-12. ACM New York.

[18] Qodmanan, H. R., Nasiri, M., and Minaei-Bidgoli, B. (2011) Multi objective association rule mining with genetic algorithm without specifying minimum support and minimum con_dence. Expert Systems with Applications, 38, 288-298.

[19] Li, J. and Cercone, N. (2005) A rough set based model to rank the importance of association rules. Lecture Notes in Computer Science, 3642/2005, 109-118.

[20] Adetunmbi, A. O., Falaki, S. O., Adewale, O. S., and Alese, B. K. (2008) Network intrusion detection based on rough set and k-nearest neighbour. International Journal of Computing and ICT Research, 2, 60-66.

[21] Grzymala-Busse, J. W. (1988) Knowledge acquisition under uncertainty - a rough set approach. Journal of Intelligent & Robotic Systems, 1, 3-16.

[22] Blake, C. L. and Merz, C. J. (2001). UCI Machine Learning Repository. Irvine, CA: University of California, Department of Information and Computer Science, http://www.ics.uci. edu / ~ mlearn/MLRepository.html.

[23] (2003). Attacks tools and information. http://packetstormsecurity.nl/index.html.

[24] Szathmary, L., Valtchev, P., and Napoli, A. (2010) Generating rare association rules using the minimal rare itemsets family. International Journal of Software Informatics, 4, 219-238.

❖❖❖

Secured Data Access in Cloud Computing

B. Krishna Prasad

Pydah College of Engineering & Technology, Affiliated to JNTUK, Visakhapatnam, India
E-mail: Kp_bommaganti@yahoo.co.in

***Abstract* -** Cloud computing is an emerging computing paradigm in which resources of the computing infrastructure are provided as services over the Internet. This paper proposed some services for data security and access control when users outsource sensitive data for sharing on cloud servers Extensive analysis shows that our proposed scheme is highly efficient and provably secures under existing security models.

I. INTRODUCTION

Cloud computing is also facing many challenges that, if not well resolved, may impede its fast growth. Data security, as it exists in many other applications, is among these challenges that would raise great concerns from users when they store sensitive information on cloud servers. These concerns originate from the fact that cloud servers are usually operated by commercial providers which are very likely to be outside of the trusted domain of the users. Data confidential against cloud servers is hence frequently desired when users outsource data for storage in the cloud. In some practical application systems, data confidentiality is not only a security/privacy issue, but also of juristic concerns. For example, in healthcare application scenarios use and disclosure of protected health information (PHI) should meet the requirements of Health Insurance Portability and Accountability Act (HIPAA) [5], and keeping user data confidential against the storage servers is not just an option, but a requirement. Furthermore, we observe that there are also cases in which cloud users themselves are content providers. They publish data on cloud servers for sharing and need fine-grained data access control in terms of which user (data consumer) has the access privilege to which types of data. In the healthcare case, for example, a medical center would be the data owner who stores millions of healthcare records in the cloud. It would allow data consumers such as doctors, patients, researchers and etc, to access various types of healthcare records under policies admitted by HIPAA. To enforce these access policies, the data owners on one hand would like to take advantage of the abundant resources that the cloud provides for efficiency and economy; on the other hand, they may want to keep the data contents confidential against cloud servers.

We address this open issue and propose a secure and scalable fine-grained data access control scheme for cloud computing. Our proposed scheme is partially based on our observation that, in practical application scenarios each data file can be associated with a set of attributes which are meaningful in the context of interest. The access structure of each user can thus be defined as a unique logical expression over these attributes to reflect the scope of data files that the user is allowed to access. As the logical expression can represent any desired data file set, fine-grainedness of data access control is achieved. To enforce these access structures, we define a public key component for each attribute. Data files are encrypted using public key components corresponding to their attributes. User secret keys are defined to reflect their access structures so that a user is able to decrypt a ciphertext if and only if the data file attributes satisfy his access structure. Such a design also brings about the efficiency benefit, as compared to previous works, in that, 1) the complexity of encryption is just related the number of attributes associated to the data file, and is independent to the number of users in the system; and 2) data file creation/deletion and new user grant operations just affect current file/user without involving system-wide data file update or re-keying. One extremely challenging issue with this design is the implementation of user revocation, which would inevitably require re-encryption of data files accessible to the leaving user, and may need update of secret keys for all the remaining users. If all these tasks are performed by the data owner himself/herself, it would introduce a heavy computation overhead on him/her and may also require the data owner to be always online. To resolve this challenging issue, our proposed scheme enables the data owner to delegate tasks of data file re-encryption and user secret key update to cloud servers without disclosing data contents or user access privilege information. We achieve our design goals by exploiting a novel cryptographic primitive, namely key policy attribute-based encryption

II. MODELS AND ASSUMPTIONS

A. System Models

Similar to , we assume that the system is composed of the following parties: the Data Owner, many Data Consumers, many Cloud Servers, and a Third Party Auditor if necessary. To access data files shared by the data owner, Data Consumers, or users for brevity, download data files of their interest from Cloud Servers and then decrypt. Neither the data owner nor users will be always online. They come online just on the necessity basis. For simplicity, we assume that the only access privilege for users is data file reading. Extending our proposed scheme to support data file writing is trivial by asking the data writer to sign the new data file on each update as does. From now on, we will also call data files by files for brevity. Cloud Servers are always online and operated by the Cloud Service Provider (CSP). They are assumed to have abundant storage capacity and computation power. The Third Party Auditor is also an online party which is used for auditing every file access event. In addition, we also assume that the data owner can not only store data files but also run his own code on Cloud Servers to manage his data files. This assumption coincides with the unified ontology of cloud computing

B. Security Models

In this work, we just consider Honest but Curious Cloud Servers as does. That is to say, Cloud Servers will follow our proposed protocol in general, but try to find out as much secret information as possible based on their inputs. More specifically, we assume Cloud Servers are more interested in file contents and user access privilege information than other secret information. Cloud Servers might collude with a small number of malicious users for the purpose of harvesting file contents when it is highly beneficial. Communication channel between the data owner/users and Cloud Servers are assumed to be secured under existing security protocols such as SSL. Users would try to access files either within or outside the scope of their access privileges. To achieve this goal, unauthorized users may work independently or cooperatively. In addition, each party is preloaded with a public/private key pair and the public key can be easily obtained by other parties when necessary.

C. Design Goals

Our main design goal is to help the data owner achieve fine-grained access control on files stored by Cloud Servers. Specifically, we want to enable the data owner to enforce a unique access structure on each user, which precisely designates the set of files that the user is allowed to access. We also want to prevent Cloud Servers from being able to learn both the data file contents and user access privilege information. In addition, the proposed scheme should be able to achieve security goals like user accountability and support basic operations such as user grant/revocation as a general one-to-many communication system would require. All these design goals should be achieved efficiently in the sense that the system is scalable.

III. SYSTEM STUDY

3.1. EXISTING SYSTEM

Our existing solution applies cryptographic methods by disclosing data decryption keys only to authorized users. These solutions inevitably introduce a heavy computation overhead on the data owner for key distribution and data management when fine grained data access control is desired, and thus do not scale well.

Disadvantages

- **Software update/patches** - could change security settings, assigning privileges too low, or even more alarmingly too high allowing access to your data by other parties.

- **Security concerns** - Experts claim that their clouds are 100% secure - but it will not be their head on the block when things go awry. It's often stated that cloud computing security is better than most enterprises. Also, how do you decide which data to handle in the cloud and which to keep to internal systems - once decided keeping it secure could well be a full-time task?

- **Control** - Control of your data/system by third-party. Data - once in the cloud always in the cloud! Can you be sure that once you delete data from your cloud account will it not exist any more... ...or will traces remain in the cloud?

3.2. PROPOSED SYSTEM

3.2.1. Main Idea

In order to achieve secure, scalable and fine-grained access control on outsourced data in the cloud, we utilize and uniquely combine the following three advanced cryptographic techniques:

- Key Policy Attribute-Based Encryption (KP-ABE).
- Proxy Re-Encryption (PRE)
- Lazy re-encryption

Advantages

- Low initial capital investment

- Shorter start-up time for new services
- Lower maintenance and operation costs
- Higher utilization through virtualization
- Easier disaster recovery

More specifically, we associate each data file with a set of attributes, and assign each user an expressive access structure which is defined over these attributes. To enforce this kind of access control, we utilize KP-ABE to escort data encryption keys of data files. Such a construction enables us to immediately enjoy fine-grainedness of access control. However, this construction, if deployed alone, would introduce heavy computation overhead and cumbersome online burden towards the data owner, as he is in charge of all the operations of data/user management. Specifically, such an issue is mainly caused by the operation of user revocation, which inevitably requires the data owner to re-encrypt all the data files accessible to the leaving user, or even needs the data owner to stay online to update secret keys for users. To resolve this challenging issue and make the construction suitable for cloud computing, we uniquely combine PRE with KP-ABE and enable the data owner to delegate most of the computation intensive operations to Cloud Servers without disclosing the underlying file contents. Such a construction allows the data owner to control access of his data files with a minimal overhead in terms of computation effort and online time, and thus fits well into the cloud environment. Data confidentiality is also achieved since Cloud Servers are not able to learn the plaintext of any data file in our construction. For further reducing the computation overhead on Cloud Servers and thus saving the data owner's investment, we take advantage of the lazy re-encryption technique and allow Cloud Servers to "aggregate" computation tasks of multiple system operations. As we will discuss in section V-B, the computation complexity on Cloud Servers is either proportional to the number of system attributes, or linear to the size of the user access structure/tree, which is independent to the number of users in the system. Scalability is thus achieved. In addition, our construction also protects user access privilege information against Cloud Servers. Accountability of user secret key can also be achieved by using an enhanced scheme of KP-ABE.

3.2.2. Definition and Notation

For each data file the owner assigns a set of meaningful attributes which are necessary for access control. Different data files can have a subset of attributes in common. Each attribute is associated with a version number for the purpose of attribute update as we will discuss later. Cloud Servers keep an attribute history list *AHL* which records the version evolution history of each attribute and PRE keys used. In addition to these meaningful attributes, we also define one dummy attribute, denoted by symbol Att_D for the purpose of key management. Att_D is required to be included in every data file's attribute set and will never be updated. The access structure of each user is implemented by an access tree. Interior nodes of the access tree are threshold gates. Leaf nodes of the access tree are associated with data file attributes. For the purpose of key management, we require the root node to be an *AND* gate (i.e., *n*-of-*n* threshold gate) with one child being the leaf node which is associated with the dummy attribute, and the other child node being any threshold gate. The dummy attribute will not be attached to any other node in the access tree. Fig.1 illustrates our definitions by an example. In addition, Cloud Servers also keep a user list *UL* which records *ID*s of all the valid users in the system.

3.2.3. Scheme Description

For clarity we will present our proposed scheme in two levels: *System Level* and *Algorithm Level*. At system level, we describe the implementation of high level operations, i.e., *System Setup*, *New File Creation*, *New User Grant*, and *User Revocation*, *File Access*, *File Deletion*, and the interaction between involved parties. At algorithm level, we focus on the implementation of low level algorithms that are invoked by system level operations.

1) System Level Operations: System level operations in our proposed scheme are designed as follows.

