IJNS
International Journal of Network Security

ISSN 1816-353X (Print)
ISSN 1816-3548 (Online)

Vol. 20, No. 3 (May 2018)

INTERNATIONAL JOURNAL OF NETWORK SECURITY

Editor-in-Chief

Prof. Min-Shiang Hwang
Department of Computer Science & Information Engineering, Asia University, Taiwan

Co-Editor-in-Chief:

Prof. Chin-Chen Chang (IEEE Fellow)
Department of Information Engineering and Computer Science, Feng Chia University, Taiwan

Publishing Editors

Shu-Fen Chiou, Chia-Chun Wu, Cheng-Yi Yang

Board of Editors

Ajith Abraham
School of Computer Science and Engineering, Chung-Ang University (Korea)

Wael Adi
Institute for Computer and Communication Network Engineering, Technical University of Braunschweig (Germany)

Sheikh Iqbal Ahamed
Department of Math., Stat. and Computer Sc. Marquette University, Milwaukee (USA)

Vijay Atluri
MSIS Department Research Director, CIMIC Rutgers University (USA)

Mauro Barni
Dipartimento di Ingegneria dell'Informazione, Università di Siena (Italy)

Andrew Blyth
Information Security Research Group, School of Computing, University of Glamorgan (UK)

Soon Ae Chun
College of Staten Island, City University of New York, Staten Island, NY (USA)

Stefanos Gritzalis
University of the Aegean (Greece)

Lakhmi Jain
School of Electrical and Information Engineering, University of South Australia (Australia)

James B D Joshi
Dept. of Information Science and Telecommunications, University of Pittsburgh (USA)

Ç etin Kaya Koç
School of EECS, Oregon State University (USA)

Shahram Latifi
Department of Electrical and Computer Engineering, University of Nevada, Las Vegas (USA)

Cheng-Chi Lee
Department of Library and Information Science, Fu Jen Catholic University (Taiwan)

Chun-Ta Li
Department of Information Management, Tainan University of Technology (Taiwan)

Iuon-Chang Lin
Department of Management of Information Systems, National Chung Hsing University (Taiwan)

John C.S. Lui
Department of Computer Science & Engineering, Chinese University of Hong Kong (Hong Kong)

Kia Makki
Telecommunications and Information Technology Institute, College of Engineering, Florida International University (USA)

Gregorio Martinez
University of Murcia (UMU) (Spain)

Sabah M.A. Mohammed
Department of Computer Science, Lakehead University (Canada)

Lakshmi Narasimhan
School of Electrical Engineering and Computer Science, University of Newcastle (Australia)

Khaled E. A. Negm
Etisalat University College (United Arab Emirates)

Joon S. Park
School of Information Studies, Syracuse University (USA)

Antonio Pescapè
University of Napoli "Federico II" (Italy)

Zuhua Shao
Department of Computer and Electronic Engineering, Zhejiang University of Science and Technology (China)

Mukesh Singhal
Department of Computer Science, University of Kentucky (USA)

Nicolas Sklavos
Informatics & MM Department, Technological Educational Institute of Patras, Hellas (Greece)

Tony Thomas
School of Computer Engineering, Nanyang Technological University (Singapore)

Mohsen Toorani
Department of Informatics, University of Bergen (Norway)

Shuozhong Wang
School of Communication and Information Engineering, Shanghai University (China)

Zhi-Hui Wang
School of Software, Dalian University of Technology (China)

Chuan-Kun Wu
Chinese Academy of Sciences (P.R. China) and Department of Computer Science, National Australian University (Australia)

Chou-Chen Yang
Department of Management of Information Systems, National Chung Hsing University (Taiwan)

Sherali Zeadally
Department of Computer Science and Information Technology, University of the District of Columbia, USA

Jianping Zeng
School of Computer Science, Fudan University (China)

Justin Zhan
School of Information Technology & Engineering, University of Ottawa (Canada)

Mingwu Zhang
College of Information, South China Agric University (China)

Yan Zhang
Wireless Communications Laboratory, NICT (Singapore)

PUBLISHING OFFICE

Min-Shiang Hwang
Department of Computer Science & Information Engineering, Asia University, Taichung 41354, Taiwan, R.O.C.
Email: mshwang@asia.edu.tw

International Journal of Network Security is published both in traditional paper form (ISSN 1816-353X) and in Internet (ISSN 1816-3548) at http://ijns.jalaxy.com.tw

PUBLISHER: Candy C. H. Lin

© Jalaxy Technology Co., Ltd., Taiwan 2005
23-75, P.O. Box, Taichung, Taiwan 40199, R.O.C.

International Journal of Network Security

Vol. 20, No. 3 (May 1, 2018)

Multi-proxy Multi-signature without Pairing from Certificateless Cryptography

Rena Ehmet[1,2], Lunzhi Deng[3,4,5], Yingying Zhang[1], Jiwen Zeng[1]
(Corresponding author: Lunzhi Deng)

School of Mathematical Sciences, Xiamen University, China[1]
School of Mathematical Sciences, Xinjiang Normal University, China[2]
School of Mathematical Sciences, Guizhou Normal University, China[3]
School of Computer Science, Guizhou University, China[4]
Guizhou University, Guizhou Provincial Key Laboratory of Public Big Data, China[5]
(Email: denglunzhi@163.com)
(Received Oct. 11, 2016; revised and accepted Feb. 20, 2017)

Abstract

In a multi-proxy multi-signature scheme, there is a group of original signers who delegate their signing rights to another group of persons called proxy group. Most of the known cryptography schemes used bilinear pairings, the computation cost of the which is much higher than that of the exponentiation in a RSA group. In this paper, we propose a certificateless multi-proxy multi-signature scheme based on the classic RSA and discrete logarithm (DL) problem. Our scheme is also constructed without using pairing which reduces the running time significantly, and it is secure against chosen message attack in random oracle model and more applicable for practical applications.

Keywords: Certificateless Cryptography; Multi-proxy; Multi-signature; RSA

1 Introduction

In traditional public key cryptography, the users first need to obtain the authenticated public keys from a certificate authority, if they want to communicate a message. In that system, the certificate management, storage space and large overhead to transfer certificates lead to increase the associated communication cost.

ID-Based Cryptography. To solve the certificate management problem in the traditional public key cryptography, Shamir [16] introduced the ID-based cryptography in 1984, which removed the need of certificate for public key and thus reduced the associated communication cost. In ID-based cryptography, the users' public and private keys are generated from their identities such as email addresses, IP addresses, etc. There is a very important problem in ID-based cryptosystem that user's private key is generated by a key generation center (KGC). It means that KGC knows user's private key. So ID-based public key cryptography has to face with key escrow problem.

Certificateless Cryptography. In 2003, Al-Riyami *et al.* [1] proposed the concept of certificateless public key cryptosystem (CLPKC). In CLPKC, a user's private key is comprised of partial private key generated by KGC and a secret value chosen by the user separately. The certificateless public key cryptography has attracted much attention since it solves the certificate management problem in the traditional public cryptography and the key escrow problem in the ID-based cryptography.

Proxy Signatures. In 1996, Mambo *et al.* proposed the first proxy signature scheme [13], which allows an entity, called original signer, to delegate his/her signing right to other entity, called proxy signer. The multi-proxy signature scheme allows a group of proxy signers generate signatures, on behalf of one original signer — a company or an organization, who delegates his/her signing right to the proxy group [9, 11, 19]. Multi-proxy multi-signature (MPMS) is a new kind of proxy signature, firstly proposed by Tzeng *et al.* [20] in 2004, in which a group of original signers can authorize a group of proxy signers under the agreement of all original and proxy signers, so that a signature could be generated by the cooperation of all proxy signers. It solves many real life problems. For example in a company, there are some conflict between the employees and the employers. All employees want to depute a lawyer group as their proxy agents.

Cryptography from RSA. RSA is one of the first practical asymmetric public-key cryptosystems and widely used for secure data transmission. In RSA cryptosystem, its asymmetry is based on the practical difficulty of factoring the product of two large prime numbers. Being as a classified difficult problem, RSA is widely used in many aspects of cryptography. Shamir [16] constructed the first ID-based signature scheme from RSA in 1985.

Herranz [7] proposed an ID-based ring signature scheme whose security is based on the hardness of RSA problem.

Using bilinear pairings, people proposed many new proxy signature scheme [2, 10, 12, 14, 15, 21, 22, 23]. All the above schemes are very practical, but they used bilinear pairings and the pairing is regarded as the most expensive cryptography primitive. In 2011, He *et al.* [5] proposed an ID-based proxy signature scheme without bilinear pairing. In 2013, He *et al.* [6] put forward a certificateless proxy signature scheme without bilinear pairing. Kim *et al.* [8] constructed a provably secure ID-based proxy signature scheme based on the lattice problems. In 2014, Deng *et al.* [4] constructed a certificateless proxy signature based on RSA and discrete logarithm problem. In 2015, Tiwari and Padhye [17] proposed a provable secure multi-proxy signature scheme without bilinear map. In 2017, Deng *et al.* [3] put forward an ID-based proxy signature from RSA without bilinear pairing. The computation cost of the pairing is much higher than that of the exponentiation in a RSA group. Therefore, certificateless schemes based on RSA primitive would still be appealing.

Our Contribution. By using the idea from [4], we propose a new certificateless multi-proxy multi-signature (CLMPMS) scheme. Security of the scheme is based on the famous RSA problem and DL (discrete-logarithm) problem. And our scheme is efficient in reducing the running time significantly because of its pairing-freeness. In addition, we analyze the security of our scheme against both of the super Type I and the super Type II adversaries. To the best of authors' knowledge, our scheme is the first certificateless multi-proxy multi-signature based on RSA and DL problem.

Roadmap. The organization of the paper is sketched as follows: The Section 2 gives some preliminaries and the formal model. We present our proposed scheme in Section 3. The security analysis and performance comparisons will be given in Sections 4 and 5 separately. Finally, we give some conclusions in Section 6.

2 Preliminaries

2.1 Elliptic Curve Group

Let E/F_p denote an elliptic curve E over a prime finite field F_p, defined by an equation

$$y^2 = x^3 + ax + b(\mathrm{mod}p),\ a, b \in F_p \text{ and } 4a^3 + 27b^2 \neq 0(\mathrm{mod}p),$$

The points on E/F_p together with an extra point O called the point at infinity form a group

$$\mathcal{G} = \{(x, y) : x, y \in F_p, E(x, y) = 0\} \cup \{O\}.$$

2.2 Notations

- N: A large composite number, the product of two prime numbers p, q.
- $\mathbb{G}$: A cyclic subgroup of $\mathcal{G}$ with prime order b and $gcd(b, \varphi(N)) = 1$.
- P: A generator of group $\mathbb{G}$.
- D_i: The partial private key of user ID_i generated by KGC.
- t_i: The secret value chosen by user ID_i.
- P_i: The public key of user ID_i.
- RSAP: Given a tuple $(N, b, \mathcal{Y})$ to find $\mathcal{X} \in \mathbb{Z}_N^*$ such that $\mathcal{X}^b = \mathcal{Y} \bmod N$.
- DLP: Given a tuple (P, xP) in $\mathbb{G}$ to compute $x \in Z_b^*$.

2.3 System Model

The proposed model involves four parties: a set of n original signers $\mathcal{N} = \{ID_{o1}, ID_{o2}, \cdots, ID_{on}\}$, a set of l proxy signers $\mathcal{L} = \{ID_{p1}, ID_{p2}, \cdots, ID_{pl}\}$, a verifier $\mathcal{V}$, and a clerk $\mathcal{B}$. Use of clerk reduces the communication cost.

Definition 1. *A multi-proxy multi-signature scheme is specified by the following polynomial time algorithms.*

Setup. Given a security parameter k, this algorithm outputs the system parameters *params*, and keeps msk as system's master secret key.

Partial private key extract. Given an identity $ID_i \in \{0, 1\}^*$, the master secret key msk, and parameters *params*, the key generation center (KGC) generates the partial private key D_i for the identity ID_i.

Set secret value. The user with identity ID_i chooses a random number as his secret value.

User's public key generation. The user with identity ID_i computes his public key.

Proxy certificate generation. This algorithm takes the warrant m_ω to be signed and generates the proxy certificate with the cooperation of all original signers and proxy signers.

Multi-proxy sign. This algorithm takes the certificate and a message $M \in \{0, 1\}^*$ as input, and outputs a multi-proxy multi-signature signed by the proxy group $\mathcal{L}$ on behalf of the original group $\mathcal{N}$.

Verify. This algorithm takes the identities of all original signers, the identities of all proxy signers, and a proxy signature as input, returns *True* if it is a valid signature on M signed by the proxy group $\mathcal{L}$, on behalf of the original group $\mathcal{N}$. Otherwise, returns $False$.

2.4 Security Model

For a certificateless multi-proxy multi-signature scheme, there are two kinds of adversaries. The adversary $\mathcal{A}_1$ is not able to access the master key, but he could replace users' public keys at his will. The adversary $\mathcal{A}_2$ can access the master key, but he is not able to replace users' public keys. The security of CLMPMS schemes are formally defined through two games played between a challenger $\mathcal{C}$ and an adversary $\mathcal{A} \in \{\mathcal{A}_1,\ \mathcal{A}_2\}$.

Definition 2. *A CLMPMS scheme is unforgeable if no polynomially bounded adversary has a non-negligible advantage in the following two games against Type I and Type II adversaries.*

Game I: It performs between the challenger $\mathcal{C}$ and a Type I adversary $\mathcal{A}_1$ for the CLMPMS scheme.

Initialization: $\mathcal{C}$ runs the setup algorithm, takes a security parameter k as input to obtain a master key msk and the system parameters $params$. $\mathcal{C}$ then sends $params$ to the adversary $\mathcal{A}_1$ and keeps msk secret. The point is that adversary $\mathcal{A}_1$ doesn't know msk.

Queries: $\mathcal{A}_1$ can get access to query the following oracles polynomially bounded number of times which are controlled by $\mathcal{C}$. Each query may depend on the answers to the previous query.

- User-Public-Key-Oracle: This oracle takes a user's identity ID_i as input. If ID_i's public key has already been queried, nothing is to be carried out. Otherwise, it generates the secret value t_i and the public key P_i. Then it returns P_i and adds (ID_i, D_i, t_i, P_i) to the list L_U.
- Partial-Private-Key-Oracle: On inputting an identity ID_i, the oracle browses the list L_U and returns the partial private key D_i as answer. Otherwise, returns 0.
- Public-Key-Replacement-Oracle: Taking an identity ID_i and a new public key P_i' as input, the oracle replaces the public key of the given identity ID_i with new one and updates the corresponding information in the list L_U.
- Secret-Value-Oracle: On inputting a created identity ID_i, the oracle browses the list L_U and returns the secret value t_i as answer. Otherwise returns 0. Note that t_i is the secret value associated with the original public key P_i. $\mathcal{A}_1$ can't query the secret value for ID_i whose public key has been replaced.
- Proxy-Certificate-Generation-Oracle: When $\mathcal{A}_1$ submits all signers' identities/public keys $(ID_i, P_i),\ ID_i \in \mathcal{N} \cup \mathcal{L}$ and a warrant m_ω to the challenger, $\mathcal{C}$ responds by running the proxy certificate generation algorithm on the warrant m_ω and the signers' full private keys $(t_i, D_i),\ ID_i \in \mathcal{N} \cup \mathcal{L}$.
- Proxy-Sign-Oracle: When $\mathcal{A}_1$ submits certificate π and a message M to the challenger, $\mathcal{C}$ responds by running the proxy sign algorithm on π, M and the proxy signers' full private keys $(t_i, D_i),\ ID_i \in \mathcal{L}$.