System Setup In this operation, the data owner chooses a security parameter κ and calls the algorithm level interface *ASetup*(*k*), which outputs the system public parameter *PK* and the system master key *MK*. The data owner then signs each component of *PK* and sends *PK* along with these signatures to Cloud Servers.

New File Creation Before uploading a file to Cloud Servers, the data owner processes the data file as follows.

- Select a unique *ID* for this data file;
- Randomly select a symmetric data encryption key $DEK \overset{R}{\leftarrow} K$, where *K* is the key space, and encrypt the data file using *DEK*;
- Define a set of attribute *I* for the data file and encrypt *DEK* with *I* using KP-ABE, i.e., $(\tilde{E}, \{E_i\}_{i\in I}) \leftarrow AEncrypt(I,DEK,PK)$.

New User Grant When a new user wants to join the system, the data owner assigns an access structure and the corresponding secret key to this user as follows.

- Assign the new user a unique identity *w* and an access structure *P*;
- Generate a secret key *SK* for *w*, i.e., $SK \leftarrow AKeyGen(P,MK)$;
- Encrypt the tuple $(P, SK,PK, \delta_{O,(P,SK,PK)})$ with user *w*'s public key, denoting the cipher-text by *C*;
- Send the tuple $(T,C, \delta_{O,(T,C)})$ to Cloud Servers, where *T* denotes the tuple $(w, \{j, sk_j\}_{jLP \setminus AttD})$. On receiving the tuple $(T,C, \delta_{O,(T,C)),}$ Cloud Servers processes as follows.
- Verify $\delta_{O,(T,C)}$ and proceed if correct;
- Store *T* in the system user list *UL*;
- Forward *C* to the user.

On receiving *C*, the user first decrypts it with his private key. Then he verifies the signature $\delta_{O,(P,SK,PK)}$. If correct, he accepts (*P, SK,PK*) as his access structure, secret key, and the system public key.

As described above, Cloud Servers store all the secret key components of *SK* except for the one corresponding to the dummy attribute Att_D. Such a design allows Cloud Servers to update these secret key components during user revocation as we will describe soon. As there still exists one undisclosed secret key component (the one for Att_D), Cloud Servers cannot use these known ones to correctly decrypt ciphertexts. Actually, these disclosed secret key components, if given to any unauthorized user, do not give him any extra advantage in decryption as we will show in our security analysis.

User Revocation - We start with the intuition of the user revocation operation as follows. Whenever there is a user to be revoked, the data owner first determines a minimal set of attributes without which the leaving user's access structure will never be satisfied. Next, he updates these attributes by redefining their corresponding system master key components in *MK*. Public key components of all these updated attributes in *PK* are redefined accordingly. Then, he updates user secret keys accordingly for all the users except for the one to be revoked. Finally, *DEK*s of affected data files are re-encrypted with the latest version of *PK*. The main issue with this intuitive scheme is that it would introduce a heavy computation overhead for the data owner to re-encrypt data files and might require the data owner to be always online to provide secret key update service for users. To resolve this issue, we combine the technique of proxy re-encryption with KP-ABE and delegate tasks of data file re-encryption and user secret key update to Cloud Servers. More specifically, we divide the user revocation scheme into two stages as is shown below.

// to revoke user *v*

// stage 1: attribute update.

The Data Owner

Cloud Servers

1. $D \leftarrow AMinimalSet(P)$, where *P* is *v*'s access structure; remove *v* from the system user list *UL*;
2. for each attribute *i* in *D* for each attribute *i* □ *D*

   $(t'_i, T'_i , rk_{i\leftrightarrow i'})\leftarrow AUpdateAtt(i,MK)$; $\xrightarrow{Att}$ store $(i, T\_i , \delta O,(i,T \_i))$;
3. send $Att = (v,D, \{i, T'_i , \delta_{O,(i,T 'i)}, rk_{i\leftrightarrow i'}\}_{i \,□\, D})$. add $rk_{i\leftrightarrow i'}$ to *i*'s history list AHL_i.

   // stage 2: data file and user secret key update.

Cloud Servers

User (*u*)

1. on receiving *REQ*, proceed if *u* □ *UL*;
2. get the tuple $(u, \{j, sk_j\}_{j□LP \setminus AttD})$;

   1. generate data file access request *REQ*; for each attribute *j* □ $LP \setminus AttD \xleftarrow{REQ}$
   2. wait for the response from Cloud Servers; $sk'_j\leftarrow AUpdateSK(j, sk_j,AHL_j)$; for each requested file *f* in *REQ*
   3. on receiving *RESP*, verify for each attribute *k* □ $I_f \xrightarrow{RESP}$ each $\delta_{O,(j,T 'j)}$ and sk'_j; proceed if correct; $E'_k\leftarrow AUpdateAtt4File(k,E_k,AHL_k)$;
   4. replace each sk_j in *SK* with sk'_j;
3. send $RESP = (\{j, sk'_j, T'_j, \delta_{O,(j,T 'j)}\}_{j□ LP \setminus AttD}, FL)$.
4. decrypt each file in *FL* with *SK*.

Description of the process of user revocation

In the first stage, the data owner determines the minimal set of attributes, redefines *MK* and *PK* for involved attributes, and generates the corresponding PRE keys. He then sends the user's *ID*, the minimal attribute set, the PRE keys, the updated public key components, along with his signatures on these components to Cloud Servers, and can go off-line again. Cloud Servers, on receiving this message from the data owner, remove the revoked user from the system user list *UL*, store the updated public key components as well as the owner's signatures on them, and record the PRE key of the latest version in the attribute history list *AHL* for each updated attribute. *AHL* of each attribute is a list used to record the version evolution history of this attribute as well as the PRE keys used. Every attribute has its own *AHL*. With *AHL*, Cloud Servers are able to compute a single PRE key that enables them to update the attribute from any historical version to the latest version. This property allows Cloud Servers to update user secret keys and data files in the "lazy" way as follows. Once a user revocation event occurs, Cloud Servers just record information submitted by the data owner as is previously discussed. If only there is a file data access request from a user, do Cloud Servers re-encrypt the requested files and update the requesting user's secret key. This statistically saves a lot of computation overhead since Cloud Servers are able to "aggregate" multiple update/re-encryption operations into one if there is no access request occurring across multiple successive user revocation events.

File Access - This is also the second stage of user revocation. In this operation, Cloud Servers respond user request on data file access, and update user secret keys and re-encrypt requested data files if necessary. As is depicted in Fig. 4, Cloud Servers first verify if the requesting user is a valid system user in *UL*. If true, they update this user's secret key components to the latest version and re-encrypt the *DEK*s of requested data files using the latest version of *PK*. Notably; Cloud Servers will not perform update/re-encryption if secret key components/data files are already of the latest version. Finally, Cloud Servers send updated secret key components as well as ciphertexts of the requested data files to the user. On receiving the response from Cloud Servers, the user first verifies if the claimed version of each attribute is really newer than the current version he knows. For this purpose, he needs to verify the data owner's signatures on the attribute information (including the version information) and the corresponding public key components, i.e., tuples of the form (j, T'_j) in Fig. 4. If correct, the user further verifies if each secret key component returned by Cloud Servers is correctly computed. He verifies this by computing a bilinear pairing between sk'_j and T'_j and comparing the result with that between the old sk_j and T_j that he possesses. If verification succeeds, he replaces each sk_j of his secret key with sk'_j and update T_j with T'_j. Finally, he decrypts data files by first calling *ADecrypt*(*P, SK,E*) to decrypt *DEK*'s and then decrypting data files using *DEK*'s.

File Deletion - This operation can only be performed at the request of the data owner. To delete a file, the data owner sends the file's unique *ID* along with his signature on this *ID* to Cloud Servers. If verification of the owner's signature returns true, Cloud Servers delete the data file. *2) Algorithm level operations:* Algorithm level operations include eight algorithms: *ASetup*, *AEncrypt*, *AKeyGen*, *ADecrypt*, *AUpdateAtt*, *AUpdateSK*, *AUpdateAtt*4*File*, and *AMinimalSet*. As the first four algorithms are just the same as *Setup*, *Encryption*, *Key Generation*, and *Decryption* of the standard KP-ABE respectively, we focus on our implementation of the last four algorithms.

AUpdateAtt - This algorithm updates an attribute to a new version by redefining its system master key and public key component. It also outputs a proxy re-encryption key between the old version and the new version of the attribute.

*AUpdateAtt*4*File* - This algorithm translates the ciphertext component of an attribute *i* of a file from an old version into the latest version. It first checks the attribute history list of this attribute and locates the position of the old version. Then it multiplies all the PRE keys between the old version and the latest version and obtains a single PRE key. Finally it apply this single PRE key to the ciphertext component E_i and returns $E^{(n)}_i$ which coincides with the latest definition of attribute *i*.

AUpdateSK - This algorithm translates the secret key component of attribute *i* in the user secret key *SK* from an old version into the latest version. Its implementation is similar to *AUpdateAtt*4*File* except that, in the last step it applies $(rk_{i\leftrightarrow i(n)})^{-1}$ to SK_i instead of $rk_{i\leftrightarrow i(n)}$. This is because t_i is the denominator of the exponent part of SK_i while in E_i it is a numerator.

AMinimalSet - This algorithm determines a minimal set of attributes without which an access tree will never be satisfied. For this purpose, it constructs the conjunctive normal form (CNF) of the access tree, and returns attributes in the shortest clause of the CNF formula as the minimal attribute set.

3.3. FEASIBILITY STUDY

The feasibility of the project is analyzed in this phase and business proposal is put forth with a very general plan for the project and some cost estimates. During system analysis the feasibility study of the proposed system is to be carried out. This is to ensure that the proposed system is not a burden to the company. For feasibility analysis, some understanding of the major requirements for the system is essential. Three key considerations involved in the feasibility analysis are

- ♦ ECONOMICAL FEASIBILITY
- ♦ TECHNICAL FEASIBILITY
- ♦ SOCIAL FEASIBILITY

3.3.1. Economical Feasibility

This study is carried out to check the economic impact that the system will have on the organization. The amount of fund that the company can pour into the research and development of the system is limited. The expenditures must be justified. Thus the developed system as well within the budget and this was achieved because most of the technologies used are freely available. Only the customized products had to be purchased.

3.3.2. Technical Feasibility

This study is carried out to check the technical feasibility, that is, the technical requirements of the system. Any system developed must not have a high demand on the available technical resources. This will lead to high demands on the available technical resources. This will lead to high demands being placed on the client. The developed system must have a modest requirement, as only minimal or null changes are required for implementing this system.

3.3.3. Social Feasibility

The aspect of study is to check the level of acceptance of the system by the user. This includes the process of training the user to use the system efficiently. The user must not feel threatened by the system, instead must accept it as a necessity. The level of acceptance by the users solely depends on the methods that are employed to educate the user about the system and to make him familiar with it. His level of confidence must be raised so that he is also able to make some constructive criticism, which is welcomed, as he is the final user of the system.

REFERENCES

[1] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. H. Katz, A. Konwinski, G. Lee, D. A. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, “Above the clouds: A berkeley view of cloud computing,” University of California, Berkeley, Tech. Rep. USB-EECS-2009-28, Feb 2009.

[2] Amazon Web Services (AWS), online at http://aws.amazon.com.

[3] Google App Engine, Online at http://code.google.com/appengine/.

[4] Microsoft Azure, http://www.microsoft.com/azure/.

[5] 104th United States Congress, “Health Insurance Portability and Accountability Act of 1996 (HIPPA),” Online at http://aspe.hhs.gov/admnsimp/pl104191.htm, 1996.

[6] H. Harney, A. Colgrove, and P. D. McDaniel, “Principles of policy in secure groups,” in Proc. of NDSS’01, 2001.

◈◈◈

Blind And Robust Audio Watermarking in Wavelet Domain

Suresh Penchala[1] & K. Sivani[2]

[1]ECE Dept., Kamala Institute of Technology & Science, Singapur, Andhra Pradesh, India.
[2] Talla Padmavathi College of Engg, Kazipet, Andhra Pradesh, India.
Email: penchalasuresh@gmail.com, k_sivani@yahoo.co.in.

Abstract - This paper presents an algorithm to embed a watermark in an audio signal and to extract using wavelet transform. The algorithm presents a method of embedding a binary image into the selected audible frequencies of an audio signal according to psychoacoustic model on the basis of SNR. The extraction algorithm doesn't require the original audio signal to extract the watermark. The extracted watermark image quality can be detected by considering normalized correlation value with a suitable scaling parameter for embedding. Experimental results demonstrate that the watermark is inaudible and this algorithm is robust to many operations of digital audio signal processing, such as, low pass filtering, additive white Gaussian noise and so on.

Key words : *Digital watermarking, discrete wavelet transform (DWT), Psychoacoustic model*

I. INTRODUCTION

The proliferation of digitized multimedia systems in distributed environments and an explosion of data exchange in the Internet facilitate digital owners that they can quickly and massively transfer multimedia documents across the Internet. This leads to wide interest in multimedia security and multimedia copyright protection of multimedia contents.

Traditionally, encryption and control access techniques were employed to protect the ownership of media. These techniques, however, do not protect against unauthorized copying after the media have been successfully transmitted and decrypted. Recently, watermark techniques are utilized to maintain the copyright. Digital watermarking is a process that embeds a perceptually undetectable signature to multimedia content. It can occur over a variety of media such as pictures or movies, but audio watermarking is of particular interest as companies to attempt to protect content from unauthorized user. Embedded watermark contains information related uniquely to the owner or distributor of the multimedia file. There are several techniques in which a watermark can be embedded in audio file.

There are several techniques in which a watermark can be embedded in an audio file [1, 2]. The watermark can be applied on wavelet transform [3], and cepstrum domain [4], that is altering the original content of the file. In time domain, the algorithm developed by Wen-Nung Lie and Li Chun-chang [5] used the relative energy relation between three consecutive sample section and finding energy maximum to embed watermark. Bassia and Pitas [6] applied a very straight forward time-domain spread spectrum watermarking technique to audio signal. Sang-kwang Lee and Yo-Sung Ho [7] proposed technique spread spectrum communication, hiding a narrow band signal in a wide band channel. Jong Won Seok and Jin Woo Hong [8] presented a novel-watermarking scheme using human auditory system based on FFT.

This paper is focused on presenting an algorithm for audio watermarking using binary image in wavelet domain based on psychoacoustic model. The rest of the paper is organized as follows; Section 2 gives the basic concepts about psychoacoustic model. Section 3 deals with the embedding algorithm and extraction algorithm. Section 4 gives the experimental results and Section 5 gives the conclusions.

II. BASICS

The hearing ability of humans depends on two facts.

Fact 1: The frequency range of the audio signal which is about 20 HZ-20 kHz and

Fact 2: Sound pressure level (SPL) in each frequency.

MPEG compression is done based on above two facts. Hence, embed a watermark where it has a higher of surviving the compression and at the same time, it won't interfere with the humans hearing.

Based on the first fact, the frequency range of 4 kHz – 15 kHz is selected to embed the watermark, because they will survive against MPEG compression and with a little noise, human ears cannot detect it. Based on the second fact, the frequency components which have SPL higher than the threshold of hearing are selected to embed the watermark, because human hears cannot detect the difference if the SPL is higher than the threshold of hearing.

III. PROPOSED WATERMARKING ALGORITHM

A. Embedding algorithm:

The procedure of embedding watermark is as follows:

1) The original audio signal (A) is decomposed using wavelet transform at level L and then obtains the decomposition vector consisting of approximated and detailed coefficients.

2) Choose the embedding area into which the watermark has to be embedded. The area to be chosen must ensure the imperceptivity of the watermark and also robustness to be considered. Choose the detailed coefficients instead of approximated coefficients to ensure imperceptivity.

3) According to the psychoacoustic model the coefficients, which are having strength above threshold level of hearing, are chosen and are stored in a vector.

4) Among the chosen coefficients, N (N depends on the image size to be embedded) strongest coefficients are selected with large magnitude and compose the vector V and store the corresponding positions in vector I.

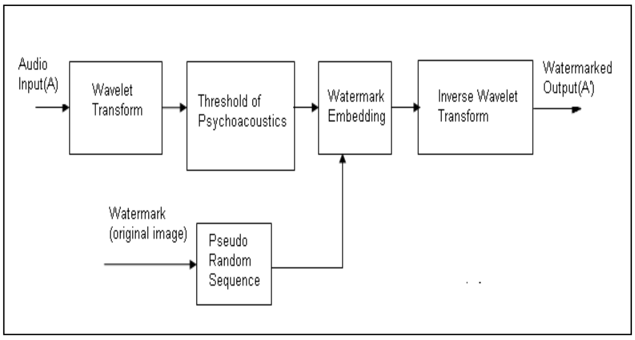


Figure 1. Watermark embedding process

5) A pseudo random sequence is generated which will be used as secret key to provide more security.

6) Binary image pixel values are sequenced and added to the pseudo random sequence to produce pseudo random watermark sequences.

7) Now the watermark sequences are added to the coefficients, which have been selected to embed the watermark by following equation.

$$CD'L\ (I\ (j)) = CDL\ (I\ (j)) + \alpha\ \omega \qquad 1 \leq j \leq N \qquad (1)$$

where I (j) represent the position of the jth important coefficients. α represents the scaling parameter which determines the watermark strength, varies 0 ~ 1. Increasing the value of α will enhance the robustness and impair the transparency. ω is the watermark coefficient vector that will be embedded.

8) Inverse wavelet transform is applied to obtain the watermarked audio signal (A′) by reconstructing the signal from altered coefficients.

B. Extraction Algorithm

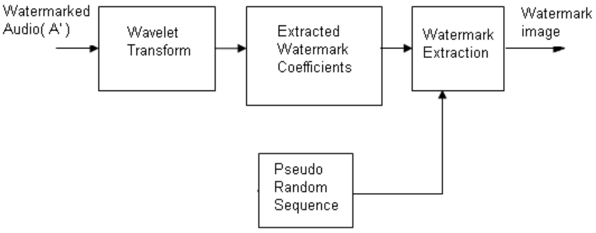


Figure 2. Block Diagram to Extract Watermark

The watermark should be extractable even if common signal processing operations are applied to the original audio signal. In extraction process (blind), original audio signal is not needed.

The arrangement steps are as followed:

1) Perform discrete wavelet transform at level L to the signal (A′) which is to be detected and extract the detailed coefficients CD′L.

2) Extract the transformed coefficients of CD′L in responding location of I and constitute the vector V′ according to psychoacoustic model.

3) Extract the watermark ω″ according to the following given equation.

$$\omega'' = (V'-V)/\ \alpha \qquad (2)$$

4) Subtract the pseudorandom sequence from the extracted watermark sequence to get original watermark sequence ω′.

4. PERFORMANCE EVALUATION AND RESULTS

The similarity between watermark image and extracted image can be measured by normalized correlation (NC).

$$NC = \left| \frac{\sum_{x=1}^{m}\sum_{y=1}^{n}\omega(x,y)*\omega'(x,y)}{\sum_{x=1}^{m}\sum_{y=1}^{n}\sqrt{\omega(x,y)*\omega(x,y)}} \right| \qquad (3)$$

where ω (x, y) and ω′ (x, y) are original watermark and extracted watermark respectively.

The performance evaluation of the signal can be determined by considering the signal to noise ratio (SNR) and Normalized correlation value (NC).

The signal to noise ratio can be found by using following equation:

$$SNR= 10 \log (\Sigma(x^2(n))/ \Sigma(x (n)-y(n))^2 \qquad (4)$$

where x (n) is mean of original audio signal and y (n) is mean of watermarked audio signal.

To evaluate the imperceptibility and robustness of the proposed embedding algorithm on audio, simulation has been done on the computer. The music used as the watermarked audio is of 0.34 seconds stereo signal (sampling frequency of 44.1 kHz and 16 bit recorded for each sampling). The audio signal is shown in Figure 3. The original watermark is a binary image of 64 X 64 pixels as shown in Figure 4.

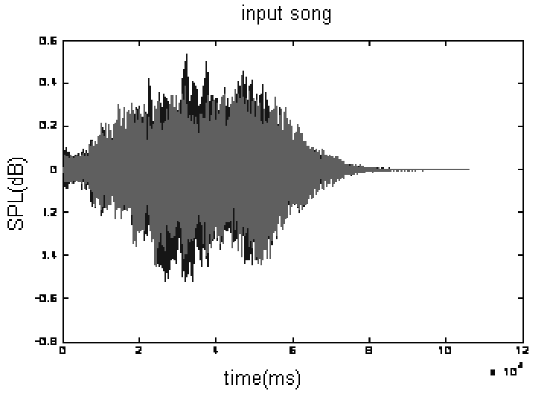


Figure 3. Original audio signal

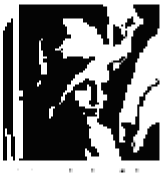


Fig. : 4. Original image   Fig. 5 : Pseudo image

Wavelet decomposition is implemented by Haar wavelet transform with 3rd level.

In order to make the watermarked signal inaudible, the watermark is embedded into low frequency part of the large magnitude in among the detailed coefficients of audio signal. In the extraction process, the dependent components V and I are used as a key for detection.