Forge: $\mathcal{A}_1$ outputs a tuple $(\pi^*, \mathcal{N} \cup \mathcal{L}, \bigcup_{ID_i \in \mathcal{N} \cup \mathcal{L}} P_i)$ or $(M^*, m_\omega^*, \sigma^*, \mathcal{N} \cup \mathcal{L}, \bigcup_{ID_i \in \mathcal{N} \cup \mathcal{L}} P_i)$. We say $\mathcal{A}_1$ wins the game, if one of the following cases is satisfied:

Case 1: $\mathcal{A}_1$ outputs a tuple $(\pi^*, \mathcal{N} \cup \mathcal{L}, \bigcup_{ID_i \in \mathcal{N} \cup \mathcal{L}} P_i)$ satisfying:

1) π^* is a valid certificate.
2) π^* is not generated from the certificate generation query.
3) There is at least one user $ID \in \mathcal{N} \cup \mathcal{L}$ whose partial private key is not queried by $\mathcal{A}_1$.

Case 2: $\mathcal{A}_1$ outputs a tuple $(M^*, m_\omega^*, \sigma^*, \mathcal{N} \cup \mathcal{L}, \bigcup_{ID_i \in \mathcal{N} \cup \mathcal{L}} P_i)$ satisfying:

1) σ^* is a valid proxy signature.
2) σ^* is not generated from the proxy signature query.
3) $(m_\omega^*, \mathcal{N} \cup \mathcal{L}, \bigcup_{ID_i \in \mathcal{N} \cup \mathcal{L}} P_i)$ didn't appear in the certificate generation query.
4) There is at least one user $ID \in \mathcal{N}$ whose partial private key is not queried by $\mathcal{A}_1$.

Case 3: $\mathcal{A}_1$ outputs a tuple $(M^*, m_\omega^*, \sigma^*, \mathcal{N} \cup \mathcal{L}, \bigcup_{ID_i \in \mathcal{N} \cup \mathcal{L}} P_i)$ satisfying:

1) σ^* is a valid proxy signature.
2) σ^* is not generated from the proxy signature query.
3) There is at least one user $ID \in \mathcal{L}$ whose partial private key is not queried by $\mathcal{A}_1$.

The advantage of $\mathcal{A}_1$ is defined as

$$Adv_{\mathcal{A}_1}^{UNF-CLMPMS} = Pr[\mathcal{A}_1\ wins].$$

Game II: It performs between the challenger $\mathcal{C}$ and a Type II adversary $\mathcal{A}_2$ for the CLMPMS scheme.

Initialization. $\mathcal{C}$ runs the setup algorithm, takes a security parameter k as input to obtain a master key msk and the system parameters $params$. $\mathcal{C}$ then sends msk, $params$ to the adversary $\mathcal{A}_2$. It means that in Game II adversary $\mathcal{A}_2$ knows msk, he/she just can't replace the public key.

Queries. $\mathcal{A}_2$ may adaptively make a polynomially bounded number of queries as in Game I.

Forge. $\mathcal{A}_2$ outputs a tuple $(\pi^*, \mathcal{N} \cup \mathcal{L}, \bigcup_{ID_i \in \mathcal{N} \cup \mathcal{L}} P_i)$ or $(M^*, m_\omega^*, \sigma^*, \mathcal{N} \cup \mathcal{L}, \bigcup_{ID_i \in \mathcal{N} \cup \mathcal{L}} P_i)$. We say $\mathcal{A}_2$ wins the game, if one of the following cases is satisfied:

Case 1: $\mathcal{A}_2$ outputs a tuple $(\pi^*, \mathcal{N} \cup \mathcal{L}, \bigcup_{ID_i \in \mathcal{N} \cup \mathcal{L}} P_i)$ satisfying:

1) π^* is a valid certificate.

2) π^* is not generated from the certificate generation query.
3) There is at least one user $ID \in \mathcal{N} \cup \mathcal{L}$ whose secret value is not queried and whose public key is not placed by $\mathcal{A}_2$.
4) $\mathcal{A}_2$ can't query the secret value for any identity if the corresponding public key has already been replaced.

Case 2: If $\mathcal{A}_2$ outputs a tuple $(M^*, m_\omega^*, \sigma^*, \mathcal{N} \cup \mathcal{L}, \bigcup_{ID_i \in \mathcal{N} \cup \mathcal{L}} P_i)$ satisfying:
1) σ^* is a valid proxy signature.
2) σ^* is not generated from the proxy signature query.
3) $(m_\omega^*, \mathcal{N} \cup \mathcal{L}, \bigcup_{ID_i \in \mathcal{N} \cup \mathcal{L}} P_i)$ didn't appear in the certificate generation query.
4) There is at least one user $ID \in \mathcal{N}$ whose secret value is not queried and whose public key is not placed by $\mathcal{A}_2$.
5) $\mathcal{A}_2$ can't query the secret value for any identity if the corresponding public key has been replaced.

Case 3: If $\mathcal{A}_2$ outputs a tuple $(M^*, m_\omega^*, \sigma^*, \mathcal{N} \cup \mathcal{L}, \bigcup_{ID_i \in \mathcal{N} \cup \mathcal{L}} P_i)$ satisfying:
1) σ^* is a valid proxy signature.
2) σ^* is not generated from the proxy signature query.
3) There is at least one user $ID \in \mathcal{L}$ whose secret is not queried and whose public key is not placed by $\mathcal{A}_2$.
4) $\mathcal{A}_2$ can't query the secret value for any identity if the corresponding public key has already been replaced.

The advantage of $\mathcal{A}_2$ is defined as

$$Adv_{\mathcal{A}_2}^{UNF-CLMPMS} = Pr[\mathcal{A}_2 \ wins].$$

3 Our Scheme

In this section we will propose a certificateless multi-proxy multi-signature scheme based on RSA problem and DL problem, with the clerk architecture and without pairings. The scheme involves a set of n original signers $\mathcal{N} = \{ID_{o1}, ID_{o2}, \cdots, ID_{on}\}$, a set of l proxy signers $\mathcal{L} = \{ID_{p1}, ID_{p2}, \cdots, ID_{pl}\}$, a verifier $\mathcal{V}$ and a clerk $\mathcal{B}$. A cooperative clerk reduces the communication cost. Our scheme is described as follows:

Setup. Given a security parameter k, KGC generates two random k-bit prime numbers p and q, then it computes $N = pq$. For some fixed parameter m (for example $m = 200$), KGC randomly chooses a prime number b satisfying $2^m < b < 2^{m+1}$ and $\gcd(b, \varphi(N)) = 1$. Then it chooses group $\mathbb{G}$ of prime order b, generator P of $\mathbb{G}$, and computes $a = b^{-1} \bmod \varphi(N)$. Furthermore, KGC chooses five hash functions as follows: $H_0 = \{0,1\}^* \to Z_N^*$, $H_i : \{0,1\}^* \to Z_b^*$ $(i = 1,2,3,4)$. Finally KGC outputs the set of public parameters $params = \{N, b, \mathbb{G}, P, H_0, H_1, H_2, H_3, H_4\}$, and the master secret key $msk = (p, q, a)$.

Partial private key extract. For an identity $ID_i \in \{0,1\}^*$, KGC computes $Q_i = H_0(ID_i)$, $D_i = Q_i^a$, then sends D_i to the user ID_i via secure channel.

Secret value set. The user with identity $ID_i \in \{0,1\}^*$ randomly chooses $t_i \in Z_b^*$.

Public key generation. The user with identity $ID_i \in \{0,1\}^*$ computes his public key $P_i = t_i P$.

Proxy certificate generation. m_ω is the warrant consisting of the identities of n original signers ID_{oi} $(i = 1,2,\cdots,n)$, l proxy signers ID_{pj} $(j = 1,2,\cdots,l)$, the certificate during and so on. On inputting the warrant m_ω, all signers ID_{oi} $(i = 1,2,\cdots,n)$ and ID_{pj} $(j = 1,2,\cdots,l)$ perform the following steps:

- Each ID_{oi} randomly selects $c_{oi} \in Z_b^*$, $A_{oi} \in Z_N^*$. Computes $S_{oi} = c_{oi}P$, $T_{oi} = A_{oi}^b \bmod N$. Broadcasts (S_{oi}, T_{oi}) to the other $n-1$ original signers, l proxy signers and clerk $\mathcal{B}$.
- Each ID_{pj} randomly selects $c_{pj} \in Z_b^*$, $A_{pj} \in Z_N^*$. Computes $S_{pj} = c_{pj}P$, $T_{p_j} = A_{p_j}^b \bmod N$. Broadcasts (S_{pj}, T_{pj}) to the other n original signers, $l-1$ proxy signers and clerk $\mathcal{B}$.
- Clerk $\mathcal{B}$ and all signers compute $S = \sum_{i=1}^n S_{oi} + \sum_{j=1}^l S_{pj}$, $T = \prod_{i=1}^n T_{oi} \cdot \prod_{j=1}^l T_{pj}$.
- Each ID_{oi} computes

$$\begin{aligned} k_{oi} &= H_1(m_\omega, ID_{oi}, P_{oi}, S, T), \\ h_{oi} &= H_2(m_\omega, ID_{oi}, P_{oi}, S, T). \end{aligned}$$

  Computes $r_{oi} = c_{oi} + t_{oi}k_{oi}$, $R_{oi} = A_{oi}D_{oi}^{h_{oi}} \bmod N$. Broadcasts (r_{oi}, R_{oi}) to clerk $\mathcal{B}$.
- Each ID_{pj} computes

$$\begin{aligned} k_{p_j} &= H_1(m_\omega, ID_{pj}, P_{pj}, S, T), \\ h_{pj} &= H_2(m_\omega, ID_{pj}, P_{pj}, S, T). \end{aligned}$$

  Computes $r_{pj} = c_{pj} + t_{pj}k_{pj}$, $R_{pj} = A_{pj}D_{pj}^{h_{pj}} \bmod N$. Broadcasts (r_{pj}, R_{pj}) to clerk $\mathcal{B}$.
- Clerk $\mathcal{B}$ does as follows:
  1) Verifies the correctness of r_{oi}, R_{oi} by checking the equations: $r_{oi}P = S_{oi} + k_{oi}P_{oi}$, $R_{oi}^b = T_{oi}Q_{oi}^{h_{oi}} \bmod N$ for $ID_{oi} \in \mathcal{N}$.
  2) Verifies the correctness of r_{pj}, R_{pj} by checking the equations: $r_{pj}P = S_{pj} + k_{pj}P_{pj}$, $R_{pj}^b = T_{pj}Q_{pj}^{h_{pj}} \bmod N$ for $ID_{pj} \in \mathcal{L}$.
  3) If all equalities hold, $\mathcal{B}$ computes

$$r = \sum_{i=1}^n r_{oi} + \sum_{j=1}^l r_{pj}, \ R = \prod_{i=1}^n R_{oi} \cdot \prod_{j=1}^l R_{pj}.$$

4) Sends $\pi = (m_\omega, r, R, S, T)$ to n original signers and l proxy signers.

Multi-proxy multi-sign. To sign a message M on behalf of the n original signers, the l proxy signers perform the following steps:

- Each proxy signer ID_{pj} selects $a_j \in Z_b^*$, $B_j \in Z_N^*$. Computes $X_j = a_j P$, $Y_j = B_j^b \bmod N$. Broadcasts (X_j, Y_j) to the other $l-1$ proxy signers.
- Each proxy signer ID_{pj} computes

$$\begin{aligned} X &= \sum_{j=1}^{l} X_j, \\ Y &= \prod_{j=1}^{l} Y_j. \\ \alpha_j &= H_3(M, m_\omega, ID_{pj}, P_{pj}, S, T, X, Y), \\ \beta_j &= H_4(M, m_\omega, ID_{pj}, P_{pj}, S, T, X, Y), \\ u_j &= r + a_j + t_{pj}\alpha_j, \\ U_j &= RB_j D_{pj}^{\beta_j}. \end{aligned}$$

Sends $(M, m_\omega, u_j, U_j, X_j, Y_j, S, T)$ to clerk $\mathcal{B}$.

- The clerk $\mathcal{B}$ checks whether all the proxy signers' partial signatures are correct.
  1) Computes $X = \sum_{j=1}^{l} X_j$, $Y = \prod_{j=1}^{l} Y_j$.
  2) Computes

$$\begin{aligned} k_{oi} &= H_1(m_\omega, ID_{oi}, P_{oi}, S, T), \\ h_{oi} &= H_2(m_\omega, ID_{oi}, P_{oi}, S, T) \\ &\qquad \text{for } i = 1, 2, \cdots, n. \\ k_{pj} &= H_1(m_\omega, ID_{pj}, P_{pj}, S, T), \\ h_{pj} &= H_2(m_\omega, ID_{pj}, P_{pj}, S, T) \\ &\qquad \text{for } j = 1, 2, \cdots, l. \\ \alpha_j &= H_3(M, m_\omega, ID_{pj}, P_{pj}, S, T, X, Y), \\ \beta_j &= H_4(M, m_\omega, ID_{pj}, P_{pj}, S, T, X, Y) \\ &\qquad \text{for } j = 1, 2, \cdots, l. \end{aligned}$$

  3) Computes

$$\begin{aligned} V &= \sum_{i=1}^{n} k_{oi} P_{oi} + \sum_{j=1}^{l} k_{pj} P_{pj}, \\ W &= \prod_{i=1}^{n} Q_{oi}^{h_{oi}} \cdot \prod_{j=1}^{l} Q_{pj}^{h_{pj}}. \end{aligned}$$

  4) Checks whether

$$\begin{aligned} u_j P &= S + V + X_j + \alpha_j P_{pj}, \\ U_j^b &= T \cdot W \cdot Y_j \cdot Q_{pj}^{\beta_j} \text{ for each } ID_j \in \mathcal{L}. \end{aligned}$$

If all equations hold, the clerk computes $u = \sum_{j=1}^{l} u_j$ and $U = \prod_{j=1}^{l} U_j$.

- Outputs multi-proxy multi-signature $\sigma = (M, m_\omega, u, U, S, T, X, Y)$.