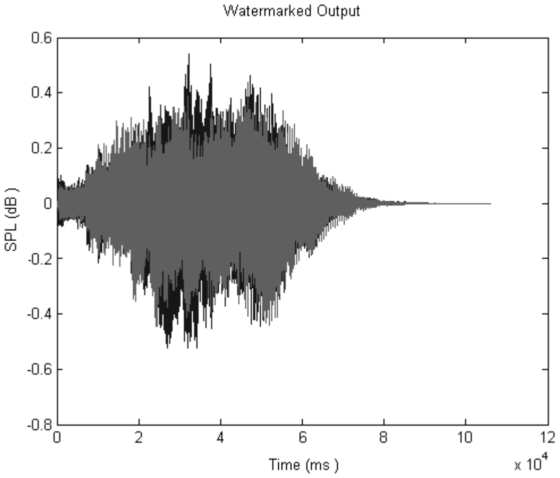
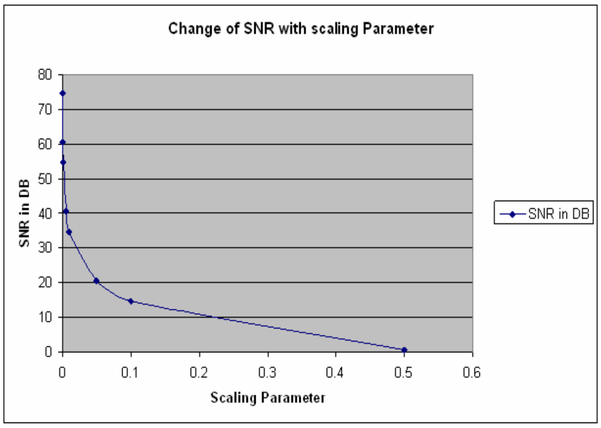
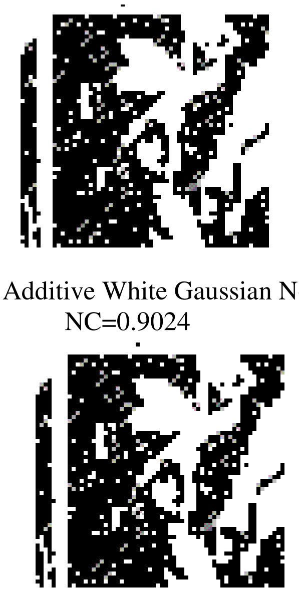


Figure 6. Watermarked Audio Signal

The experimental result show in Figure 6 are the magnitudes of the scaling parameter α among 0.0001~0.5 and SNR. In order to compromise the inaudibility and robustness of the watermark, α=0.0001 is selected as the scaling parameter in this experiment.



To detect the robustness of watermarks, the watermarked audio signal was done with several digital signal processes, such as, additive Gaussian noise, low pass filtering and resampling. Figure 8 shows the extracted watermarks after the above-mentioned attacks.



a) Additive White Gaussian Noise
NC=0.9024

b) Low pass filtering
NC=0.9394

Figure 8. Extracted watermarks after attacks

V. CONCLUSIONS

The watermarking technique based on discrete wavelet transform is proposed. The paper proposes embedding of an image watermark into audio signal in the frequency range 4 kHz-15 kHz and uses psychoacoustic model for selecting embedded position. The main characteristic of this algorithm is to adjust the watermark embedding intensity. Experimental results show that, the normalized correlation of watermark using selected scaling parameter is more than 85%. The advantage of the presented algorithm is the extraction process does not require original audio signal. And also the algorithm is robust to many attacks, such as, addition of noise, signal resampling and low pass filtering.

REFERENCES

[1] X. Tang, Y. Niu, H. Yue and Z. Yin, “A digital audio watermark embedding algorithm with WT and CCT,” in Proc. Conf. Microwave, Antenna, Propagation and EMC technologies for Wireless Communications, MAPE’05, Beijing, CHINA, pp. 970–973.

[2] X. Li, M. Zhang and R. Zhang, “A new adaptive audio watermarking algorithm,” in Proc. 5th World congress on Intelligent Control and Automation, WCICA’04, Hangzhou, CHINA, pp. 4357–4361.

[3] Y. Fu, Z. Ma and G. Song, “A robust audio watermarking algorithm based on wavelet transform,” Journal of Information & Computational Science, vol. A247, pp. 7-11, 2005.

[4] I. J. Cox, J. Kilian and T. Shamoon, “Secure Spread Spectrum Watermarking for multimedia,” IEEE Trans. Image Processing., vol. 6, pp. 1673-1687, 1997.

[5] Lie, W.N. and Chang, L.C.,“Robust and High Quality Time-Domain Audio Watermarking Subject to Psychoacoustic Masking”, Proceeding of IEEE, Vol. 75, No. 6, April 2001, pp. 45-48.

[6] Bassia, P., Pitas, I. and Nikolaidis, N., “Robust Audio Watermarking in the Time Domain”, IEEE Transactions on Multimedia, Vol. 3, No. 2, June 2001, pp. 232-241.

[7] Jong, W.S. and Jin, W.H., “Audio Watermarking for Copyright Protection of Digital Audio Data”, Electronics Letters, Vol. 37, No.1, January 2001, pp. 60-61.

[8] Lee, S. and Ho, Y., “Digital Audio Watermarking in the Ceptrum Domain”, IEEE Transactions on Consumer Electronics, Vol. 46, No. 3, August 2000, pp. 744-749.

An Improved Protocol for Proxy Based Certification Authority for MANETs

Bibhudendu Panda & Pabitra Mohan Khilar
Dept of CSE, NIT, Rourkela
E-mail: bibhu_panda25@rediffmail.com, pmkhilar@nitrkl.ac.in

Abstract **-** This paper includes developing framework for proxy based certificate Authority for MANET. Key management is a central aspect for security in mobile ad hoc networks. In mobile ad hoc networks, the computational load and complexity for key management is strongly subject to restriction of the node's available resources and the dynamic nature of the topology. PKI has been recognized as one of the most effective tools for providing security for dynamic networks. However, providing such an infrastructure in MANETs is a challenging task due to their infrastructure-less nature. In this paper we have considered such challenges in detail, identify the requirements for such solutions, and propose a practical PKI service for MANETs. We used threshold cryptography to distribute the CA functionality. To reduce the overhead research work employs a proxy and a timer during multicasting to achieve certification service. Results from simulation establish the effectiveness and the efficiency of the approach.

Keywords*-PKI,MANET,AODV,CA*

I. INTRODUCTION

In the past few years, more attention has been drawn to the security of mobile ad hoc network (MANET)[1][2]. Due to its glorious success in securing Internet computing, the Internet de facto standard public key cryptography including encryption and digital signature becomes natural choice as a fundamental building block to secure MANET. However, since MANET is significantly different from the Internet, a salient issue is how to adopt the technology to the new environment. As we know, successful application of PKI relies on the ubiquitous capability of verifying the binding between a public key and the owner principal. In the internet, mainstream solution is to have a third party centrally trusted entity, called certificate Authority(CA); vouch for the authenticity of the binding by issuing digital certificates, which in essence is a statement of the binding digitally signed by the CA. In practice, CAs and digital certificates are organized and maintained by PKI.[3][4]. It is questionable yet if PKI can be implemented in MANET because PKI requires well protected CAs and constant connectivity between users and CAs. However, MANET is composed of a group of mobile devices communicating with each other through a wireless link without a backbone or infrastructure. In such an environment, all devices are exposed to hacking to the same extent and no device can be assumed to be significantly more secure than the others. Moreover, devices roam around, run out of power or just stop functioning, which lead to volatile connectivity among them and CAs. Research proposals have been seen in[5][6][7] etc. , to address the two issues by distributing the CA's functionality across a set of network nodes and use threshold signature[8] to achieve tolerance up to the threshold number of faulty nodes. These methods are suitable for small MANETs with a single CA. This is partially due to inherent high communication cost. A more fundamental reason is that it is difficult if not feasible for a CA to get familiar with all other principals. Some researchers take another approach based on the concept of "web of trust" first appearing in PGP[9]. In these methods each principal is it's own CA[10][7] and keeps a certificate directory. To authenticate a certificate signed by another principal,a principal has to find a certificate path between them but this lead to difficulty of finding such a path without incurring a lot of broadcasting cost or forcing each principal to save a large number of certificates in it's local directory.

II. RELATED WORK

In [11], a secure and efficient key management framework(SEKM) for mobile ad hoc networks is proposed. SEKM builds PKI by applying a secret sharing scheme and an underlying multicast server group. In SEKM, the server group creates view of the certification authority(CA) and provides certificate update service for all nodes, including the server themselves. A ticket scheme is introduced for efficient certificate service. In addition, an efficient server group updating scheme is proposed.

In [12], a locality driven key management architecture that achieves robust key authentication and facilitates timely and efficient establishment of distributed trust is proposed. The architecture reflects application oriented view of MANET and is based on threshold cryptography to achieve high fault tolerance against network partition and malicious nodes. On top of it,distributed trust protocols are designed to help set up trust relations on the fly.

In [13], a practical PKI service for ad hoc networks is proposed. Threshold cryptography is employed to distribute the CA functionality over specially selected nodes based on the security and the physical characterstics of nodes. The selected nodes that collectively provides PKI functionality are called MOCAs, an efficient and effective communication protocol for correspondence with MOCAs for certification services is presented.

Pathak et al proposed in [14] a voting based scheme for both public key authentication and group membership control. In this method, the decision of trust is made collectively by a group of n principals via voting. The system achieves high fault tolerance when it satisfies Byzantine condition. Compared to the above threshold based CA solutions, the method does not require a shared trusted principal (the dealer) and therefore does not have any single point of failure. How ever, the group[18] does not own a single signing key. Consequently each individual principal has to know all public keys of the n voters and perform n signature verifications to authenticate one public key.

III. PROXY BASE CERTIFICATE AUTHORITY FOR MANETS

In this part of the research, Proxy Based Certificate Authority (PBCA) is proposed. In this frame work, all the N nodes in the network provide CA (Certificate Authority). Using threshold cryptography, any r nodes can reconstruct the full CA key. Threshold cryptography is an application of secret sharing that was first proposed by Shamir [15]. The basic idea of secret sharing is that it is mathematically possible to divide up a secret to n pieces in such way that anybody who requires the full secrete can collect any k piece out of those n to reconstruct the full secret. k becomes the threshold needed to reconstruct the secret. Threshold cryptography applies this technique to the keys for cryptographic operations. Frankel and Desmedt [16] proposed to use secret sharing for the private key of public key cryptography and Shoup proposed a way to generate a digital signature from key pieces without reconstructing the full key at any point.

In the Fig 1.1, we assume that there are 6 nodes, a proxy node in the network. A new node is joining in the network and sending a request to proxy, in turn multicast the request to all the other nodes in network. Let us assume that the minimum number of nodes need to reply be 2.

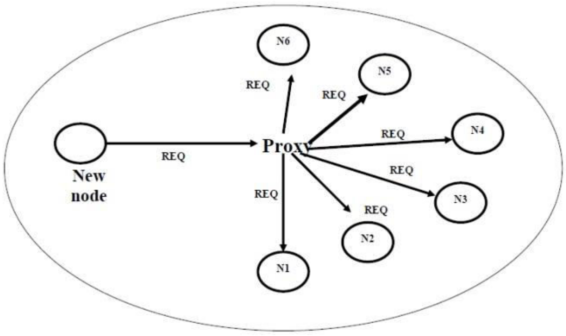
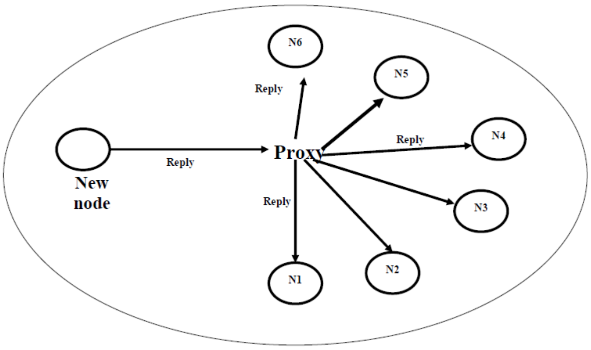


Figure 1.1: New Node sending request to proxy and proxy multicast to other nodes in the network

In Fig 1.2, it can be observed that proxy receives reply with their corresponding key share from only 3 nodes in network. As the threshold condition is satisfied, the proxy reconstructs the full signature and sends to the new node.