Multi-proxy multi-signature verify. To verify the validity of the signature $\sigma = (M, m_\omega, u, U, S, T, X, Y)$ on message M, the verifier does as follows:

1) Checks whether the message M conforms to the warrant m_ω. If not, stops. Otherwise, continues.
2) Checks whether the l proxy signers are authorized by the original group $\mathcal{N}$ in the warrant m_ω. If not, stops. Otherwise, continues.
3) Computes

$$\begin{aligned} k_{oi} &= H_1(m_\omega, ID_{oi}, P_{oi}, S, T), \\ h_{oi} &= H_2(m_\omega, ID_{oi}, P_{oi}, S, T) \\ &\qquad \text{for } i = 1, 2, \cdots, n. \\ k_{pj} &= H_1(m_\omega, ID_{pj}, P_{pj}, S, T), \\ h_{pj} &= H_2(m_\omega, ID_{pj}, P_{pj}, S, T) \\ &\qquad \text{for } j = 1, 2, \cdots, l. \\ \alpha_j &= H_3(M, m_\omega, ID_{pj}, P_{pj}, S, T, X, Y), \\ \beta_j &= H_4(M, m_\omega, ID_{pj}, P_{pj}, S, T, X, Y) \\ &\qquad \text{for } j = 1, 2, \cdots, l. \end{aligned}$$

4) Computes

$$\begin{aligned} V &= \sum_{i=1}^{n} k_{oi} P_{oi} + \sum_{j=1}^{l} k_{pj} P_{pj}, \\ W &= \prod_{i=1}^{n} Q_{oi}^{h_{oi}} \cdot \prod_{j=1}^{l} Q_{pj}^{h_{pj}}. \end{aligned}$$

5) Checks whether the equations below hold. If both hold, accepts. Otherwise, rejects.

$$\begin{aligned} uP &= l(S + V) + X + \sum_{j=1}^{l} (\alpha_j P_{pj}), \\ U^b &= (TW)^l \cdot Y \cdot \prod_{j=1}^{l} Q_{pj}^{\beta_j}. \end{aligned}$$

Correctness:

$$\begin{aligned} uP &= \sum_{j=1}^{l} u_j P = \sum_{j=1}^{l} \left(r + a_j + t_{pj}\alpha_j\right) P \\ &= lrP + \sum_{j=1}^{l} (a_j P + \alpha_j t_{pj} P) \\ &= lrP + \sum_{j=1}^{l} (X_j + \alpha_j P_{pj}) \\ &= l(\sum_{i=1}^{n} r_{oi} + \sum_{j=1}^{l} r_{pj})P + \sum_{j=1}^{l} (X_j + \alpha_j P_{pj}) \\ &= \sum_{i=1}^{n} l r_{oi} P + \sum_{j=1}^{l} (l r_{pj} P + X_j + \alpha_j P_{pj}) \end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{n} l(c_{oi} + t_{oi}k_{oi})P \\
&\quad + \sum_{j=1}^{l}(l(c_{pj} + t_{pj}k_{pj})P + X_j + \alpha_j P_{pj}) \\
&= \sum_{i=1}^{n}(lS_{oi} + lk_{oi}P_{oi}) \\
&\quad + \sum_{j=1}^{l}(lS_{pj} + lk_{pj}P_{pj} + X_j + \alpha_j P_{pj}) \\
&= lS + X + \sum_{i=1}^{n} lk_{oi}P_{oi} + \sum_{j=1}^{l}(lk_{pj} + \alpha_j)P_{pj} \\
&= l(S+V) + X + \sum_{j=1}^{l}(\alpha_j P_{pj}).
\end{aligned}
$$

$$
\begin{aligned}
U^b &= \big(\prod_{j=1}^{l} U_j\big)^b = \prod_{j=1}^{l}\big(RB_j D_{pj}^{\beta_j}\big)^b = R^{lb}\cdot\prod_{j=1}^{l} Y_j Q_{pj}^{\beta_j} \\
&= (\prod_{i=1}^{n} R_{oi})^{lb}\cdot(\prod_{j=1}^{l} R_{pj})^{lb}\cdot\prod_{j=1}^{l} Y_j Q_{pj}^{\beta_j} \\
&= \prod_{i=1}^{n}(A_{oi}D_{oi}^{h_{oi}})^{lb}\cdot\prod_{j=1}^{l}(A_{pj}D_{pj}^{h_{pj}})^{lb} Y_j Q_{pj}^{\beta_j} \\
&= \prod_{i=1}^{n} T_{oi}^{l} Q_{oi}^{lh_{oi}}\cdot\prod_{j=1}^{l} T_{pj}^{l} Q_{P_j}^{lh_{pj}+\beta_j} Y_j \\
&= T^l\cdot Y\cdot\prod_{i=1}^{n} Q_{oi}^{lh_{oi}}\cdot\prod_{j=1}^{l} Q_{pj}^{lh_{pj}+\beta_j} \\
&= (TW)^l\cdot Y\cdot\prod_{j=1}^{l} Q_{pj}^{\beta_j}.
\end{aligned}
$$

4 Security Results

Theorem 1. *The scheme is unforgeable against the type I adversary $\mathcal{A}_1$ if the RSA problem is hard in random oracle model.*

Proof. Suppose the challenger $\mathcal{C}$ receives a random instance $(N, b, \mathcal{Y})$ of the RSA problem and has to find an element $\mathcal{X} \in Z_N^*$ such that $\mathcal{X}^b = \mathcal{Y}$. Challenger $\mathcal{C}$ will run $\mathcal{A}_1$ as a subroutine and act as $\mathcal{A}_1$'s challenger in the UNF-CLMPMS Game I.

Setup: At the beginning of the game, $\mathcal{C}$ runs the setup algorithm with the parameter k and gives $\mathcal{A}_1$ the system parameters $params = \{N, b, \mathbb{G}, P, H_0, H_1, H_2, H_3, H_4\}$ and $\mathcal{A}_1$ doesn't know the master secret key $msk = (p, q, a)$.

Queries: Without loss of generality we assume that all the queries are distinct and $\mathcal{A}_1$ will make H_0 query for ID_i before ID_i is used in any other queries.

1) H_0 queries: $\mathcal{C}$ maintains the list L_0 of tuple (ID_i, A_i). The list is initially empty. When $\mathcal{A}_1$ makes a query $H_0(ID_i)$, $\mathcal{C}$ responds as follows: At the j^{th} H_0 query, $\mathcal{C}$ sets $H_0(ID^*) = \mathcal{Y}$. For $i \neq j$, $\mathcal{C}$ randomly picks a value $A_i \in \mathbb{Z}_N^*$ and sets $H_0(ID_i) = A_i^b$. Then the query and the answer will be stored in the list L_0.
2) H_1 queries: $\mathcal{C}$ maintains the list L_1 of tuple (γ_i, k_i). The list is initially empty. When $\mathcal{A}_1$ makes a query $H_1(\gamma_i)$, $\mathcal{C}$ randomly picks a value $k_i \in \mathbb{Z}_b^*$ and sets $H_1(\gamma_i) = k_i$. The query and the answer will be stored in the list L_1.
3) H_2 queries: $\mathcal{C}$ maintains the list L_2 of tuple (γ_i, h_i). The list is initially empty. When $\mathcal{A}_1$ makes a query $H_2(\gamma_i)$, $\mathcal{C}$ randomly picks a value $h_i \in \mathbb{Z}_b^*$ and sets $H_2(\gamma_i) = h_i$. The query and the answer will be stored in the list L_2.
4) H_3 queries: $\mathcal{C}$ maintains the list L_3 of tuple (η_i, α_i). The list is initially empty. When $\mathcal{A}_1$ makes a query $H_3(\eta_i)$, $\mathcal{C}$ randomly picks a value $\alpha_i \in \mathbb{Z}_b^*$ and sets $H_3(\eta_i) = \alpha_i$. The query and the answer will be stored in the list L_3.
5) H_4 queries: $\mathcal{C}$ maintains the list L_4 of tuple (η_i, β_i). The list is initially empty. When $\mathcal{A}_1$ makes a query $H_4(\eta_i)$, $\mathcal{C}$ randomly picks a value $\beta_i \in \mathbb{Z}_b^*$ and sets $H_4(\eta_i) = \beta_i$. The query and the answer will be stored in the list L_4.
6) User-Public-Key queries: $\mathcal{C}$ maintains the list L_U of tuple (ID_i, t_i, P_i). When $\mathcal{A}_1$ makes user public key query for ID_i, $\mathcal{C}$ randomly chooses $t_i \in \mathbb{Z}_b^*$, sets $P_i = t_i P$. Then sends the P_i to $\mathcal{A}_1$. The tuple (ID_i, t_i, P_i) will be stored in the list L_U.
7) User-Public-Key-Replacement: $\mathcal{C}$ maintains the list L_R of tuple (ID_i, P_i, P_i'). When $\mathcal{A}_1$ makes a user public key replacement for ID_i with a new value P_i', $\mathcal{C}$ replaces the current public key P_i with the value P_i' and the tuple (ID_i, P_i, P_i') will be stored in the list L_R.
8) Partial-Private-Key queries: $\mathcal{C}$ maintains the list L_K of tuple (ID_i, A_i). When $\mathcal{A}_1$ makes a partial private key query for ID_i, If $ID_i = ID^*$, $\mathcal{C}$ fails and stops, otherwise $\mathcal{C}$ finds the tuple (ID_i, A_i) in list L_0 and responds with A_i. The tuple (ID_i, A_i) will be stored in the list L_K.
9) Secret-Value queries: $\mathcal{C}$ maintains the list L_S of tuple (ID_i, t_i). When $\mathcal{A}_1$ makes a secret value query for ID_i, $\mathcal{C}$ finds the tuple (ID_i, t_i, P_i) in list L_U and responds with t_i. The tuple (ID_i, t_i) will be stored in the list L_S. $\mathcal{A}_1$ can't query the secret value for ID_i whose public key has been replaced.
10) Proxy-Certificate-Generation : When $\mathcal{A}_1$ submits all signers' identities/public keys (ID_i, P_i), $ID_i \in \mathcal{N} \cup \mathcal{L}$ and a warrant m_ω to the challenger,

$\mathcal{C}$ responds by running the certificate generation algorithm on the warrant m_ω and the signers' full private key $(t_i, D_i), ID_i \in \mathcal{N} \cup \mathcal{L}$, then outputs a certificate as follows:

If $ID^* \notin \mathcal{N} \cup \mathcal{L}$ and $(\mathcal{N} \cup \mathcal{L}) \bigcap L_R = \emptyset$, $\mathcal{C}$ gives a certificate by calling the certificate generation algorithm. Otherwise, $\mathcal{C}$ does the following:

a. Randomly selects $r \in Z_b^*, R \in Z_N^*$.
b. Randomly selects $k_i, h_i \in Z_b^*$ for each $ID_i \in \mathcal{N} \cup \mathcal{L}$.
c. Computes $S = rP - \sum_{ID_i \in \mathcal{N} \bigcup \mathcal{L}} k_i P_i$ and $T = R^b \cdot \prod_{ID_i \in \mathcal{N} \bigcup \mathcal{L}} Q_i^{-h_i}$.
d. Stores the relation $k_i = H_1(m_\omega, ID_i, P_i, S, T)$ and $h_i = H_2(m_\omega, ID_i, P_i, S, T)$ for each $ID_i \in \mathcal{N} \cup \mathcal{L}$. Repeats the steps (1)-(3) if collision occurs.
e. Outputs $\pi = (m_\omega, r, R, S, T)$ as the proxy certificate.

11) Multi-proxy Multi-sign: When $\mathcal{A}_1$ submits certificate $\pi = (m_\omega, r, R, S, T)$ and a message M to the challenger, $\mathcal{C}$ outputs a signature by running the multi-proxy multi-sign algorithm on π and M as follows: If $ID^* \notin \mathcal{L}$ and $\mathcal{L} \bigcap L_R = \emptyset$, $\mathcal{C}$ gives a signature by calling the multi-proxy multi-sign algorithm. Otherwise, $\mathcal{C}$ does the following:

a. Randomly selects $u \in Z_b^*, U \in Z_N^*$.
b. Randomly selects $\alpha_j, \beta_j \in Z_b^*$ for each $ID_j \in \mathcal{L}$.
c. Computes $X = uP - lS - \sum_{i=1}^{n} lk_{oi} P_{oi} - \sum_{j=1}^{l} (lk_{pj} + \alpha_j) P_{pj}$ and $Y = U^b \cdot T^{-l} \cdot \prod_{i=1}^{n} Q_{oi}^{-lh_{oi}} \cdot \prod_{j=1}^{l} Q_{pj}^{-lh_{pj}-\beta_j}$.
d. Stores the relation $\alpha_j = H_3(M, m_\omega, ID_{p_j}, P_{p_j}, S, T, X, Y)$ and $\beta_j = H_4(M, m_\omega, ID_{p_j}, P_{p_j}, S, T, X, Y)$. Repeats the steps (1)-(3) if collision occurs.
e. Outputs $\sigma = (M, m_\omega, u, U, S, T, X, Y)$ as the proxy signature.

Forge: $\mathcal{A}_1$ outputs the tuple $\{\pi = (m_\omega, r, R, S, T), \mathcal{L} \cup \mathcal{N}, \bigcup_{ID_i \in \mathcal{L} \cup \mathcal{N}} P_i\}$ or $\{\sigma = (M, m_\omega, u, U, S, T, X, Y), \mathcal{L} \cup \mathcal{N}, \bigcup_{ID_i \in \mathcal{L} \cup \mathcal{N}} P_i\}$.

Solve RSAP. If $\mathcal{A}_1$'s output satisfies none of the following 3 cases in UNF-CLMPMS Game I, $\mathcal{C}$ aborts. Otherwise, $\mathcal{C}$ can solve the RSA problem as follows:

Case 1. The final output is $\{\pi = (m_\omega, r, R, S, T), \mathcal{L} \cup \mathcal{N}, \bigcup_{ID_i \in \mathcal{L} \cup \mathcal{N}} P_i\}$ and the output satisfies the requirement of Case 1 as defined in UNF-CLMPMS Game I. In fact, π is the signature on m_ω. If $ID^* \in \mathcal{N} \cup \mathcal{L}$, we can solve the RSA problem as follows.

Without loss of generality, we may assume that $ID^* = ID_\lambda$. By Forking Lemma for generic signature scheme, we can get another $\pi' = (m_\omega, r, R', S, T)$. To do so we maintain all the random tapes in two invocations are the same except the λ^{th} result returned by H_2 query of the forged message. In other words $h_\lambda \neq h'_\lambda$ and $h_i = h'_i$ for $i \neq \lambda$. The relation becomes $(R' \cdot R^{-1})^b = \mathcal{Y}^{h'_\lambda - h_\lambda} \bmod N$. Since $h_\lambda, h'_\lambda \in \mathbb{Z}_b^*$, we have that $|h'_\lambda - h_\lambda| < b$. By the element b is a prime number, then $gcd(b, h'_\lambda - h_\lambda) = 1$. This means that there exists two integers μ, ν such that $\mu b + \nu(h'_\lambda - h_\lambda) = 1$. Finally, the value $\mathcal{X} = (R' R^{-1})^\nu \mathcal{Y}^\mu \bmod N$ is the solution of the given instance of the RSA problem.

$$\begin{aligned} \mathcal{X}^b &= (R' R^{-1})^{b\nu} \mathcal{Y}^{b\mu} \\ &= \mathcal{Y}^{\nu(h'_\lambda - h_\lambda)} \mathcal{Y}^{b\mu} \\ &= \mathcal{Y}^{b\mu + \nu(h'_\lambda - h_\lambda)} \\ &= \mathcal{Y}. \end{aligned}$$

Probability of success. Let q_{H_i} $(i = 0, 1, 2, 3, 4)$, q_U, q_K, q_C and q_P be the number of $H_i (i = 0, 1, 2, 3, 4)$ queries, user public key queries, partial private key queries, proxy certificate generation queries, multi-proxy multi-signature queries, respectively.

The probability that $\mathcal{C}$ doesn't fail during the queries is $\frac{q_{H_0} - q_K}{q_{H_0}}$. The probability that $ID^* \in \mathcal{L} \cup \mathcal{N}$ is $\frac{n+l-1}{q_K} \cdot \frac{1}{q_{H_0} - q_K}$. So the combined probability is $\frac{q_{H_0} - q_K}{q_{H_0}} \cdot \frac{n+l-1}{q_K} \cdot \frac{1}{q_{H_0} - q_K} = \frac{n+l-1}{q_K \cdot q_{H_0}}$. Therefore, if $\mathcal{A}_1$ can succeed with the probability ε within time $\mathcal{T}$, then $\mathcal{C}$ can solve RSAP with the probability $\frac{(n+l-1)\varepsilon}{q_K \cdot q_{H_0}}$. The running time required for $\mathcal{C}$ is: $2\mathcal{T} + [q_{H_0} + (3n + 3l + 2)q_D + 2lq_P]T_N + [q_U + (2n + 2l + 2)q_D + lq_P)]T_E$, where T_N denotes the time for a modular operation and T_E denotes the time for a exponentiation in $\mathbb{G}$.

Case 2. The final output is $\{\sigma = (M, m_\omega, u, U, S, T, X, Y), \mathcal{L} \cup \mathcal{N}, \bigcup_{ID_i \in \mathcal{L} \cup \mathcal{N}} P_i\}$ and the output satisfies the requirement of Case 2 as defined in UNF-CLMPMS Game I. If $ID^* \in \mathcal{N}$, we can solve the RSA problem as follows.

Without loss of generality, we may assume that $ID^* = ID_\lambda$. By the Forking Lemma for generic signature scheme, we can get another signature $(M, m_\omega, u, U', S, T, X, Y)$. To do so we maintain all the random tapes in two invocations are the same except the λ^{th} result returned by H_2 query of the forged message. In other words $h_\lambda \neq h'_\lambda$ and $h_i = h'_i$ for $i \neq \lambda$. The relation becomes $(U' \cdot U^{-1})^{bl^{-1}} = \mathcal{Y}^{h'_\lambda - h_\lambda} \bmod N$. Since $h_\lambda, h'_\lambda \in \mathbb{Z}_b^*$, we have that $|h'_\lambda - h_\lambda| < b$. By the element b is a prime number, then $gcd(b, h'_\lambda - h_\lambda) = 1$. This means that there exists two integers μ, ν such that $\mu b + \nu(h'_\lambda - h_\lambda) = 1$. Finally, the value $\mathcal{X} = (U' U^{-1})^{l^{-1}\nu} \mathcal{Y}^\mu \bmod N$ is the solution of

the given instance of the RSA problem.

$$\begin{aligned}\mathcal{X}^b &= (U'U^{-1})^{bl^{-1}\nu}\mathcal{Y}^{b\mu}\\ &= \mathcal{Y}^{\nu(h'_\lambda - h_\lambda)}\mathcal{Y}^{b\mu}\\ &= \mathcal{Y}^{b\mu+\nu(h'_\lambda-h_\lambda)}\\ &= \mathcal{Y}.\end{aligned}$$

Probability of success. Let q_{H_i} $(i = 0, 1, 2, 3, 4)$, q_U, q_K, q_C and q_P be the number of $H_i(i = 0, 1, 2, 3, 4)$ queries, user public key queries, partial private key queries, proxy certificate generation queries, multi-proxy multi-signature queries, respectively.

The probability that $\mathcal{C}$ doesn't fail during the queries is $\frac{q_{H_0}-q_K}{q_{H_0}}$. The probability that $ID^* \in \mathcal{N}$ is $\frac{n-1}{q_K}\cdot\frac{1}{q_{H_0}-q_K}$. So the combined probability is $\frac{q_{H_0}-q_K}{q_{H_0}}\cdot\frac{n-1}{q_K}\cdot\frac{1}{q_{H_0}-q_K} = \frac{n-1}{q_K\cdot q_{H_0}}$. Therefore, if $\mathcal{A}_1$ can succeed with the probability ε within time $\mathcal{T}$, then $\mathcal{C}$ can solve RSAP with the probability $\frac{(n-1)\varepsilon}{q_K\cdot q_{H_0}}$. The running time required for $\mathcal{C}$ is the same as the time in Case 1.

Case 3. The final output is $\{\sigma = (M, m_\omega, u, U, S, T, X, Y), \mathcal{L}\cup\mathcal{N}, \bigcup_{ID_i\in\mathcal{L}\cup\mathcal{N}} P_i\}$ and the output satisfies the requirement of *Case 3* as defined in UNF-CLMPMS Game I. If $ID^* \in \mathcal{L}$, we can solve the RSA problem as follows.

Without loss of generality, we may assume that $ID^* = ID_\lambda$. By the Forking Lemma for generic signature scheme, we can get another signature $(M, m_\omega, u, U', S, T, X, Y)$. To do so we maintain all the random tapes in two invocations are the same except the λ^{th} result returned by H_4 query of the forged message. In other words $\beta_\lambda \neq \beta'_\lambda$ and $\beta_j = \beta'_j$ for $j \neq \lambda$. The relation becomes $(U'\cdot U^{-1})^b = \mathcal{Y}^{\beta'_\lambda-\beta_\lambda} \bmod N$. Since $\beta_\lambda, \beta'_\lambda \in \mathbb{Z}^*_b$, we have that $|\beta'_\lambda - \beta_\lambda| < b$. By the element b is a prime number, then $gcd(b, \beta'_\lambda - \beta_\lambda) = 1$. This means that there exists two integers μ, ν such that $\mu b + \nu(\beta'_\lambda - \beta_\lambda) = 1$. Finally, the value $\mathcal{X} = (U'U^{-1})^\nu\mathcal{Y}^\mu \bmod N$ is the solution of the given instance of the RSA problem.

$$\begin{aligned}\mathcal{X}^b &= (U'U^{-1})^{b\nu}\mathcal{Y}^{b\mu}\\ &= \mathcal{Y}^{\nu(\beta'_\lambda-\beta_\lambda)}\mathcal{Y}^{b\mu}\\ &= \mathcal{Y}^{b\mu+\nu(\beta'_\lambda-\beta_\lambda)}\\ &= \mathcal{Y}.\end{aligned}$$

Probability of success. Let $q_{H_i}(i = 0, 1, 2, 3, 4)$, q_U, q_K, q_C and q_P be the number of $H_i(i = 0, 1, 2, 3, 4)$ queries, user public key queries, partial private key queries, proxy certificate generation queries, multi-proxy multi-signature queries, respectively.

The probability that $\mathcal{C}$ doesn't fail during the queries is $\frac{q_{H_0}-q_K}{q_{H_0}}$. The probability that $ID^* \in \mathcal{L}$ is $\frac{l-1}{q_K}\cdot\frac{1}{q_{H_0}-q_K}$. So the combined probability is $\frac{q_{H_0}-q_K}{q_{H_0}}\cdot\frac{l-1}{q_K}\cdot\frac{1}{q_{H_0}-q_K} = \frac{l-1}{q_K\cdot q_{H_0}}$. Therefore, if $\mathcal{A}_1$ can succeed with the probability ε within time $\mathcal{T}$, then $\mathcal{C}$ can solve RSAP with the probability $\frac{(l-1)\varepsilon}{q_K\cdot q_{H_0}}$. The running time required for $\mathcal{C}$ is the same as the time in Case 1.

□

Theorem 2. *The scheme is unforgeable against the type II adversary $\mathcal{A}_2$ if the DL problem is hard in randomly oracle model.*

Proof. Suppose the challenger $\mathcal{C}$ receives a random instance (P, xP) of the DL problem and has to compute $x \in Z^*_b$. Challenger $\mathcal{C}$ will run $\mathcal{A}_2$ as a subroutine and act as $\mathcal{A}_2$'s challenger in the UNF-CLMPMS Game II.

Setup: At the beginning of the game, $\mathcal{C}$ runs the setup algorithm with the parameter k and gives $\mathcal{A}_2$ the system parameters $params = \{N, b, \mathbb{G}, P, H_0, H_1, H_2, H_3, H_4\}$ and the master secret key $msk = (p, q, a)$.

Queries: Without loss of generality we assume that all the queries are distinct and $\mathcal{A}_2$ will make user's public key query for ID_i before ID_i is used in any other queries.

1) User-Public-Key queries: $\mathcal{C}$ maintains the list L_U of tuple (ID_i, t_i, P_i). When $\mathcal{A}_2$ makes public key query for ID_i, $\mathcal{C}$ responds as follows:

   At the j^{th} query, $\mathcal{C}$ sets $ID_j = ID^*$ and $P^* = xP$. For $i \neq j$, $\mathcal{C}$ randomly chooses $t_i \in \mathbb{Z}^*_b$, sets $P_i = t_iP$. Then the query and the answer will be stored in the list L_U.

2) H_0 queries: $\mathcal{C}$ maintains the list L_0 of tuple (ID_i, A_i). The list is initially empty. When $\mathcal{A}_2$ makes a query $H_0(ID_i)$, $\mathcal{C}$ randomly picks a value $A_i \in \mathbb{Z}^*_N$ and sets $H_0(ID_i) = A^b_i$. Then the query and the answer will be stored in the list L_0.

3) H_1, H_2, H_3, H_4 queries and User-Public-Key-Replacement are the same as those in Theorem 1.

4) Partial-Private-Key queries: $\mathcal{C}$ maintains the list L_K of tuple (ID_i, A_i). When $\mathcal{A}_2$ makes a partial private key query for ID_i, $\mathcal{C}$ finds the tuple (ID_i, A_i) in list L_0 and responds with A_i. The tuple (ID_i, A_i) will be stored in the list L_K.

5) Secret-Value queries: $\mathcal{C}$ maintains the list L_S of tuple (ID_i, t_i). When $\mathcal{A}_2$ makes a secret value query for ID_i, If $ID_i = ID^*$, $\mathcal{C}$ fails and stops, otherwise $\mathcal{C}$ finds the tuple (ID_i, t_i, P_i) in list L_U and responds with t_i. The tuple (ID_i, t_i) will be stored in the list L_S. $\mathcal{A}_2$ can't query the secret value for ID_i whose public key has been replaced.

6) proxy certificate generation queries and Multi-proxy Multi-sign queries are the same as those in Theorem 1.

Forge: $\mathcal{A}_2$ outputs the tuple $\{\pi = (m_\omega, r, R, S, T), \mathcal{L} \cup \mathcal{N}, \bigcup_{ID_i \in \mathcal{L} \cup \mathcal{N}} P_i\}$ or $\{\sigma = (M, m_\omega, u, U, S, T, X, Y), \mathcal{L} \cup \mathcal{N}, \bigcup_{ID_i \in \mathcal{L} \cup \mathcal{N}} P_i\}$.

Solve DLP. If $\mathcal{A}_2$'s output satisfies none of the following 3 cases in UNF-CLMPMS Game II, $\mathcal{C}$ aborts. Otherwise, $\mathcal{C}$ can solve the DL problem as follows:

Case 1. The final output is $\{\pi = (m_\omega, r, R, S, T), \mathcal{L} \cup \mathcal{N}, \bigcup_{ID_i \in \mathcal{L} \cup \mathcal{N}} P_i\}$ and the output satisfies the requirement of Case 1 as defined in UNF-CLMPMS Game II. In fact, π is the signature on m_ω. If $ID^* \in \mathcal{N} \cup \mathcal{L}$, we can solve the DL problem as follows.

Without loss of generality, we may assume that $ID^* = ID_\lambda$. By Forking Lemma for generic signature scheme, we can get another $\pi'(m_\omega, r', R, S, T)$. To do so we maintain all the random tapes in two invocations are the same except the λ^{th} result returned by H_1 query of the forged message. In other words $k_\lambda \neq k'_\lambda$ and $k_i = k'_i$ for $i \neq \lambda$. We note that $r = c_\lambda + k_\lambda x + \sum_{ID_i \in \mathcal{N} \cup \mathcal{L} \setminus \{ID*\}} (c_i + k_i t_i)$, $r' = c_\lambda + k'_\lambda x + \sum_{ID_i \in \mathcal{N} \cup \mathcal{L} \setminus \{ID*\}} (c_i + k_i t_i)$. It follows that $x = \frac{r - r'}{k_\lambda - k'_\lambda}$.

Probability of success. Let $q_{H_i} (i = 0, 1, 2, 3, 4)$, q_U, q_K, q_S, q_C and q_P be the number of $H_i (i = 0, 1, 2, 3, 4)$ queries, user public key queries, partial private key queries, secret value queries, proxy certificate generation queries, multi-proxy multi-signature queries, respectively.

The probability that $\mathcal{C}$ doesn't fail during the queries is $\frac{q_U - q_S}{q_U}$. The probability that $ID^* \in \mathcal{L} \cup \mathcal{N}$ is $\frac{n+l-1}{q_S} \cdot \frac{1}{q_U - q_S}$. So the combined probability is $\frac{q_U - q_S}{q_U} \cdot \frac{n+l-1}{q_S} \cdot \frac{1}{q_U - q_S} = \frac{n+l-1}{q_S \cdot q_U}$. Therefore, if $\mathcal{A}_2$ can succeed with the probability ε within time $\mathcal{T}$, then $\mathcal{C}$ can solve DL problem with the probability $\frac{(n+l-1)\varepsilon}{q_S \cdot q_U}$. The running time required for $\mathcal{C}$ is: $2\mathcal{T} + [q_{H_0} + (3n + 3l + 2)q_D + 2lq_P]T_N + [q_U + (2n + 2l + 2)q_D + lq_P)]T_E$, where T_N denotes the time for a modular operation and T_E denotes the time for a exponentiation in $\mathbb{G}$.

Case 2. The final output is $\{\sigma = (M, m_\omega, u, U, S, T, X, Y), \mathcal{L} \cup \mathcal{N}, \bigcup_{ID_i \in \mathcal{L} \cup \mathcal{N}} P_i\}$ and the output satisfies the requirement of Case 2 as defined in UNF-CLMPMS Game II. If $ID^* \in \mathcal{N}$, we can solve the DL problem as follows.

Without loss of generality, we may assume that $ID^* = ID_\lambda$. By the Forking Lemma for generic signature scheme, we can get another signature $(M, m_\omega, u', U, S, T, X, Y)$. To do so we maintain all the random tapes in two invocations are the same except the λ^{th} result returned by H_1 query of the forged message. In other words $k_\lambda \neq k'_\lambda$ and $k_i = k'_i$ for $i \neq \lambda$. We note that $u = l(c_\lambda + k_\lambda x + \sum_{ID_i \in \mathcal{N} \cup \mathcal{L} \setminus \{ID*\}} (c_i + k_i t_i)) + \sum_{ID_j \in \mathcal{L}} (a_j + \alpha_j t_j)$, $u' = l(c_\lambda + k'_\lambda x + \sum_{ID_i \in \mathcal{N} \cup \mathcal{L} \setminus \{ID*\}} (c_i + k_i t_i)) + \sum_{ID_j \in \mathcal{L}} (a_j + \alpha_j t_j)$. It follows that $x = \frac{u - u'}{l(k_\lambda - k'_\lambda)}$.

Probability of success. Let $q_{H_i} (i = 0, 1, 2, 3, 4)$, q_U, q_K, q_S, q_C and q_P be the number of $H_i (i = 0, 1, 2, 3, 4)$ queries, user public key queries, partial private key queries, secret value queries, proxy certificate generation queries, multi-proxy multi-signature queries, respectively.