Any client requiring a certificate service must contact all the nodes in network with its request. This request is sent to the nodes in the network via a proxy node. So, a new client unicast the request message to the proxy and then the request message is multicast then the timer is specified which indicates that the node need to reply within the time limit or else need no reply. When the timer expires, the nodes need not reply. This technique reduces the overhead cost. The nodes which send reply will generate a partial signature over the received data. The client needs to collect at least r such partial signatures to reconstruct the full signature and successfully receive the certificate service. The reconstruction of the full signature is done by proxy node and is sent to the client. This process reduces the cost without implementing in every new client.

Maintaining information on revoked certificates is one of the key tasks of the CA and this topic has got much of the attention in recent years [17]. In our approach, the certificate can be revoked only upon the agreement of minimum r nodes in the network. So ,when r nodes come to an agreement to revoke the certificate, each node generates the revocation certificate with it's partial key. This revoke certificate is then sent to the proxy node. This proxy upon receiving r revoke certificates ,it reconstruct the full revocation certificate. This avoid false revocation. The threshold value r must be chosen carefully, such overhead does not increase or security decreases. In this protocol, the chance of not receiving the reply from the nodes and possibility of failure in reconstructing the full signature is very less because the request message is sent to all the nodes in the network unlike the MOCA.



The request and reply message are similar to the Route Request(RREQ) and Route Reply(RREP) message in on demand ad hoc routing protocol like AODV and DSR. The management of routing information is also similar to these protocols. As a request packet passes through a node, the reverse path to the sender is established. If no reply is returned within the time-out period, the reverse path entry in the routing table expires and is purged. If a reply traverses back through the previously setup reverse path to the sender, the routing table entries are refreshed and the bidirectional path remains in the routing table for potential reuse. This similarity to on demand routing presents a potential for our certification protocol and the existing on demand routing protocols to benefits from each other by sharing routing information. In our protocol, there is requirement of unicast –based optimization as the request is sent to all the nodes in the network. When any client leaves the network, then the proxy sends a request message to revoke the existing certification and to build the new certificate. Then the nodes reply back with revoke certificate with their key share. The nodes include the key share also in the revoke certificate. When proxy receives minimum of r replies from the nodes, the certificate is revoked and new certificate is generated and the respective key shares are multicast to all the nodes in the network.

IV. RESULTS

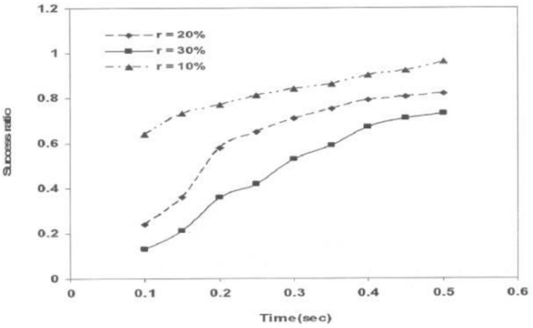
The focus of our evaluation of the PBCA framework is effectiveness and efficiency (or cost). Effectiveness is measured using the success ratio of certification requests.

$$\text{Success ratio}=\frac{\text{number of successful certification request}}{\text{Number of total certification request}}$$

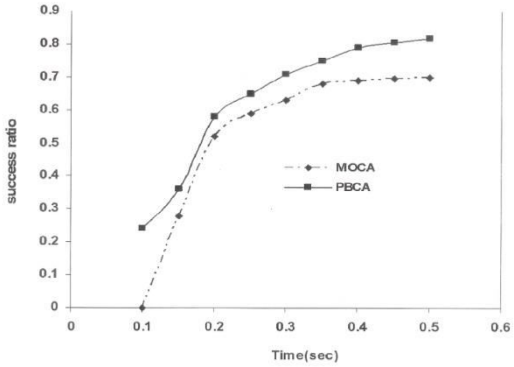
The cost of a certification protocol can be evaluated using the two metrics.

1. Packet overhead
2. Additional communication delay caused by the certification process.

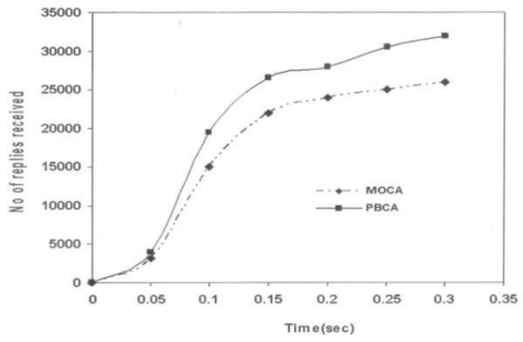
Certification delay: The most frequent use of a certification service is to acquire the communicating peer's public key certificate. The delay to get the certification service is added to the start up latency of any secure communication relying on PKI. Fig 5.5 shows the distribution of arrival times of CREP packets



It can be observed that the no of replies received is more with the proposed protocol when compared to MOCA. This is because the request is sent to all the nodes in the network but not only to selected nodes. The limitation might be the slight overhead in multicasting. The comparison of MOCA and



PBCA in terms of success ratio is shown in fig 5.6. The success ratio is more in PBCA as shown in fig 3. The minimum no of replies is considered as 20% of the total requests. Success ratio depends on the minimum of replies required to reconstruct the full certificate.



When we vary the minimum number of required replies r, the success ratio varies. When r is increased then success ratio decreases and vice versa. This variation is shown in fig 5.7 for PBCA.

V. CONCLUSIONS

In this paper we have introduced a key management framework for ad hoc wireless network and proposed proxy based certificate authority protocol. The PKI framework has many problems to be implemented in ad hoc networks. These problems are resolved using the proposed framework. The need of unicast-based optimization is eliminated. The overhead of the system is reduced by using a proxy node and the timer to receive the replies from the nodes with their key share. Once the timer is expired, it is implied to the nodes no more replies are accepted. If any node replies, it becomes an invalid data. Efficient and effective communication protocols are developed which are similar to AODV and DSR. These communication protocols give the reserved reversible path for destination to the sender. The simulation results obtained in this paper establishes that the proposed protocol is better when compared to the legacy system.

Developing PKI infrastructure solutions for MANETs is not easy and this paper made an attempt to study the issues of PKI and proposed a method based on proxy certificate with permissible delay. However, extending this solution to other networks would be an interesting problem of research. Also, studying about faulty nodes vs proxy certificates would be another future work. The proposed model can be implemented using Elliptic Curve Cryptography (ECC) to minimize the mobile node's resource utilization.

REFERENCES

[1] C.E. Perkins, E. Royer, and S.R D," AODV routing",. In 2nd IEEE workshop on mobile Computing Systems and Application (WMCSA'99), 1999.

[2] Y.C Hu, A Perig, and D.B Johnson Ariadne, " a secure on demand routing protocol for ad hoc networks", in the proceedings of the 8th annual ACM/IEEE International conference on mobile computing and Networking (Mobicom) 2002.

[3] S.Berkovits, S. Chokhani, J furlong, J.geiter, and J.Guild, public key infrastructure study final report. MITRE report, 1994.

[4] K.Drira, H.Seba, H.Kheddouci, "ECGK: An efficient clustering scheme for group key management in MANETs", computer communication 33(2010) 1094-1107.

[5] L. Zhou and Z.J Haas, " securing ad hoc networks", In IEEE networks, volume 13(6).1999.

[6] J.kong, P. Zerfos, H.Luo, S.Lu, and L.Zhang," providing robust and ubiquitious security support for mobile ad hoc networks", in proceedings of ICNP '01.

[7] S. Yi and R Kravets. Moca: ," Mobile Certificate Authority for wireless adhoc networks" , in the proceeding of the 10th IEEE International Conference on Networks protocols (ICNP'02), 2002.

[8] V. Shoup. Practical threshold signatures, theory and Application of Cryptographic Techniques 2000.

[9] P.Zimmermann, The official PGP user's guide, MIT Press,1995.

[10] T.Aura and S.M"aki", " Towards a survivable security architecture for ad hoc networks. In Security protocols", 9th International workshop, Cambridge,UK, april 2001, Volume 2467 of Lecture notes in computer science. Springer-verlag, Berlin. Springer- verlag Berlin Heidelberg 2002.

[11] Capkun, S., Hubaux, J., and Buttyan, l.2006. " mobility helps peer-to-peer security". IEEE Trans. Mobile Comput. 5, 1, 43-51..

[12] Hao yang, Haiyun Luo, fan Ye, Songwu Lu, and Lixia Zhang, " security in mobile Ad hoc networks- Challenges and solutions, " IEEE Transactions on wireless communications , Vol 11, no.1, pp. 38-47, 2004.

[13] S. Yi and R Kravets. Moca: ," Mobile Certificate Authority for wireless adhoc networks" , in the proceeding of the 2nd annual PKI research workshop(PKI'03),2003.

[14] V.Pathak and L. Iftode, “Byzantine fault tolerant authentication “, technical report, Dept of Computer science, Rutgers University , 2003

H.Luo and s.Lu,’ URSA: ubiquitous and Robust access Control for mobile Ad hoc networks’, UCLA, 2004

[15] A.Shamir,” How to share a secret”, Communication of the ACM, 1979.

[16] Y.Frankel and Y.G desmedt,” parallel Reliable Threshold Multisignature “ technical report Tr 94-04-02, University of Wisconsin Theory . IT-28:714-720.

[17] The Network Simulator- NS-2. Available at http://www.isi.edu/nsnam/ns/

[18] A.Renuka, and K.C Shet, “ Cluster based Group key Management in Mobile Ad hoc Networks,” IJCSNS International journal of Computer Science and Networks, Vol . 9, no. , pp., 42-49,2009

❖❖❖

Development of Empirical Mode Decomposition

R. Samba Siva Nayak, Satish Kumar Mandava & K. Suresh Babu

Deptt. of ECE, Don Bosco Institute of Tech & Science
Deptt. of ECE, DVR college of Engg. & Tech.
Deptt. of ECE, Paladugu Parvathi Devi college of Engg. & Tech.

Abstract - In this paper, one of the tasks for which empirical mode decomposition is potentially useful is nonparametric signal denoising, an area for which wavelet thresholding has been the dominant technique for many years. In this paper, the wavelet thresholding principle is used in the decomposition modes resulting from applying EMD to a signal. We show that although a direct application of this principle is not feasible in the EMD case, it can be appropriately adapted by exploiting the special characteristics of the EMD decomposition modes. In the same manner, inspired by the translation invariant wavelet thresholding, a similar technique adapted to EMD is developed, leading to enhanced denoising performance.

Key words - *Empirical mode decomposition, Signal denoising, Wavelet thresholding*

I. INTRODUCTION

The empirical mode decomposition method is an algorithm for the analysis of multi component signals that breaks them down into a number of amplitude and frequency modulated zero-mean signals, termed intrinsic mode functions. EMD expresses the signal as an expansion of basis functions that are signal-dependent and are estimated via an iterative procedure called sifting. Although many attempts have been made to improve the understanding of the way EMD operates or to enhance its performance, EMD still lacks a sound mathematical theory and is essentially described by an algorithm. it has found a vast number of diverse applications to name a few biomedical, watermarking, and audio processing. Apart from the specific applications of EMD listed above, a more generalized task in which EMD can prove useful is signal denoising. In this paper, inspired by standard wavelet thresholding and translation invariant thresholding, a number of EMD-based denoising techniques are developed1 and tested in different signal scenarios and white Gaussian noise. It is shown that although the main principles between wavelet and EMD thresholding are the same, in the case of EMD, the thresholding operation has to be properly adapted in order to be consistent with the special characteristics of the signal modes resulting from EMD.