The probability that $\mathcal{C}$ doesn't fail during the queries is $\frac{q_U - q_S}{q_U}$. The probability that $ID^* \in \mathcal{N}$ is $\frac{n-1}{q_S} \cdot \frac{1}{q_U - q_S}$. So the combined probability is $\frac{q_U - q_S}{q_U} \cdot \frac{n-1}{q_S} \cdot \frac{1}{q_U - q_S} = \frac{n-1}{q_S \cdot q_U}$. Therefore, if $\mathcal{A}_2$ can succeed with the probability ε within time $\mathcal{T}$, then $\mathcal{C}$ can solve DL problem with the probability $\frac{(n-1)\varepsilon}{q_S \cdot q_U}$. The running time required for $\mathcal{C}$ is the same as the time in Case 1.

Case 3. The final output is $\{\sigma = (M, m_\omega, u, U, S, T, X, Y), \mathcal{L} \cup \mathcal{N}, \bigcup_{ID_i \in \mathcal{L} \cup \mathcal{N}} P_i\}$ and the output satisfies the requirement of *Case 3* as defined in UNF-CLMPMS Game II. If $ID^* \in \mathcal{L}$, we can solve the DL problem as follows.

Without loss of generality, we may assume that $ID^* = ID_\lambda$. By the Forking Lemma for generic signature scheme, we can get another signature $(M, m_\omega, u', U, S, T, X, Y)$. To do so we maintain all the random tapes in two invocations are the same except the λ^{th} result returned by H_3 query of the forged message. In other words $\alpha_\lambda \neq \alpha'_\lambda$ and $\alpha_j = \alpha'_j$ for $j \neq \lambda$. We note that $u = lr + \sum_{ID_j \in \mathcal{L} \setminus \{ID^*\}} (a_j + \alpha_j t_j) + (a_\lambda + \alpha_\lambda x)$, $u' = lr + \sum_{ID_j \in \mathcal{L} \setminus \{ID^*\}} (a_j + \alpha_j t_j) + (a_\lambda + \alpha'_\lambda x)$, It follows that $x = \frac{u - u'}{\alpha_\lambda - \alpha'_\lambda}$.

Probability of success. Let $q_{H_i} (i = 0, 1, 2, 3, 4)$, q_U, q_K, q_S, q_C and q_P be the number of $H_i (i = 0, 1, 2, 3, 4)$ queries, user public key queries, partial private key queries, secret value queries, proxy certificate generation queries, multi-proxy multi-signature queries, respectively.

The probability that $\mathcal{C}$ doesn't fail during the queries is $\frac{q_U - q_S}{q_U}$. The probability that $ID^* \in \mathcal{L}$ is $\frac{l-1}{q_S} \cdot \frac{1}{q_U - q_S}$. So the combined probability is $\frac{q_U - q_S}{q_U} \cdot \frac{l-1}{q_S} \cdot \frac{1}{q_U - q_S} = \frac{l-1}{q_S \cdot q_U}$. Therefore, if $\mathcal{A}_2$ can succeed with the probability ε within time $\mathcal{T}$, then $\mathcal{C}$ can solve DL Problem with the probability $\frac{(l-1)\varepsilon}{q_S \cdot q_U}$. The running time required for $\mathcal{C}$ is the same as the time in Case 1.

□

5 Efficiency and Comparison

Our scheme is constructed without using bilinear pairing. In the following, we compare the performance of our scheme with several MPMS schemes in Table 2. We define some notations as follows:

- T_P: A pairing operation.
- E_P: A pairing-based scalar multiplication operation.
- T_E: A scalar multiplication operation in the elliptic curve group $\mathbb{G}$.
- T_N: A modular exponent operation in $\mathbb{Z}_N$.

Through PIV 3-GHZ processor with 512-MB memory and a Windows XP operation system. He *et al.* [5] obtained the running time for cryptographic operations. To achieve 1024-bit RSA level security, they use the Tate pairing defined over a super singular curve $E/F_p : y^2 = x^3 + x$ with embedding degree 2, where q is a 160-bit Solinas prime $q = 2^{159} + 2^{17} + 1$ and p is a 512-bit prime satisfying $p + 1 = 12qr$. To achieve the same security level, they employed the parameter secp160r1 [18], where $p = 2^{160} - 2^{31} - 1$. The running times are listed in Table 1.

Table 1: Cryptographic operation time (in milliseconds)

T_P	T_N	E_P	T_E
20.04	5.31	6.38	2.21

To evaluate the computation efficiency of different schemes, we use a simple method. For example in [10], system costs $3n+3l+2$ pairing-based scalar multiplication operations and $3n+3l$ pairing operations in Proxy Certificate Generation, system costs $3l + 3$ pairing-based scalar multiplication operations and $3l + 6$ pairing operations in Multi-Proxy Multi-Sign and Verify. Hence system costs $3n + 6l + 5$ pairing-based scalar multiplication operations and $3n+6l+6$ pairing operations in total. To facilitate the comparison, we let $n = l = 10$. So the resulting computation time is $95 \times 6.38 + 96 \times 11.20 = 2721.34$. The detailed comparison results of several different MPMS schemes are illustrated in Table 2 and Table 3.

Table 2: Comparison of several CLMPMS schemes

Scheme	Public key form	Secure base
Li [10]	ID-base	CDHP
Sahu [15]	ID-base	CDHP
Our scheme	Certificateless	RSAP and DLP

6 Conclusion

In a multi-proxy multi-signature scheme, the group of original signers delegate their signing rights to the proxy group. RSA is a key cryptography technique and provides various interfaces for the applied software in real-life scenarios. Although some good results were achieved in speeding up the computation of pairing function in recent years, the computation cost of the pairing is much higher than that of the exponentiation in a RSA group and also much higher than the scalar multiplication over the elliptic curve group. In this paper, we propose a certificateless multi-proxy multi-signature scheme and prove that our scheme is unforgeable under the strongest security model where the Type I/II adversary is a super Type I/II adversary. The analysis shows that our scheme is more efficient than the related schemes. Due to the very good properties of our scheme, it is very useful for practical applications.

Acknowledgments

This research is supported by the National Natural Science Foundation of China under Grants 61562012, 11261060, the Innovation Group Major Research Projects of Department of Education of Guizhou Province under Grant No.KY[2016]026. The authors gratefully acknowledge the anonymous reviewers for their valuable comments.

References

[1] S. Al-Riyami and K. Paterson, "Certificateless public key cryptography," in *Advances in Cryptology (Asiacrypt'03)*, LNCS 2894, pp. 452-473, 2003.

[2] M. K. Chande, C. C. Lee, C. T. Li, "Message recovery via an efficient multi-proxy signature with self-certified keys," *International Journal of Network Security*, vol. 19, no. 3, pp. 340-346, 2017.

[3] L. Deng, H. Huang and Y. Qu, "Identity based proxy signature from RSA without pairings," *International Journal of Network Security*, vol. 19, no. 2, pp.229-235, 2017.

[4] L. Deng, J. Zeng and Y. Qu, "Certificateless proxy signature from RSA," *Mathematical Provlems in Engineering*, vol. 2014, pp. 1-10, 2014.

[5] D. He, J. Chen and J. Hu, "An ID-based proxy signature schemes without bilinear pairings," *Annals of Telecommunications*, vol. 66, pp. 657-662, 2011.

[6] D. He, Y. Chen and J. Chen, "An efficient certificateless proxy signature scheme without pairing," *Mathmatical and Computer Modelling*, vol. 57, pp. 2510-2518, 2013.

[7] J. Herranz, "Identity-based ring signatures from RSA," *Theoretical Computer Science*, vol. 389, pp. 100-117, 2007.

Table 3: Comparison of several CLMPMS schemes

Scheme	Proxy certificate generation	MPMSign and verify	Time(n=l=10)
Li [10]	$(3n+3l+2)E_P+(3n+3l)T_P$	$(3l+3)E_P+(3l+6)T_P$	2721.34
Sahu [15]	$(2n+2l+2)E_P+(3n+3l)T_P$	$(2l+3)E_P+(3l+6)T_P$	2338.54
Our scheme	$(3n+3l)T_E+(4n+4l+4)T_N$	$(2n+7l+2)T_E+(2n+7l+5)T_N$	1275.19

[8] K. Kim, D. Hong, and I. Jeong, "Identity-based proxy signature from lattices," *Journal of Communications and Networks*, vol. 15, no. 1, pp. 1-7, 2013.

[9] L. H. Li, S. F. Tzeng, M. S. Hwang, "Generalization of proxy signature based on discrete logarithms", *Computers & Security*, vol. 22, no. 3, pp. 245–255, 2003.

[10] X. Li and K, Chen, "ID-based multi-proxy signature, proxy multi-signature and multi-proxy multi-signature schemes from bilinear pairings," *Applied Mathematics and Computation*, vol. 169, pp. 437-450, 2005.

[11] E. J. L. Lu, M. S. Hwang, and C. J. Huang, "A new proxy signature scheme with revocation", *Applied Mathematics and Computation*, vol. 161, no. 3, PP. 799-806, Feb. 2005.

[12] Y. Lu and J. Li, "Provably secure certificateless proxy signature scheme in the standard model," *Theoretical Computer Science*, vol. 639, pp. 42-59, 2016.

[13] M. Mambo, K. Usuda, and E. Okamoto, "Proxy signatures: Delegation of the power to sign messages," *IEICE Transactions of The Fundamentals of Electronics, Communications and Computer Sciences*, vol. 79, no. 9, pp. 1338-1353, 1996.

[14] S. Mashhadi, "A novel non-repudiable threshold proxy signature scheme with known signers," *International Journal of Network Security*, vol. 15, no. 4, pp. 274-279, 2013.

[15] R. Sahu and S. Padhye, "An ID-based multi-proxy multi-signature scheme," *International Conference on Computer and Communication Technology*, pp. 60-63, 2010.

[16] A. Shamir, "Identity-based cryptosystem and signature scheme," in *Advances in Cryptology (Crypto'84)*, LNCS, vol. 196, pp. 47-53, 1984.

[17] N. Tiwari and S. Padhye, "Provable secure multi-proxy signature scheme without bilinear maps," *International Journal of Network Security*, vol. 17, no. 6, pp. 736-742, 2015.

[18] The Certicom Corporation, *SEC2: Recommended Elliptic Curve Domain Parameters*, Dec. 23, 2017. (`www.secg.org/collateral/sec2-final.pdf`)

[19] S. F. Tzeng, M. S. Hwang, C. Y. Yang, "An improvement of nonrepudiable threshold proxy signature scheme with known signers", *Computers & Security*, vol. 23, no. 2, pp. 174-178, Apr. 2004.

[20] S. F. Tzeng, C. Y. Yang, M. S. Hwang, "A Nonrepudiable Threshold Multi-Proxy Multi-Signature Scheme with Shared Verification", *Future Generation Computer Systems*, vol. 20, no. 5, PP. 887-893, June 2004.

[21] J. Xun H. SUN, Q. Wen and H. Zhang, "Improved certificateless multi-proxy signature," *The Journal of China Universities of Posts and Telecommunications*, vol. 19, no. 4, pp. 94-105, 2012.

[22] H. Xiong, F. Li and Z. Qin, "A provably secure proxy signature scheme in certificateless cryptography," *Informatica*, vol. 21, no. 2, pp. 277-294, 2010.

[23] L. Zhang, F. Zhang and Q. Wu, "Delegation of signing rights using certificateless proxy signatures," *Information Sciences*, vol. 184, pp. 298-309, 2012.

Biography

Rena Ehmet received her B.S. from Xinjiang University, Urumqi, China, in 2005; M.S. from Xinjiang University, Urumqi, China, in 2008. She is now a Ph.D candidate in the School of Mathematical Sciences, Xiamen University, Xiamen, China, and she is a teacher in the School of Mathematical Sciences, Xinjiang Normal University, Xinjiang, China. Her recent research interests include cryptography and information safety.

Lunzhi Deng received his B.S. from Guizhou Normal University, Guiyang, China, in 2002; M.S. from Guizhou Normal University, Guiyang, China, in 2008; and Ph.D. from Xiamen University, Xiamen, China, in 2012. He is currently a professor in the School of Mathematical Sciences, Guizhou Normal University, Guiyang, China. His recent research interests include cryptography and information security.

Yingying Zhang received her B.S. from Xinjiang Normal University, Urumqi, PR China, in 2011; M.S. from Xinjiang Normal University, Urumqi, China, in 2014. She is now a Ph.D candidate in the School of Mathematical Sciences, Xiamen University, Xiamen, China. Her recent research interests include cryptography and information safety.

Jiwen Zeng received his B.S. from Gannan Normal University, Ganzhou, China, in 1981; M.S. from Wuhan University, Wuhan, China, in 1988; and Ph.D from Peking University, Beijing, China, in 1995. He is currently a professor in the School of Mathematical Sciences, Xiamen University, Xiamen, China. His recent research interests include group and cryptography.

A Lightweight Authentication Protocol in Smart Grid

Debsmita Ghosh, Celia Li, Cungang Yang
(Corresponding author: Cungang Yang)
Department of Electrical and Computer Engineering, Ryerson University
350 Victoria St, Toronto, ON M5B 2K3, Canada
(Email: cungang@ee.ryerson.ca)
(Received Oct. 16, 2016; revised and accepted Feb. 21 & June 5, 2017)

Abstract

Smart grids allow automated meter readings and facilitate two-way communications between the smart meters and utility control centers. As the smart grid becomes more intelligent, it becomes increasingly vulnerable to cyber-attacks. Smart grid security mainly focuses on mutual authentication and key management techniques. An impeding factor in grid security is the memory and processing constraints of the smart meters. The aim of this paper is to propose a lightweight mutual authentication protocol between a residential smart meter and a gateway. The authentication protocol provides source authentication, data integrity, message confidentiality, and non-repudiation. The security analysis renders this protocol robust against several attacks. Its performance analysis provides meticulous results as to how the proposed protocol is efficient in terms of computation overhead, average delay and buffer occupancy at the gateway.

Keywords: Authentication Protocol; Key Management; Smart Grid

1 Introduction

The coexistence of the intelligent devices and the traditional power grid is termed as smart grid technology. Smart grid follows a distributed mode of control over the power system, as opposed to the centralized approach adopted by the traditional grid. The traditional power grid allows one-way electricity flow from a few power plants towards a large customer base. The NIST 3.0 framework, released in October 2014, mentions that the smart grid is the inclusion of communication and information technologies to the traditional power grid, and enabling duplex communication between smart meters and utility control centers [26]. If an active adversary is successful in obtaining and manipulating the meter readings, he may alter the readings to reflect incorrect usage. If this happens on a large-scale, it will significantly hamper the restricted energy resources and the economy as well.

A passive adversary, on the other hand, may collect reports for a long duration of time for a specific house. By analyzing the meter readings, the attacker will be able to understand the number of occupants in the house, the time at which the house is empty or the occupants are asleep, and other information describing the activity occurring inside the house. The attacker may use this information to launch an attack on the house. Hence, meter readings are extremely sensitive and must be protected. A limiting factor is the memory and processor capabilities of the smart meter device. For instance, a Home Area Network (HAN) smart meter configuration may comprise of MSP430-F4270 microcontroller along with 128 KB of flash and RAM memory [14]. Efficient protocols and mutual authentication schemes are already in use in the smart grid industry, but they also incur additional overhead. A few instances that increase overhead are long key sizes, ciphers and certificates, maintenance of Public Key Infrastructure (PKI), keeping track of Certificate Revocation Lists and timers. Furthermore, as the grid becomes smarter, it becomes increasingly vulnerable to software attacks. Smart meter devices depend on communication protocols such as TCP/IP, HTTP and FTP to exchange data. By default, these protocols do not have security built into them [3]. These conditions highlight the need for a lightweight authentication protocol between smart meters.