II. EMD

Adaptively decomposes a multicomponent signal into a number L of the so-called IMFs $\mathrm{h^i(t)}, 1 \leq \mathrm{i} \leq \mathrm{L}$

$$\mathrm{x(t)} = \sum_{\mathrm{i=1}}^{\mathrm{L}} \mathrm{h^i(t)} + \mathrm{d(t)}$$

Where $\mathrm{d(t)}$ is a reminder that is a non-zero-mean slowly varying function with only few extrema. Each one of the IMFs, say, the ith one$\mathrm{h(t)}$, is estimated with the aid of an iterative process, called sifting, applied to the residual multi component signal

$$\mathrm{x^i(t)} = \begin{cases} \mathrm{x(t)} & i = 1 \\ \mathrm{x(t)} - \sum_{\mathrm{j=1}}^{\mathrm{i-1}} \mathrm{h^i(t)} & \mathrm{i} \geq 2 \end{cases}$$

The sifting process is as follows the (n+1) th sifting iteration, the temporary IMF estimate $\mathrm{h_n}$(t) is improving according to the following steps.

1) Find the local maxima and minima of $\mathrm{h_n}$(t).
2) Interpolate, $\mathrm{h_n}$(t) estimated in the first step in order to form an upper and a lower envelope.
3) Compute the mean of the two envelopes.
4) Obtain the refined estimate $\mathrm{h_{n+1}^{(i)}}$(t)of the IMF by subtracting. $\mathrm{h_n^{(i)}}$(t)
5) Proceed from step 1) again unless a stopping criterion has been fulfilled. Which is actually the corresponding IMF, i.e., $\mathrm{h^i(t)} = \mathrm{x^i(t)} - \mathrm{m^i(t)}$ Each IMF

III. SIGNAL DENOISING

Digital signal denoising can be described as follows. Having a sampled noisy signal x(t) given by$x(t) = \bar{x}(t) + \sigma n(t) \;\; t = 1,2,\ldots,N$ Where $\bar{x}(t)$ is the noiseless signal and n(t) are independent random variables Gaussian distributed N(0,1), produce an estimate $\tilde{x}(t)$) of $\bar{x}(t)$ signal . The novelty of this paper lies in the introduction of new nonparametric thresholding techniques applied to the decomposition modes resulting from EMD instead of the wavelet components. As will be seen, thresholding in EMD is not a straightforward application of the concepts used in wavelet thresholding.

Wavelet Based Denoising : Employing a chosen orthonormal wavelet basis, an orthogonal $N \times N$ Matrix W is the discrete wavelet transform (DWT) $\boldsymbol{c = Wx}$ where $x = [x(1), x(2), x(3), \ldots\ldots x(N)]$ and $c = [c(1), c(2), \ldots\ldots, c(N)]$contains the resultant wavelet coefficients. Using major thresholding operators—hard and soft , the estimated denoised signal is given by $\tilde{x} = W^T \tilde{c}$ where$\widetilde{c} = [\rho_T\ (c_1), \rho_T\ (c_2), \ldots\ldots, \rho_T\ (c_N)]$ and W^TWith respect to the threshold selection, the universal Threshold $T = \sigma\sqrt{2lnN}$.Such a threshold guarantees with high probability

Conventional EMD Denoising : The EMD as a denoising tool emerged from the need to know whether a specific IMF contains useful information or primarily noise. Then, the noise-only IMF energies can be approximated according to $\acute{E}_K = \frac{E_1^2}{\beta} \rho^{-k}$, $k = 2,3,4,\ldots$where E_1s is the energy of the first IMF and depend mainly on the number of sifting iterations used.

IV. IMF THRESHOLDING-BASED DENOISING

In wavelet thresholding a generalized reconstruction of the denoised signal is given by

$$x\acute{(}t) = \sum_{k=M_1}^{M_2} h\acute{^{(i)}} + \sum_{k=M_2+1}^{L} h^{(i)}(t)$$

Where the parameters M1 and M2 gives us flexibility. In the study of thresholds, multiples of the IMF-dependent universal thresholds, i.e. $T_k = C\sqrt{E_k 2lnN}$ where C is a constant, are used.

Thresholding Adapted to EMD Characteristics: This newly developed EMD hard thresholding, as EMD interval thresholding (EMD-IT), translates to

$$\widetilde{}h^{(i)}\left(z_j^{(i)}\right) = \begin{cases} h^i\left(z_j^{(i)}\right), & h^{(i)}\left(r_j^{(i)}\right) > T_i \\ 0, & h^{(i)}\left(r_j^{(i)}\right) \leq T_i \end{cases}$$

For $j = 1,2,..,N_z^{(i)}$,where $h^i\left(z_j^{(i)}\right)$ indicates the samples from instant $z_j^{(i)}$ to $z_{j+1}^{(i)}$ of the i th IMF.

Iterative EMD Interval-Thresholding : This EMD is summarized in the following steps.

1) Perform an EMD expansion of the signal

2) Perform a partial reconstruction using the last $L - 1$ IMFs only,

$$x_p(t) = \sum_{i=2}^{L} h^{(i)}(t)$$

3) Randomly alter the sample positions of the first IMF$h_a^{(1)}(t) = ALTER\ (h^{(1)}(t))$

4) Construct a different noisy version of the original signal $x_a(t) = x_p(t) + h_a^{(1)}(t)$

5) Perform EMD on the new altered noisy signal$x_a(t)$.

6) Perform the EMD-IT denoising on the IMFs $x_a(t)$ of to obtain a denoised version of $\widetilde{x_1}\ (t) of\ x$.

7) Iterate k-1 times between steps 3)–6), where k is the number of averaging iterations in order to obtain k denoised versions of x , i.e $\widetilde{x_{1,}}, \widetilde{x_{2,}}\widetilde{x_{,3,}}, \ldots, \widetilde{x_{K,}}$

8) Average the resulted denoised signals

$$\tilde{x}(t) = (1/K) \sum_{k=1}^{K} \tilde{x}_k\ (t)$$

Clear Iterative EMD Interval Thresholding: When the noise is relatively low, denoising can be achieved with a variant called clear iterative interval-thresholding (EMD-CIIT).EMD-IIT has to be replaced with the following four steps.

1) Perform an EMD expansion of the original noisy signal.

2) Perform a thresholding operation to the first IMF of $x(t)$) to obtain a denoised version of $\widetilde{h^1}(t)$ of $h^{(1)}(t)$

3. Compute the actual noise signal that existed in $h^1(t)$,$h^{(1)}(t) = \widetilde{h^1}(t) - h^1(t)$.

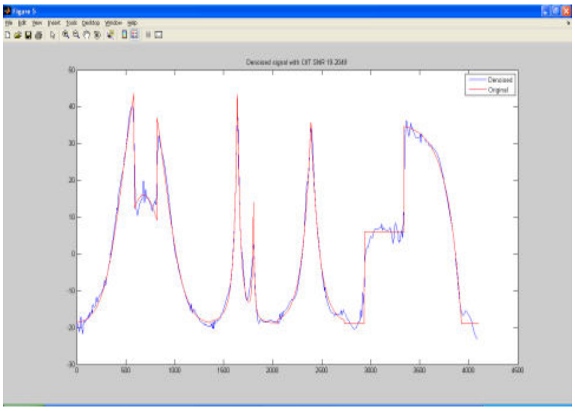
4) Perform a partial reconstruction using the last L-1 IMFs plus the information signal contained in the first IMF

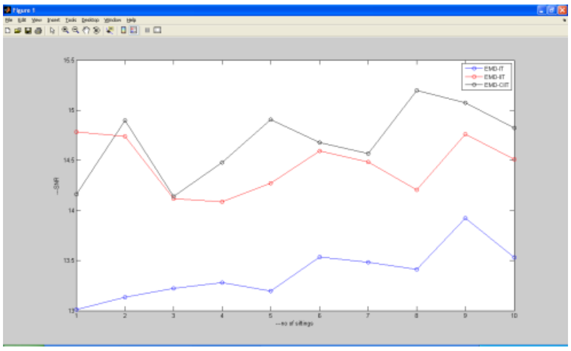
$$x_p(t) = \textstyle\sum_{i=2}^{L}\ h^{(i)}(t) + \widetilde{h^1}(t).$$

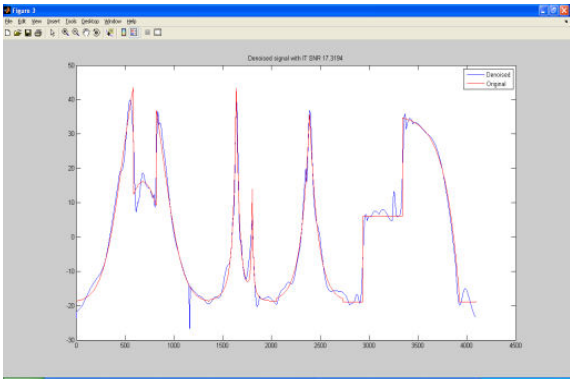
5) Randomly alter the sample positions of the noise-only part of the first IMF

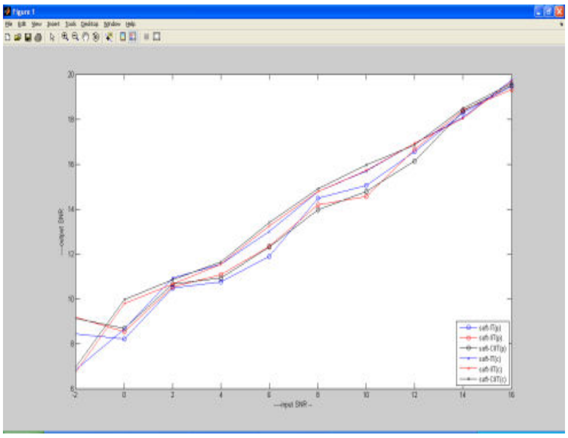
$$h_a^{(1)}(t) = ALTER\ (h_n^{(1)}(t)$$

V. RESULTS









The effect SNR performance adapting IMF with respect to the sifting iterations is studied in the above figure by taking various signals. According to irregularities e.g., the piece-regular signal case, the best performance is achieved with a relatively low number of sifting iterations. These results have been evaluated with other regular and irregular signals. The balanced tradeoff between number of sifting and performance is realized with about eight sifting iterations. Secondly, it is apparent that the sifting-dependent IMF curves do not offer significant advantages over fixed curves since the performance difference never exceeds 0.2 dB. In addition, the sifting-dependent curves can even lead to slight performance deterioration in the case of EMD-CIIT when the signal has intense irregularities and a small number of sifting iterations are used.

VI. CONCLUSION

In this paper, the principles of wavelet thresholding were appropriately modified to develop denoising methods suited for thresholding EMD modes. Presented denoising in the cases when the signal SNR is low and/or the sampling frequency is high and enhanced performance compared to wavelet. These results suggest further efforts for improvement of EMD based denoising when denoising the signals with moderate to high SNR

REFERENCE

[1] N. E. Huang et al., “The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis,” Proc.

[2] G. Rilling and P. Flandrin, “One or two frequencies? The empiricalmode decomposition answers,”

[3] Y. Kopsinis and S. McLauglin, “Investigation and performance enhancement of the empirical mode decomposition method based on a heuristic search optimization approach,”

[4] S. Mallat, A Wavelet Tour of Signal Processing, 2nd ed.

[5] A. Antoniadis and J. Bigot, “Wavelet estimators in nonparametric regression:
A comparative simulation study,”

◈◈◈

Evaluation of Probabilistic Broadcasting Method in Mobile Ad Hoc Network in Different Scenario

Sukant Kishoro Bisoyi, Mohit Ranjan Panda & Manas Kumar Swain

C. V. Raman College of Engineering, Bhubaneswar, India
E-mail: sukantabisoyi@yahoo.com, mohit1146@gmail.com , mkswain2004@yahoo.co.in

Abstract - A Mobile Ad-Hoc Network (MANET) is a collection of wireless mobile nodes forming a temporary network without using any centralized access point, infrastructure, or centralized administration. Broadcasting is one of the most fundamental operations in mobile ad hoc networks. Broadcasting is the process in which a source node sends a message to all other nodes in MANET. However, broadcasting induces what is known as the "broadcast storm problem" which causes severe degradation in network performance due to excessive redundant retransmission, collision, and contention. Traditional implementation of flooding suffers from the problems of excessive redundancy of messages, resource contention, and signal collision. This causes high protocol overhead and interference to the existing traffic in the networks. Probabilistic broadcast has been widely used as a flooding optimization mechanism to alleviate the effect of broadcast storm problem (BSP) in mobile ad hoc networks (MANETs). In the paper, we have investigated the performance probabilistic flooding using NS2 simulator with a regular grid and random network. We found probabilistic approach to flooding is a means of reducing redundant rebroadcasts and alleviating the broadcast storm problem.

Key words *- MANET, Broadcasting, Flooding, Probability, NS2.*

I. INTRODUCTION

MANET is an autonomous system of mobile routers (and associated hosts) connected by wireless links - the union of which forms an arbitrary graph. The routers are free to move randomly and organize themselves arbitrarily; thus, the network's wireless topology may change rapidly and unpredictably. Such a network may operate in a standalone fashion, or may be connected to the larger Internet. Most of the major routing protocols, such as DSR [1], AODV [2], etc., rely on flooding for disseminating route discovery, route maintenance, or topology update packets. MANET protocol involves unicast, multicast and broadcast communications. In a multi hop scenario the packets originated from the source host are relayed by several intermediate hosts before reaching the destination. Broadcasting is a fundamental operation in MANETs whereby a source node transmits a message that is to be disseminated to all the nodes in the network. However, broadcasting induces what is known as the "broadcast storm problem" which causes severe degradation in network performance due to excessive redundant retransmission, collision, and contention. To solve the broadcast storm problem, several schemes have been proposed to reduce the redundancy in flooding operations. Paper [3], [4] describes the notable work of flooding. However, these algorithms either perform poorly in reducing redundant transmissions. The broadcast mechanism is used for data, transmission of large amount data or stream media which requires a broadcast routing to find an efficient route before the actual transmission of data, so that data can be transmitted efficiently along the pre-found route. In contrast, flooding is usually used for dissemination of control packets, which is a one-off operation. It does not need routing before hand.

II. BROADCASTING IN MANET

The broadcast mechanism is used for data, transmission of large amount data or stream media which requires a broadcast routing to find an efficient route before the actual transmission of data, so that data can be transmitted efficiently along the pre-found route. Broadcast method follows *one-to-all* model, where transmission by each node can reach all nodes that are within its transmission radius, while in the *one-to-one* model, each transmission is directed towards only one neighbour using narrow beam directional antennas or separate frequencies for each node [5]. In MANET, a host may not communicate with another directly indirectly. So a multi hop scenario occurs, where the packets originated from the source host are relayed by several intermediate hosts before reaching the destination. The broadcast problem refers to the sending of a message to other hosts in the network. The problem considered here has the following characteristics. In a broadcast process, each node decides its forwarding status based on given neighborhood information and the corresponding broadcast protocol. Most existing broadcast schemes assume either the underlying network topology is static during the broadcast process such that the neighborhood information can be updated in a timely manner.

Broadcasting is a common operation in many applications, e.g., graph-related problems and distributed computing problems. It is also widely used to resolve many network layer problems. In a MANET in particular, due to host mobility, broadcastings are expected to be performed more frequently (e.g., for paging a particular host, sending an alarm signal, and finding a route to a particular Host. Broadcasting may also be used in LAN emulation [6] or serve as a last resort to provide multicast services in networks with rapid changing topologies. Problem with the broadcast message is :

The broadcast is spontaneous - Any obile host can issue a broadcast operation at any time.

The broadcast is unreliable- No acknowledgement mechanism will be used.2 However, attempt should be made to distribute a broadcast message to as many hosts as possible without paying too much effort.

One of the earliest broadcast mechanisms proposed in the literature is simple or “blind” flooding [7] where each node receives and then re-transmits the message to all its neighbours. The only ‘optimization’ applied to this technique is that nodes remember broadcast messages received and do not act if they receive repeated copies of the same message. However, a straightforward flooding of the network with broadcast messages is usually costly and results in serious redundancy and collisions in the network; such a scenario has often been referred to as the *broadcast storm problem* [5, 8], and has generated many challenging research issues.

III. BROADCASTING METHODS

Broadcasting methods have been categorized into four types utilizing the IEEE 802.11 MAC specifications [9].

- Simple flooding
- Probabilirty based
- Area based
- Neighbourhood based

(1) Simple flooding method

In simple flooding, in which each node retransmits the received message when it receives it for the first time starting at the source node. This process continues until all reachable nodes have received and retransmit the broadcast message. This simple scheme guarantees that a flooding message can reach all nodes if there is no collision and the network is connected. However, it generates excessive amount of redundant network traffic, because all nodes in the network transmit the flooding message. This will consume a lot of energy resource of mobile nodes and cause the congestion of the network. Furthermore, due to the broadcast nature of radio transmissions, there is a very high probability of signal collisions when all nodes flood the message in the network at the same time, which would cause more re-transmissions or some nodes failing to receive the message. Figure 1 provides a brief outline of this scheme.

Algorithm 1: Flooding (m)

```
1. On receiving a broadcast message m at node
   P do the following:
2. If message m received for the first time Then
   broadcast (m) {this is the basic local
   broadcast primitive to nodes within range
   only}
3. End if
```

Figure 1: Simple flooding algorithm for broadcasting in MANETs.

Advantages:

(i) Simple

(ii) Reliable.

Disadvantages:-

(i) It costs **n** transmissions in a network of **n** reachable nodes.

(ii) **Redundant Retransmissions:** When a mobile node decides to retransmit a broadcast message to its neighbors, all its neighbors already have the message.

(iii) **Contention**: After a mobile node retransmits a message, if many of its neighbors decide to retransmit the message, these transmissions (which are all from nearby nodes) may severely contend with each other.

(iv) **Collision**: Because of the deficiency of back-off mechanism, the lack of request to send/clear to send (RTS/CTS) dialogue, and the absence of collision detection, collisions are more likely to occur and cause more damage.

(2) Probability based approach

The probabilistic scheme [10] is widely-used for flooding optimization during route discovery in

MANETs and aim at reducing redundancy through rebroadcast timing control in an attempt to alleviate the broadcast storm problem (BSP). In this scheme, when receiving a broadcast message for the first time, a node rebroadcasts the message with a pre-determined probability p and with probability $(1-p)$ it discards the packet, so that every node has the same probability to rebroadcast the message, regardless of its number of neighbors. In dense networks, multiple nodes share similar transmission range. Therefore, these probabilities control the frequency of rebroadcasts and thus could save network resources without affecting delivery ratios. In flooding, a mobile node rebroadcasts all routing request packets that are received for the first time. Therefore, there are N-1 possible rebroadcasts, where N is the total number of nodes. In general probabilistic approach, each node decides to rebroadcast or not according to a fixed probability P. Since their decisions are independent, the total number of rebroadcasts is P*(N-1) on average. Figure 2 provides brief of this scheme.

Probabilistic Flooding (m, p)

```
1. On receiving a broadcast message m at node P
   do the following:
2. If message m received for the first time Then
   broadcast (m)with probability p {this is the
   basic local broadcast primitive to nodes within
   range only}
3. End if
```

Figure 2: Probabilistic flooding algorithm for broadcasting in MANETs.

(3) Area based methods

Area based flooding requires a node to evaluate the additional area covered by its rebroadcasting. If this additional coverage is less than some threshold value, the node will give up its rebroadcasting. A node using an Area Based Method can evaluate additional coverage area based on all received redundant transmissions. We note that area based methods only consider the coverage area of a transmission; they don't consider whether nodes exist within that area.

(4) Neighbor Knowledge method

The simplest of the Neighbor Knowledge Methods is what Lim and Kim refer to as Flooding with Self Pruning [11]. This protocol requires that each node have knowledge of its 1-hop neighbors, which is obtained via periodic "Hello" packets. A node includes its list of known neighbors in the header of each broadcast packet. A node receiving a broadcast packet compares its neighbor list to the sender's neighbor list. If the receiving node would not reach any additional nodes, it refrains from rebroadcasting; otherwise the node rebroadcasts the packet.

IV. SIMULATION SCENARIO AND RESULTS

Simulation experiment of probabilistic flooding algorithm is carried out using the ns-2 simulator [12] in two different scenarios. We adopted the "Random way-point" model to simulate nodes movement. In scenario 1 and scenario 2, we simulate the algorithm in regular grid network and random network respectively. Parameters used in scenarios are shown in table 1. We ran the simulation for probability values: 0.1, 0.3, 0.5, 0.7, 0.8, 0.9, and 1.0. When probability=1.0, it is equivalent to complete flooding. Then following things were observed after simulation: Received messages per node, total broadcasted messages, fraction of nodes that receive message at least once (for 100 and 36 nodes only), Success Rate.

From the figure 3 it can be concluded that average number of message received per node increases for lesser number of nodes. Figure 4 show that total number of broadcast message received increases with increase of number of nodes. When probability increases total number of broadcast message received also increases. The success rate is defined as the average no of received messages per node by total broadcast message. In our simulation we found that success rate of the algorithm decreases with increase of probability (shown in figure 5). But success rate is high if the number of node is less. It is due to the increase in the number of broadcast messages as number of node increases. In the figure 6, with the increase in probability the fractions of nodes that receive the messages at least once increases initially but later on it remain constant. In the random network as shown in figure 7, average received messages per node increases with increase in the probability, however as the grid size increases the average received messages per node decreases. In the figure 9 it shows that the total number of broadcasted messages increases with the increase in probability. But total broadcast message increases with large number of nodes as compare to small number of nodes. The success rate is high for less number of node in regular network (in figure 10). But unlike regular grid network in random network (shown in figure 11) average number of message received per node increases with increase of grid size but decrease with increase of probability.