2 Related Work

Extensive research is being conducted in devising lightweight approaches using techniques such as Diffie-Hellman, ECC-based cryptography and ID-based cryptography. H. So proposes a zero-configuration signcryption protocol to ensure safe and secure communications between two ends [29]. The communication overhead for encryption and signature schemes of the protocol increases with the degree of encryption. Also, the security

level of the signature is directly proportional to the degree of encryption. The advantage of this protocol is it doesn't use asymmetric key algorithms. This protocol assures key protection but it is too expensive, keeping in mind that several smart meters generate packets every 15 minutes.

Nicanfar *et al.* proposes a mutual authentication scheme between a HAN smart meter and an authentication server [25]. They use Secure Remote Password protocol (SRP) and decrease the number of steps in SRP from five to three. The proposed protocol also reduces the number of exchanges from four to three. It is essentially based upon Enhanced ID-based Cryptography (EIBC). EIBC essentially uses True Random Number Generator and Pseudorandom Number Generator to keep changing the secret master key, along with the public/private keys of the meters. The paper is robust against several attacks. Also, the key renewal mechanism is efficient in terms of refreshing the public/private and multicast keys. However, it requires synchronization of three timers between the smart meter and authentication server. This adds to the overhead of the protocol. Also, as mentioned previously, having two random number generators means memory consumption for storing the generator states. Hence, the protocol doesn't favor scaling of the smart grid environment.

Fouda *et al.* proposes a lightweight mutual authentication protocol and generates a shared session key on the basis of computational Diffie-Hellman exchange protocol [8]. The protocol is applicable between the HAN smart meters and Building Area Networks (BAN) gateway, each of which have a public/private key pair issued by a certificate authority. The paper describes the protocol steps after the HAN smart meter and BAN gateway have extracted and verified their certificates. Fouda *et al.* use computational Diffie-Hellman scheme to establish mutual authentication. The generated shared session key is then combined with hash-based authentication code techniques to authenticate messages between the two entities. The proposed protocol is successful in establishing a semantically secure shared key in the mutual authentication environment. The main disadvantage of this protocol is usage of RSA protocol to establish authentication. The involvement of certificate authority and certificate revocation lists is a costly process for limited devices like HAN smart meters.

Li *et al.* propose a protocol that uses homomorphic encryption to attain secure demand response exchanges in a smart grid environment [15]. The protocol achieves forward secrecy, by renewing the users' key after appropriate intervals. It also achieves entity authentication, and message integrity and confidentiality. Homomorphic encryption is a method in which plaintext is encrypted using algebraic expression. They combine homomorphic encryption with pairing-based cryptography to create the mutual authentication process. In this paper, authentication process is applicable between control center and BAN, as well as between HAN and BAN. Once two entities have successfully established a session key between each other, then message exchange commences. The messages are signed using ID-based signature mechanism. A drawback of the protocol is the absence of explicit key confirmation. As the session key is generated separately at the two entities, it is advisable to confirm the key before commencing message exchange.

3 Background Knowledge

3.1 Topology of Smart Grid

The topology of the smart grid has been adapted from the NIST Conceptual Reference Model for Smart Grid that is shown in Figure 1. Smart grid architecture consists of four main domains: generation, transmission, distribution and consumers. The generation domain consists of the large-scale power plants and small-scale DERs that generate electricity. This is followed by the transmission domain consisting of step-up transformers (transmission voltage), transmission substations and transmission lines that aid in transmitting the electricity to the next domain, the distribution domain. The distribution domain consists of step-down transformers (distribution voltage and service voltage), distribution substations and distribution lines.

Lastly, the consumers include the smart meters at the homes or businesses that directly use electricity. Consumers may be residential or commercial. Electrical sensors and circuit breakers are placed along the entire length of the communication medium between smart meters and generators to constantly monitor voltage and flow. Smart meters also have a hierarchy of their own. The lowest level of the hierarchy consists of the meters installed at the home/business, and is called the HAN smart meter. Several HAN smart meters regularly send their meter readings to a designated BAN gateway, which is the next level in the hierarchy. Lastly, a number of BAN gateway send the collection of meter readings to the Neighborhood Area Network (NAN) gateway. The NAN gateway then forward these meter readings to the utility center. The utility centers are located in the distribution substations.

3.2 Smart Grid Communications

The communication technology used in the smart grid is a combination of wireless and wired technology [7,28]. The generation and transmission domains are entirely based on wired technology such as optical fiber or power line carriers (PLC). Optical fiber technology is advantageous because it is flexible, suitable for the core network, and capable to carry high volume of traffic with the least latency [8]. The consumer domain favors wireless technology to communicate with the distribution domain. The distribution domain consists of wireless technology at the end connected to the consumer domain, and wired technology for the end connected to the transmission domain.

The Smart Grid environment primarily consists of three areas - HAN, BAN and WAN. The potential technologies for HAN are ZigBee, Wi-Fi, Ethernet, Z-Wave

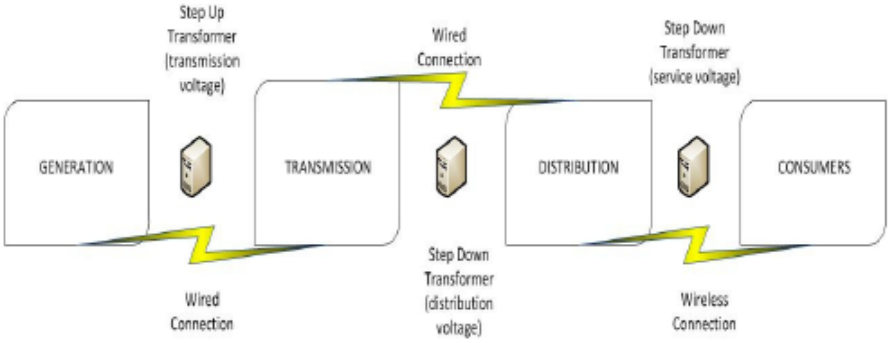

Figure 1: Reference model for smart grid

and PLC. ZigBee is preferred over other wireless technologies such as Wi-Fi or Bluetooth because it consumes the least amount of power and delivers high performance. In BAN, ZigBee, Wi-Fi, PLC and cellular technologies may be used. Wi-Fi and WiMAX are preferred over PLC because they are cost-effective and flexible. As the coverage distance is in tens of kilometers for WAN, potential technologies are Ethernet, microwave, WiMAX, 3G/LTE, and fiber optic links. Wired technologies are typically favored in the WAN connections because wired connections are more robust and secure compared to the wireless connections [2].

3.3 ID-based Cryptography (IBC)

This technique replaces traditional digital certificates with unique identifying attributes, such as email addresses or phone numbers, for encryption and signature verification [21]. IBC replaces the certificate authority with a Private Key Generator (PKG). Before the system nodes enter into mutual authentication processes with one another, the PKG generates a master private-public key pair. The master public key is distributed to all the system nodes. The following procedure describes the encryption and decryption using IBC:

Encryption Process: Node A uses Node B's identifier and the master public key to encrypt Message M. This produces the cipher text C. Node A sends C to Node B. The ease of using IBC is that Node A did not have to make prior arrangements to be able to send a message to Node B, unlike in the traditional certificate process.

Decryption Process: Upon receiving C, Node B contacts PKG to get its secret private key to decrypt C. The PKG then transmits Node B's private key to it over a secure channel. This secure channel may be an SSL link that allows Node B to download its private key. Node B is now able to successfully decrypt C to obtain plaintext M.

Signature: Node A wants to send a signed message to Node B. Upon receiving its private key from the PKG, Node A creates a signature S for Message M and sends it to Node B, along with the plaintext Message M. Signature ensures data integrity as well as non-repudiation of a message. In other words, because the message is signed with a private key and private keys are secret, hence the sender cannot deny having sent the signed message.

Verification: Upon receiving M and S from Node A, Node B applies Node A's identifier and the master public key on M. If the generated signature is the same as S, then Node B accepts the Message M. Else, it rejects Message M.

3.4 Bilinear Pairing

Bilinear degenerate maps are mathematical functions, which when used in combination with ID-based cryptography, produces computationally efficient cryptographic systems [12, 18]. A bilinear map is a pairing function which produces a mapping of elements from one cyclic group to another cyclic group, provided both cyclic groups is of the same prime order [9,10]. The discrete log problem of the first group is hard. Bilinear maps are considered to be secure because they are chosen as one-way functions. In other words, it is easy to calculate the result from a known set of pair of elements, but it is hard vice-versa.

3.5 Zero-knowledge Password Proof

Zero-knowledge password proof (ZKPP) is a technique in which Node A (prover) proves to Node B (verifier) that it possesses knowledge of a password without actually knowing the password [1]. This possession of knowledge about the password works as a verification that the node may be trusted. The password belongs to Node B and never leaves Node B. Node B generates a verifier related to this password and conveys this verifier to all nodes it wants to communicate with. This technique works as an advantage for systems using password-authenticated key agreement (PAKE) protocol because it is robust against off-line dictionary attacks, as is mentioned in IEEE P1363.2. In IEEE P1363.2, ZKPP is defined as "An interactive zero knowledge proof of knowledge of password-driven data shared between a prover and the corresponding verifier."

3.6 Secure Remote Password Protocol

Secure Remote Password Protocol (SRP) is also a modified password-authenticated key agreement protocol [5]. SRP is more secure than SSH protocol, and faster than Diffie-Hellman key exchange in terms of user authentication and data integrity [13,16,17]. Compared to Kerberos protocol, SRP doesn't rely on third parties. The SRP is instantiated with the client node selecting a small random salt. The client node also shares a password with the server node [11, 20, 24]. At the end of the exchange protocol, the client and server nodes now have a symmetric session key. The two nodes need to explicitly confirm that their keys match in order to complete authentication process.

4 Proposed Mutual Authentication Protocol

The proposed protocol ensures a lightweight mutual authentication and key renewal mechanism between the HAN smart meter and the BAN gateway. It also provides confidentiality, authentication and integrity; the three essential requirements for Smart Grid security mentioned by NIST [26].

4.1 Pre-authentication Protocol

Let G be an additive cyclic group of prime order q, and GT be a multiplicative cyclic group of prime order q; let g be the generator of these cyclic groups. In order to build cryptographic systems, pairing-based cryptography utilizes a symmetric bilinear pairing between two elements of an additive group to an element of a multiplicative group [19]. In the proposed protocol, we have used the map e: $G1 \times G2 = GT$, where $G1 = G2 = G$. This mapping satisfies the properties stated below:

Bilinearity: $\forall x, y \in Z_q*$, $\forall A, B \in G : e(A^x, B^y) = e(A, B) \in G_T$;

Non-degeneracy:] $e(A, B) \neq 1$;

Computability:] There exists an efficient algorithm to compute e.

Let there be a bilinear parameter generator which runs an algorithm that takes in as input a security parameter L, and outputs the system's 5-tuple (q, g, G, GT, e).

In the proposed protocol, the BAN gateway also acts as the public key generator (PKG). Hence, as is the function of the PKG, the BAN gateway determines the 5-tuple (q, g, G, GT, e) by providing input L to the bilinear parameter generator. The BAN gateway randomly chooses a master secret key s that belongs to Z_q*. It does not convey the master secret key to any other entity. The cryptographic secure hash functions are determined by the BAN gateway. This protocol utilizes 5 hash functions:

$$H_1(\cdot) : (0,1)* \times (0,1)* \rightarrow (0,1)*$$
$$H_2(\cdot) : (0,1)* \rightarrow G*$$
$$H_3(\cdot) : Z_N* \times G* \rightarrow G*$$
$$H_4(\cdot) : G_T* \rightarrow Z_q*$$
$$H_5(\cdot) : Z_q* \times Z_q* \times G_T* \rightarrow G*.$$

Two messages are exchanged prior to commence of the mutual authentication protocol between the smart meters which is shown in Figure 2. Owing to its steady rise in popularity, WMN is considered as the communication protocol running between the HAN and BAN smart meters.

Each smart meter/gateway bears a unique identifier. Also, each HAN smart meter contains a password its corresponding verifier, which is required to execute the Zero Knowledge Password Proof. The first message conveys the identifier and the verifier of a HAN smart meter (HAN SM) to the BAN gateway (BAN GW) through a secure channel [6]. On receiving the message, the BAN GW stores the received identifier and verifier in its memory. A BAN GW has ten times more memory than a HAN [8]. After receiving the first message, the BAN GW functions as the public key generator. In other words, the BAN GW uses hash function H1 to generate the public key for the HAN SM. Next, it applies its master secret key on this newly-generated public key to create the HAN SM's private key. In this manner, the HAN SM initiates the authentication process between itself and the BAN GW.

The second message serves as an acknowledgement for the HAN SM having sent its authentication request to the designated BAN GW. The private key as well as the public system parameters (q, g, G, GT, e, H_1, H_2, H_3, H_4, H_5) are conveyed via the message from the BAN GW to the HAN SM through a secure channel [6]. Once the HAN SM receives the message, it stores its private key and the public parameters. It then uses hash function H1 from the list of public parameters to create the public key for the BAN GW using the identifier sent by the BAN GW.

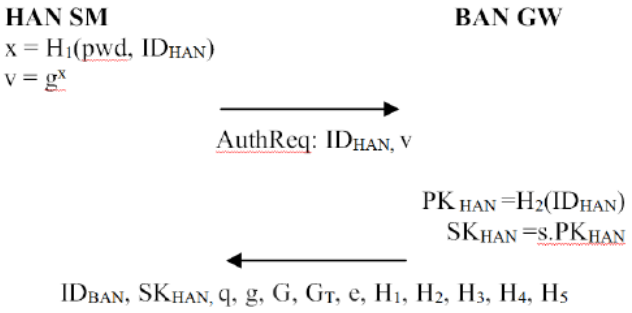

Figure 2: Pre-authentication phase exchange

4.2 Authentication Protocol

In Figure 3, the HAN SM randomly choose a number Z_q*. The HAN SM generates the public key of the BAN as well as variable A using a random number. The HAN SM encrypts A using the public key of the BAN GW. The first message contains the HAN SM's identifier and the encrypted value of A. At the BAN GW, the variable A is first decrypted using the BAN GW's private key. The identifier sent by the HAN SM is used to lookup the verifier corresponding to that identifier. The BAN GW will not be successful if it has the wrong verifier corresponding to a HAN SM identifier. This is because the verifier is derived from the password associated with that specific HAN SM. After storing the value of A and locating the verifier corresponding to the identifier of the HAN SM, the BAN GW then chooses a random number that belongs to Z_q*, which is used to compose B which is utilized in calculating variable J. Variable B also requires k. Referring the Secure Remote Password Protocol, k is derived by applying hash function $H3$ on N and g. N is a safe prime which is equal to $2q+1$, where q is the prime order of the cyclic groups used in pairing-based cryptography; g is the generator of these cyclic groups.

The variable k is calculated by both sides HAN and BAN. Furthermore, when an active adversary impersonates a smart meter, then variable k helps to eliminate 2-for-1 guess. The next step for the BAN GW is to use these values to calculate variable B and variable J. Variable J holds the bilinear pairing map using the private key of the HAN SM, variable A received from HAN SM, and the random number selected by the BAN GW itself. Variable A is sent to the BAN GW in an encrypted manner. Variable J helps enforce one half of the pairing based cryptography because it stores the bilinear mapping at the BAN GW. After J is calculated, W is also devised by applying hash function on J. BAN GW then sends B and W to the HAN SM. On receiving the second Message from the BAN GW, the HAN SM stores B. As mentioned above, k is calculated once again.

The variable J' is constituted of B and other parameters. J' forms the other half of the bilinear pairing because it comprises of the bilinear mapping held at the HAN SM. Based on the properties of bilinear pairing, variables J and J' have to be equal in order for the HAN SM to be able to authenticate the BAN GW. Hence, if W and W' are not equal to each other, then that reflects inconsistencies in its constituent variables resulting in the abortion of the authentication process. If W is equal to W' then the HAN SM authenticates the BAN GW. To complete the mutual authentication protocol, the BAN GW should also trust the HAN SM. To do so, the HAN SM introduces a valid-period and a sequence number initialized to 0. It then forms the first session key, $K1$, by applying a hash function on the valid-period, J' and the sequence number.

Furthermore, this session key is signed with the private key of the HAN SM. In Message 3, the sequence number, valid-period and the signed session key, Signk1, are sent to BAN GW. Upon receiving Sign $k1$, the BAN GW extracts the key, $K1$, using the identifier of the HAN SM on the cipher text received in Message 3. Using the received valid-period and sequence number and the previously calculated J, the BAN GW calculates a session key k' at its end. This key K' is then compared with $K1$. If they are the same, then BAN GW authenticates HAN SM as well.

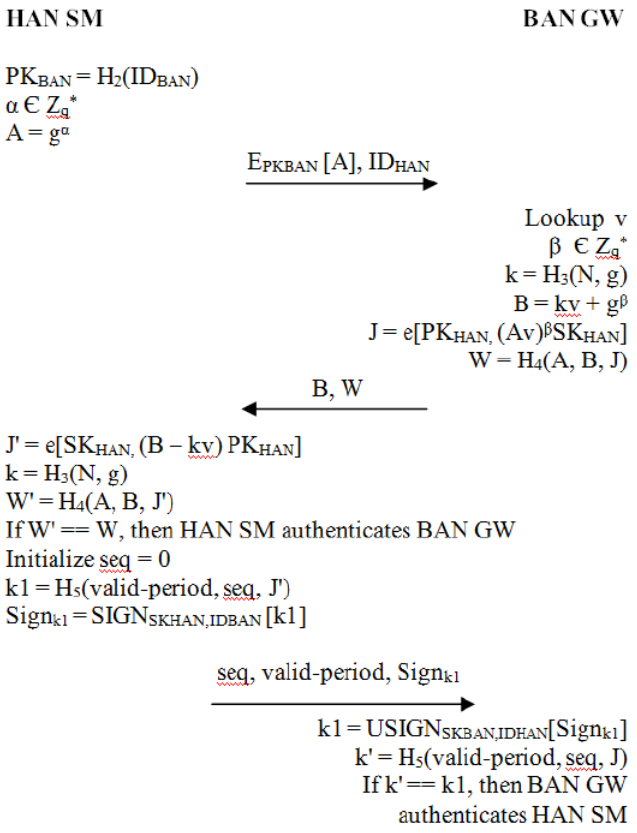

Figure 3: Mutual authentication protocol

Being an public-key based authentication protocol, the proposed authentication protocol provides data integrity, message confidentiality, and non-repudiation.

5 Security Analysis

The proposed protocol is resistant to several attacks. Here, we assume that an active or a passive attack can be made. A passive attacker may observe all exchanges between two smart meters. An active attacker, on the other hand, may make changes to the actual messages. He may further enters into a smart meter and takes control of its operation.

Replay Attack: The protocol is resistant against the replay attack because the adversary does not know the private key of the BAN smart meter, it cannot decrypt the value of A. Owing to the one-directional nature of hash functions, it is extremely difficult and

time-consuming to feed the hash function with random inputs until the known output is achieved. Also, the adversary has no previous knowledge about the password verifier and random value A.

Man-in-the-Middle Attack: A combination of random values, passwords, hashed messages and keys create a strong mutual protocol. Furthermore, the HAN smart meter's private key is exchanged through a secure channel. The attacker will not be able to decrypt or verify any signed messages. The advantage of the proposed protocol lies in the fact that if a message contains a sensitive value, then that is encrypted or signed. Otherwise, the message contains a hash of a value, which can only be verified by being recomputed at the sender's end.

Known Session Key Attack: After an exchange of verifier, random values and keys, the sender and receiver use a different combination of variables and apply a hash function on it. These constituent variables are complex and difficult to guess. The HAN smart meter uses a combination of its private key and the BAN smart meter's public key, along with random value B. The BAN smart meter uses its private key, the HAN smart meter's public key, and random variable A. Both of these values give the same session key but with different set of variables.

Impersonation Attack: During the authentication process, the adversary attempting such an attack will always be unsuccessful owing to several factors. Firstly, the adversary might have access to the HAN smart meter's identifier and public key. But owing to ZKPP, the impersonating attacker will not know the password or the corresponding verifier of the actual HAN SM. Since the mutual authentication depends on the password and its related verifier, hence impersonation attack will definitely not succeed.

Key Control Attack: In the proposed protocol, the session key and mutual authentication process depends on the verifier and its corresponding password, and also between two random numbers exchanged between the two smart meters. The key is calculated individually at both the ends and then verified. Hence, it does not depend on one end alone. Furthermore, bilinear mapping involves using the public and private keys of the entities.

6 Performance Analysis

6.1 Computational Costs

This section evaluates the performance of the proposed protocol. We compare our proposed protocol with a smart grid mutual authentication protocol put forward by Nicanfar *et al.* [4]. The protocol proposed by Nicanfar *et al.* is an efficient mutual authentication scheme between a HAN smart meter and an authentication server. Notations used in this section are shown in Table 1.

Table 1: Parameter notations for performance analysis

T_{bm}	Latency of a bilinear map operation
T_{mul}	Latency of a scalar multiplication operation
T_{add}	Latency of an addition operation
T_{sub}	Latency of a subtraction operation
T_{xor}	Latency of an XOR operation
T_{exp}	Latency of a modular exponentiation operation
T_{pow}	Latency of a power operation
T_h	Latency of a hash operation

Table 2: Computational costs of the proposed protocol

	Proposed Protocol
HAN Side	$5T_h + 2T_{exp} + 1T_{bm} + 1T_{mul} + 1T_{sub}$
BAN Side	$4T_h + 1T_{exp} + 1T_{bm} + 3T_{mul} + 1T_{sub}$ $+1T_{add}$
BAN Side	

Table 3: Computational costs of Nicanfar *et al.* protocol

	Nicanfar *et al.* protocol
SM Side	$10T_h + 2T_{exp} + 1T_{mul} + 1T_{sub}$ $+1T_{xor} + 1T_{add} + 1T_{pow}$
SAS Side	$8T_h + 1T_{exp} + 3T_{mul} + 1T_{bm}$ $+1T_{add} + 2T_{pow}$

The computational costs of our protocol and Nicanfar *et al.* are shown in Tables 2 and 3. To summarize, in regards to the HAN side/SM side, the proposed protocol uses less number of operations compared to the protocol proposed by Nicanfar *et al.* The proposed protocol uses 5 hash operations, whereas the other protocol uses 10 hash operations. Hash operations are one of the least computationally intensive operations, but keeping in mind the memory constraints of the HAN smart meter, it is best to save memory and CPU power under any circumstances. The number of scalar multiplication, modular exponentiation and subtraction operations used in both the protocols is the same. Coming to the BAN side/SAS side, the proposed protocol uses 4 hash function operations and 2 power operations, as compared to the protocol proposed by Nicanfar *et al.*, which uses 8 hash functions and 3 power operations. The number of addition, scalar multiplication and modular exponentiations between the two protocols is the same. In case of similar operations, the protocol proposed in this thesis is always using a lesser

number of operations as compared to the other protocol. As is shown in Table 4 the proposed protocol ensures mutual authentication with one encryption/decryption operation, and one sign/verify operation.

Table 4: Encryption/Signature in the proposed protocol

	HAN Side	BAN Side
Encryption/Decryption	1 Encryption	1 Decryption
Sign/Verify	1 Sign	1 Verify

Table 5: Encryption/Signature in Nicanfar *et al.* protocol

	SM Side	SAS Side
Encryption/Decryption	1 Encryption 1 Decryption	1 Decryption 1 Encryption
Sign/Verify	1 Sign 1 Verify	1 Verify 1 Sign

On the other hand, Table 5 shows that the other protocol has 2 encryption/decryption operations, and 2 sign/verify operations. As these operations are expensive for the resource-constrained smart meters, hence the proposed protocol proves to be more lightweight by using the common principles of SRP protocol, ID-based cryptography, bilinear mapping and ZKPP.

6.2 Simulation Results

The following subsection compares the performance of the proposed protocol with the Elliptic Curve Digital Signature Algorithm (ECDSA). The simulation parameters are shown in Table 6. As the proposed protocol uses AES-128 and SHA-256, therefore its equivalent algorithm ECDSA-256 has been chosen. The reason for choosing ECDSA is that it is a standardized algorithm, already in practice in the smart grid environment. MATLAB [22] and OpenSSL [27] have been used to generate results which depict that the proposed protocol is a better alternative.

Communication Overhead: The communication overhead of the proposed protocol and ECDSA is shown in Figure 4. When the number of HAN smart meters is 50, the communication overhead experienced by the BAN smart meter is around 100 KB for the proposed protocol, and around 300 KB for the ECDSA algorithm. As the number of HAN smart meters increase to 125, the disparity between these two methods increase further. ECDSA algorithm has a communication overhead close to 775 KB. On the other hand, the proposed protocol displays a communication overhead of less than 300 KB. In the last scenario, when the number of HAN smart meters is 250, the communication overhead for ECDSA algorithm is almost 1550 KB, whereas the proposed protocol consumes 600 KB. This value is less than half of the communication overhead experienced in the ECDSA algorithm. Hence, with an increasing number of smart meters at the HAN, the communication overhead will increase greatly for the ECDSA algorithm. This may act as a barrier in further expansion of the smart meter network.

Table 6: Simulation parameters

Simulation Parameters	Value
Interval of Message Generation	Once every hour
Simulation Time	24 Hours
TCP Header	20 Bytes
IPv4 Header	20 Bytes
Ethernet Header	26 Bytes
Payload	32 Bytes
SHA-256 Header	32 Bytes
ECDSA Signature Size	64 Bytes
ECDSA Certificate Size	125 Bytes
Number of HAN Smart Meters	Maximum of 250 per BAN Gateway

Average Delay: Average delay refers to the mean time take to perform decryption/signature verification for the cipher text. In Figures 5 and 6, the BAN gateway experiences an average delay of 0.075 seconds for 250 smart meters, each generating meter reports once every hour over 24 hours. Considering the highest number of smart meters in this simulation, which is 250, the proposed protocol experiences an average delay of 0.0085ms for 50 smart meters. Similarly, for 175 and 250 smart meters, the average delay generated is 0.095 ms and about 0.01 ms respectively. On the other hand, the ECDSA algorithm experiences an average delay of 0.05 seconds for 250 smart meters. This comparison shows the wide difference between the ECDSA algorithm and the proposed protocol. Furthermore, RSA displays a very high delay of 0.075 seconds for the same number of smart meters.

Considering the highest number of smart meters in this simulation, which is 250, the proposed protocol experiences an average delay of 0.0085ms for 50 smart meters. Similarly, for 175 and 250 smart meters, the average delay generated is 0.095 ms and about 0.01 ms respectively. On the other hand, the ECDSA algorithm experiences an average delay of 0.05 seconds for 250 smart meters. This comparison shows the wide difference between the ECDSA algorithm and the proposed protocol. Furthermore, RSA displays a very high delay of 0.075 seconds for the same number of smart meters.

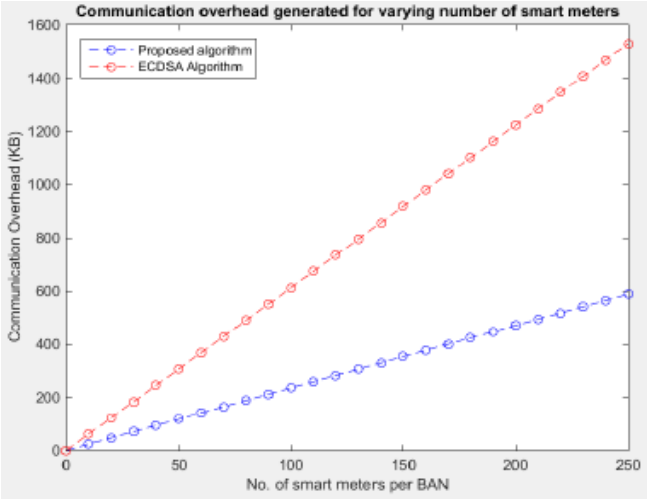

Figure 4: Communication overhead

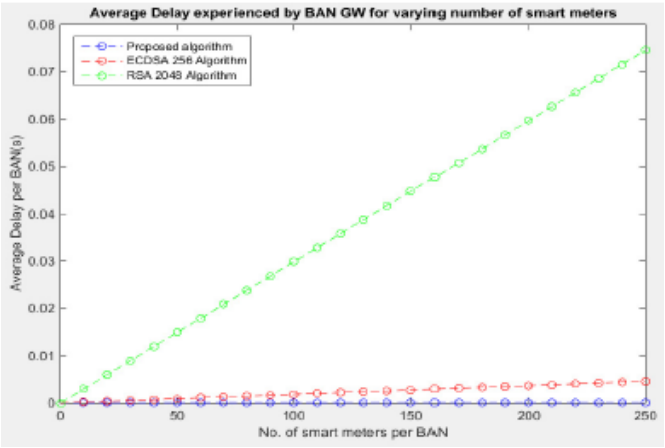

Figure 5: Average delay experienced for proposed protocol

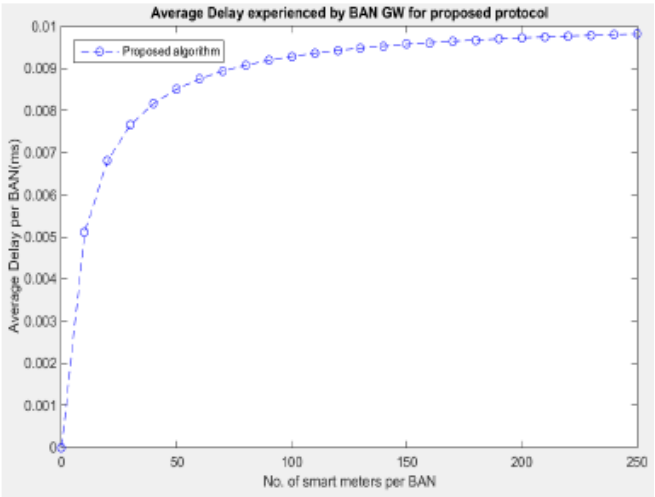

Figure 6: Average delay experienced by BAN GW

7 Conclusion

The proposed authentication protocol describes an efficient lightweight scheme to provide mutual authentication between the HAN smart meter and BAN gateway. This scheme provides source authentication, data integrity, message confidentiality, and non-repudiation as well. The proposed protocol is secure against Replay, Man-in-the-Middle, Known Session Key, Impersonation and Key Control Attacks. On comparison with the efficient mutual authentication protocol proposed by Nicanfar *et al.* [25], the proposed protocol utilizes lesser number of computation operations, while achieving the same results in terms of message security. To be specific, the main difference between these two protocols is that the proposed scheme uses Pairing-based Cryptography, whereas the other protocol uses Enhanced ID-based Cryptography (EIBC). In addition, the proposed protocol is compared with ECDSA, which is currently used in smart meter authentication. The parameters of comparison between these two schemes are as follows: communication overhead, average delay, and buffer occupancy. In each case, the proposed protocol proves to be more lightweight and efficient. Firstly, the proposed protocol incurs a communication overhead of 98 bytes, whereas ECDSA incurs 255 bytes (mainly owing to the ECDSA signature and certificate). Secondly, the BAN gateway has an average delay of 0.01ms and 0.05 seconds in the proposed protocol and the ECDSA scheme respectively. Lastly, using ECDSA, the BAN gateway exhausts its 1128 KB buffer while handling an incoming message rate of 100 messages every 15 minutes across a simulation period of 8 hours. On the other hand, the proposed protocol consumes 775 KB in the same simulation environment. Hence, proving that it is scalable as well as lightweight.