Table 1: parameter used in simulation

PARAMETER	VALUE
Channel type	Wireless channel
Number of nodes	36, 100, 225, 400
Routing protocol	AODV
Grid Area	500 X 500 and 1000 X 1000 sq meter
MAC Types	802_11
Node Placement	Random
Mobility model	Random way point
Antenna model	Omni
Size of interface queue	50
Time of simulation	10 msec.
Area of simulation	500*500

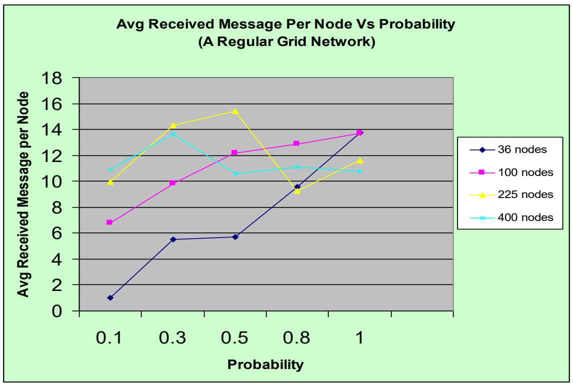


Fig. 3: Average message received Vs Probability

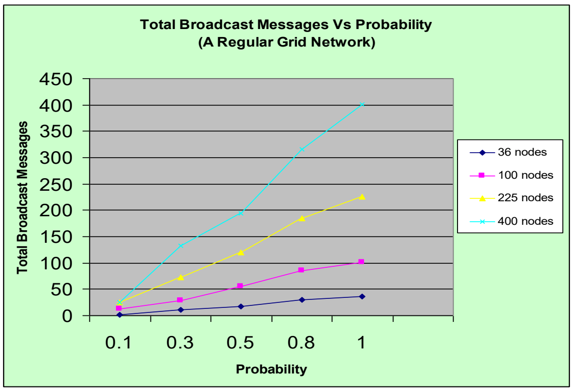


Fig. 4: Total Broadcast message received Vs Probability

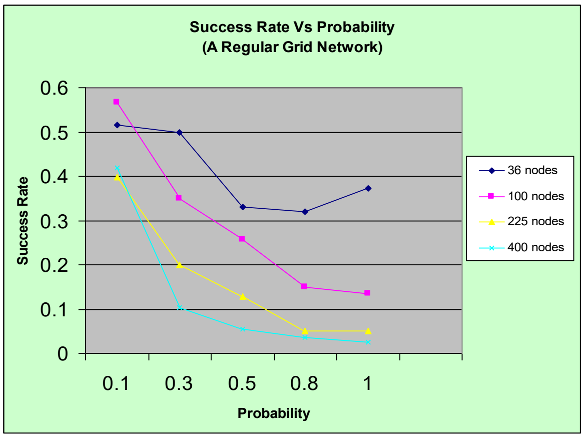


Fig. 5: Success rate Vs Probability

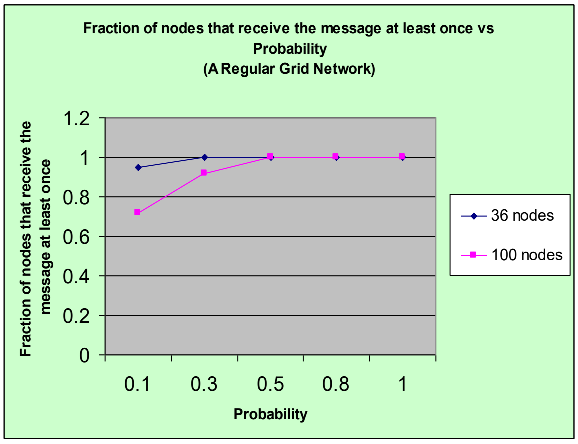


Fig. 6: Fraction of node that receive the message at once Vs Probability

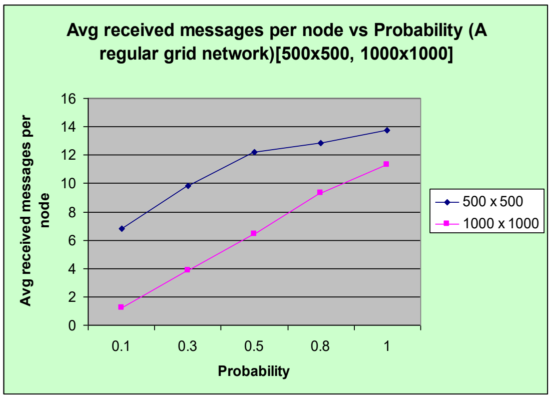


Fig. 7: Average message received per node Vs Probability

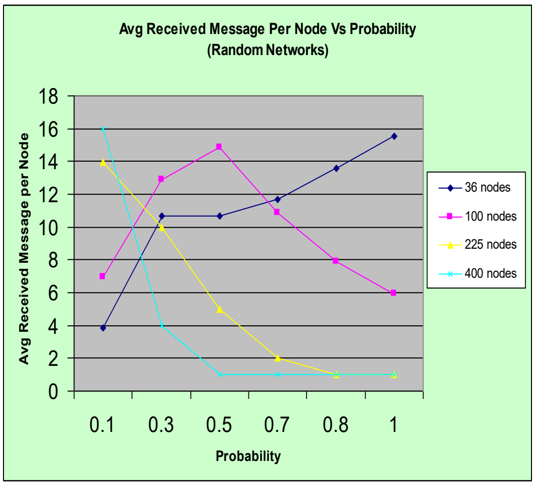


Fig. 8: Average message received Vs Probability

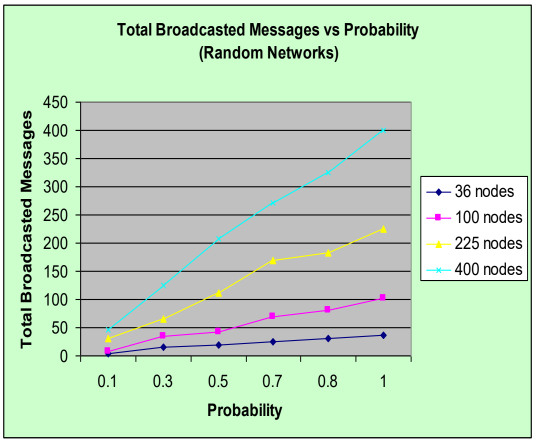


Fig. 9: Total Broadcast message received Vs Probability

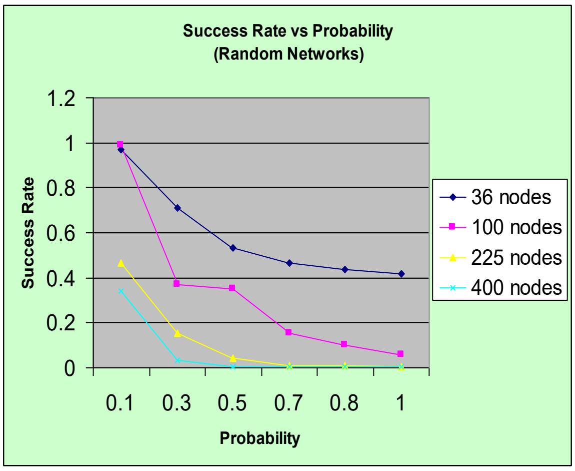


Fig. 10: Success rate Vs Probability

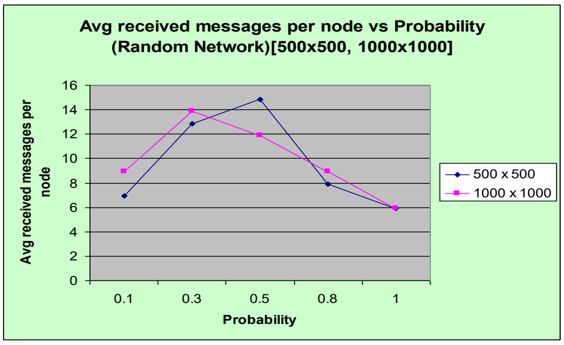


Fig. 11: Average message received per node Vs Probability

VI. CONCLUSION

Broadcasting is an active research topic in MANETs. An important problem is how to minimize the number of rebroadcast packets while good retransmission latency and packets reachability are maintained. This paper has evaluated and analyzed the performance of probabilistic flooding on the AODV protocol which is based on simple flooding in MANETs.

From the experiment we find that the average broadcast messages increases with the increase in probability for both regular grid network and random networks. Because of increase in the number of broadcast messages success rate gradually decreases. Therefore probabilistic approach to flooding is a means of reducing redundant rebroadcasts and alleviating the broadcast storm problem.

REFERENCES

[1] D. Johnson and D. A. Maltz, “Dynamic Source Routing in Ad Hoc Wireless Networks,” in Mobile Computing, T. Imielinski and H. F. Korth, Eds., pp. 153–181. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1996.

[2] C.E. Perkins, “Ad Hoc On-Demand Distance Vector (AODV) Routing,” INTERNET DRAFT - Mobile Ad hoc NETworking (MONET) Working group of the Internet Engineering Task Force (IETF), Nov. 1997.

[3] Y. Cai, K.A. Hua, and A. Phillips, “Leveraging 1-hop Neighborhood Knowledge for Efficient Flooding in Wireless Ad Hoc Networks,” 24th IEEE International Performance Computing and Communications Conference (IPCCC), Apr. 7-9, 2005.

[4] C.C. Yang and C.Y. Chen, “A Reachability-Guaranteed Approach for Reducing the Broadcast

Storms in MANETs," Proceedings of IEEE Semiannual Vehicular Technology Conference (VTC-2002 Fall), Sep. 2002

[5] B. Williams, T. Camp, Comparison of broadcasting techniques for mobile ad hoc networks. Proc. ACM Symposium on Mobile Ad Hoc Networking & Computing(MOBIHOC 2002), pp. 194–205, 2002.

[6] R. Bar-Yehuda, 0. Goldreich, and A. Itai. Efficient emulation of single-hop radio network with collision detection on multi-hop radio network with no collision detection. Distributed Computing, 5(2):67-72, 1991.

[7] S.-Y. Ni, Y.-C. Tseng, Y.-S. Chen, J.-P. Sheu, The broadcast storm problem in a mobile ad hoc network, Proc. Mobicom'99, 1999.

[8] Y.-C. Tseng, S.-Y. Ni, E.-Y. Shih, Adaptive approaches to relieving broadcast storm in a wireless multihop mobile ad hoc network, IEEE Transactions Computers, vol. 52, no 5, 2003.

[9] I.S.Committee.Wireless LAN Medium ACCESS CONTROL (MAC) and Physical Layer Specifications. IEEE 802.11 Standards. IEEE, New York, ISBN 1559379359, 1997.

[10] Y. Sasson, D. Cavin, A. Schiper, Probabilistic broadcast for flooding in wireless mobile ad hoc networks, Proc. IEEE Wireless Communications & Networking Conference (WCNC 2003), pp. 1124-1130, March 2003.

[11] H. Lim and C. Kim. Multicast tree construction and flooding in wireless ad hoc networks. In Proceedings of the ACM International Workshop on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWIM), 2000.

[12] NS-2, The ns Manual (formally known as NS Documentation) available at http: //www. isi.edu/nsnam/ns/doc.

◈◈◈