References

[1] A. Ahmed, A. Younes, A. Abdellah, Y. Sadqi, "Strong zero-knowledge authentication based on virtual passwords," *International Journal of Network Security*, vol. 18, no. 4, pp. 601-616, 2016.

[2] M. Badra, S. Zeadally, "An improved privacy solution for the smart grid," *International Journal of Network Security*, vol. 18, no. 3, pp. 529-537, 2016.

[3] M. Balakrishnan, *Security in Smart Meters*, Aug. 2012. (`https://cache.freescale.com`)

[4] M. Balakrishnan, M. Mienkina, *Designing Smart Meters for the Smart Grid*, Jan. 6, 2018. (`http://docplayer.net/27262011-Designing-smart-meters-for-the-smart-grid.html`)

[5] T. Y. Chang, W. P. Yang, M. S. Hwang, "Simple authenticated key agreement and protected password change protocol", *Computers & Mathematics with Applications*, vol. 49, pp. 703–714, 2005.

[6] C. Chou, K. Tsai, C. Lu, "Two ID-based authenticated schemes with key agreement for mobile envi-

ronments," *Journal of Supercomputing*, vol. 66, no. 2, pp. 973-988, Nov. 2013.

[7] C. Chrysoulas, "Shielding the grid world: An overview," *International Journal of Electronics and Information Engineering*, vol. 1, no. 1, pp. 23-28, 2014.

[8] M. M. Fouda, Z. M. Fadlullah, N. Kato, L. Rongxing, S. Xuemin, "A lightweight message authentication scheme for smart grid communications," *IEEE Transactions on Smart Grid*, vol. 2, no. 4, pp. 675-685, Dec. 2011.

[9] X. Fu, X. Nie, and F. Li, "Outsource the ciphertext decryption of inner product predicate encryption scheme based on prime order bilinear map," *International Journal of Network Security*, vol. 19, no. 2, pp. 313-322, 2017.

[10] P. Hu, H. Gao, "A key-policy attribute-based encryption scheme for general circuit from bilinear maps," *International Journal of Network Security*, vol. 19, no. 5, pp. 704-710, 2017.

[11] M. S. Hwang, C. C. Lee, Y. L. Tang, "An improvement of SPLICE/AS in WIDE against guessing attack", *International Journal of Informatica*, vol. 12, no. 2, pp.297–302, Apr. 2001.

[12] B. King, "A dynamic threshold decryption scheme using bilinear pairings," *International Journal of Network Security*, vol. 17, no. 6, pp. 771-778, 2015.

[13] P. Kuacharoen, "An anti-phishing password authentication protocol," *International Journal of Network Security*, vol. 19, no. 5, pp. 711-719, 2017.

[14] H. Li, *Enabling Secure and Privacy Preserving Communications in Smart Grids*, Springer, Aug. 2015.

[15] H. Li, X. Lin, H. Yang, X. Liang,R. Lu, X. Shen, "EPPDR: An efficient privacy-preserving demand response scheme with adaptive key evolution in smart grid," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 8, pp. 2053-2064, Apr. 2013.

[16] C. W. Lin, C. S. Tsai, M. S. Hwang, "A new strong-password authentication scheme using one-way hash functions", *International Journal of Computer and Systems Sciences*, vol. 45, no. 4, pp. 623–626, July 2006.

[17] C. H. Ling, C. C. Lee, C. C. Yang, M. S. Hwang, "A secure and efficient one-time password authentication scheme for WSN," *International Journal of Network Security*, vol. 19, no. 2, pp. 177-181, 2017.

[18] J. Ling, G. Zhao, "An improved anonymous password authentication scheme using nonce and bilinear pairings," *International Journal of Network Security*, vol. 17, no. 6, pp. 787-794, 2015.

[19] N. Liu, J. Chen, L. Zhu, J. Zhang, Y. He, "A key management scheme for secure communications of advanced metering infrastructure in smart grid," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 10, pp. 4746-4756, Aug. 2013.

[20] Y. Liu, C. C. Chang and S. C. Chang, "An efficient and secure smart card based password authentication scheme," *International Journal of Network Security*, vol. 19, no. 1, pp. 1–10, 2017.

[21] C. March and C. Youngblood, *An Introduction to Identity-based Cryptography*, Jan. 6, 2018. (`https://www.researchgate.net`)

[22] MatLab, *MATLAB and Statistics Toolbox Release 2014b*, Jan. 6, 2018. (`https://www.mathworks.com`)

[23] J. Menezes, O. P. C. Van, S. A. Vanstone, *Handbook of Applied Cryptography*, Boca Raton: CRC Press, 1997.

[24] J. Moon, D. Lee, J. Jung, D. Won, "Improvement of efficient and secure smart card based password authentication scheme," *International Journal of Network Security*, vol. 19, no. 6, pp. 1053-1061, 2017.

[25] H. Nicanfar, P. Jokar, K. Beznosov, V. C. M. Leung, "Efficient authentication and key management mechanisms for smart grid communications," *IEEE Systems Journal*, vol. 8, no. 2, pp. 629-640, July 2014.

[26] NIST, *NIST Framework and Roadmap for Smart Grid Interoperability Standards, Release 2.0*, NIST Special Publication 1108R2, Aug. 2015.

[27] OpenSSL, *OpenSSL*, Jan. 6, 2018. (`https://www.openssl.org/`)

[28] R. Singh and M. S. Manu, "An energy efficient grid based static node deployment strategy for wireless sensor networks," *International Journal of Electronics and Information Engineering*, vol. 7, no. 1, pp. 32-40, 2017.

[29] H. K. H. So, S. H. Kwok, E. Y. Lam, and K. S. Lui, "Zero-configuration identity-based signcryption scheme for smart grid," in *Proceedings of First IEEE International Conference on Smart Grid Communications*, pp. 321-326, Oct. 2010.

Authors' Biographies

Debsmita Ghosh completed her MASc student in computer networks at Ryerson University. Her research interest is on security and privacy of smart grid and power system computer networks.

Celia Li completed her Ph.D degree in electrical engineering and computer science department in 2015 at York University. Her research is focused on security and privacy, role-based access control and wireless mesh network security.

Cungang Yang completed his Ph.D degree in computer science in 2003 at University of Regina, Canada. In 2003, he joined the Ryerson University as an assistant professor in the Department of Electrical and Computer Engineering. His research areas include security and privacy, enhanced role-based access control model, information flow control, web security and secure wireless networks.

Soft Biometrics Traits for Continuous Authentication in Online Exam Using ICA Based Facial Recognition

Prakash Annamalai[1], Krishnaveni Raju[1], and Dhanalakshmi Ranganayakulu[2]
(Corresponding author: A. Prakash)
Department of Computer Science and Engineering, Hindustan Institute of Technology and Science[1]
Old Mahabalipuram Road, Padur 603103, India
(Email: prakash1712@yahoo.com)
Department of Computer Science and Engineering, KCG College of Technology, Chennai, India[2]
(Received Nov. 8, 2016; revised and submitted Feb. 1, 2017)

Abstract

Biometric authentication has been getting widespread attention over the past decade with growing demands in automated secured personal identification. Continuous Authentication (CA) system verifies the user continuously once a person is logged in. Continuous Authentication system prevents the intruders from invoking the system. A new framework for continuous user authentication that primarily uses hard and soft biometric traits using Independent Components Analysis (ICA) dimension reducing method for video frames. The proposed framework automatically registers (enrolls) soft biometric traits every time the user logs in and fuses soft biometric matching with the conventional authentication schemes. Different soft biometrics are considered to obtain the matching score value, here optimize the soft biometrics weights using Grey Wolf Optimization (GWO) technique is used. Finally the authentication is performed and evaluated using standard evaluation metrics then produce the maximum accuracy compared to existing methods.

Keywords: Biometric Traits; Continuous Biometric Authentication; Face Recognition; Online Exam Authentication

1 Introduction

In a modern life personal authentication is a common concern to both industries and academia due to its numerous applications such as physical access control, computer security, banking, airport, computer system login, and mobile phones law enforcement, etc [16, 17, 21]. Biometric measurement is a key component of several personal authentication systems that only render services to legitimately enrolled users [8, 10, 12]. The most known and often used modalities are fingerprints, face, hand geometry, knuckle print, palm and iris. These are widely deployed in large-scale systems such as border control and biometric passports [2, 6, 20]. Biometric information stored in a database may leak biometric features which can be used to reconstruct a biometric image [1]. Biometric traits are difficult to counterfeit and hence results in higher accuracy when compared to other methods such as using passwords and ID cards [7]. Biometrics identifies the person by what the person is rather than what the person carries, unlike the conventional authorization systems like smart cards [9, 13]. Hand-based person identification provides a reliable, low-cost and user-friendly viable solution for a range of access control applications. Palm print is one of the relatively new hand-based biometrics due to its stable and unique characteristics [15]. Authentication with respect to fingerprints implies that recognition is based on matching the features of a live fingerprint against those of fingerprints that are already stored in a server database. In addition, a digital signature of a fingerprint can be used for reliability [18].

2 Literature Review

In 2014, Gao *et al.* [3] Had proposed the Competitive Coding (Comp Code) scheme, which extracts and codes the local dominant orientation as features, had been widely used in Finger Knuckle Print (FKP) verification. However, Comp Code may lose some valuable information such as multiple orientation and texture of the FKP image. To remedy the above drawback, a novel multiple orientation and texture information integration scheme is proposed in the process. As compared with Comp Code, the proposed scheme not only considers more orientations, but also introduces a multilevel image thresholding scheme to perform orientation coding on each Gabor filtering response. For texture features extraction, LBP

maps are first obtained by performing Local Binary Pattern (LBP) operator on each Gabor filtering response, and then a similar coding scheme is applied on these LBP maps.

In 2014, Gupta *et al.* [4] had proposed an efficient algorithm to segment all finger tips from a slap-image and to identify them into their corresponding indices i.e. index, middle, ring or little finger of left/right hand. Geometrical and spatial properties have been used to identify these fingertips. The proposed algorithm can handle various challenges like the presence of dull prints, large rotational angles of the hand, small variation in the orientation of the fingertips and non-elliptical shape of components. It has been tested on a database of 6732 images of 1122 subjects. Experimental results reveal the segmentation of all fingertips from slap-images with an accuracy of 99.02%.

Tan *et al.* [19] had presented that discrimination of Used Frying Oil (UFO) from Edible Vegetable Oil (EVO), the estimation of the using time of UFO, and the determination of the adulteration of EVO with UFO. Both the heating time of laboratory prepared UFO and the adulteration of EVO with UFO could be determined by Partial Least Squares Regression (PLSR). To simulate the EVO adulteration with UFO, for each kind of oil, fifty adulterated samples at the adulterant amounts range of 1-50% were prepared. PLSR was then adopted to build the model and both full (leave-one-out) cross-validation and external validation were performed to evaluate the predictive ability.

Lai *et al.* [11] 2016 had proposed a new lip feature representation for lip biometrics which can portray the static and dynamic characteristics of a lip sequence. The new representation catches both the physiological and behavioral parts of the lip and is strong against varieties brought about by various speaker position and posture. In our approach, a lip sequence is initially partitioned into a few subsequences along the fleeting measurement. For every subsequence, sparse coding (SC in short) is received to portray the details of the lip locale and its development in little spatiotemporal cells. At long last, notwithstanding when there is one and only preparing test per speaker, the proposed feature still accomplishes high discriminative power (an exactness of 98.39% and HTER of 2.62%).

Gupta *et al.* [5] 2016 had exhibited that hand dorsal images procured under infrared light are utilized to outline a precise individual authentication framework. Another quality estimation algorithm is proposed to assess the nature of palm dorsal which appoints ease back qualities to the pixels containing hair or skin surface. Matching scores are acquired by matching palm dorsal veins and infrared hand geometry features. These are in the long run combined for authentication. For execution assessment, a database of 1500 hand images gained from 300 unique hands is made. Exploratory results exhibit the predominance of the proposed framework over existing frameworks.

3 Biometrics

Biometric identifiers are the distinctive, measurable characteristics used to label and describe individuals. Biometric identifiers are often categorized as physiological versus behavioral characteristics. Biometrics is used to refer to the field of technology devoted to identification of individuals using biological traits, such as those based on retinal or iris scanning, fingerprints, or face recognition. Biometrics provides a convenient and low-cost additional tier of security. It eliminates problems caused by shared passwords by using physiological attributes. This work Authentication process considers the some natural and soft bio metrics are considered to the online exam process.

4 Proposed Methodology

In the current investigation, an earnest effort is made to design an effective technique for the multimodal biometric recognition employing the soft bio metrics in online exam process. Initially prepare the video frame database for the authentication process. User authentication merely at the very first login session is one among these severe issues, which is normally found in majority of the currently available computer and network systems. This proposed method for continuous user authentication is proposed that continuously collects soft biometric information. In particular, in this method the colors of user's clothing and face as the soft biometric traits are used. This proposed approach having four models such as Initial login authentication, continuous authentication, re login authentication and enrollment template are considered to bio metric authentication process. Proposed block diagram shown in Figure 1. Then the dimension reduction process Independent Component Analysis (ICA) is used and also obtain the matching score value Grey Wolf Optimization (GWO) techniques are used.

This issue is a massively serious security issue, particularly in systems with high security requirement, since an imposter is permitted to access the resources in the system in the period between user log in and user log out. Therefore, this paper introduces a continuous biometric authentication system, wherein, the system is observed incessantly from the time the user logs in.

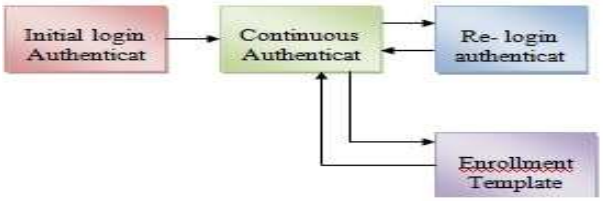

Figure 1: Block Diagram of the proposed method

Figure 2 shows the diagram of template registration and continuous authentication process consider the

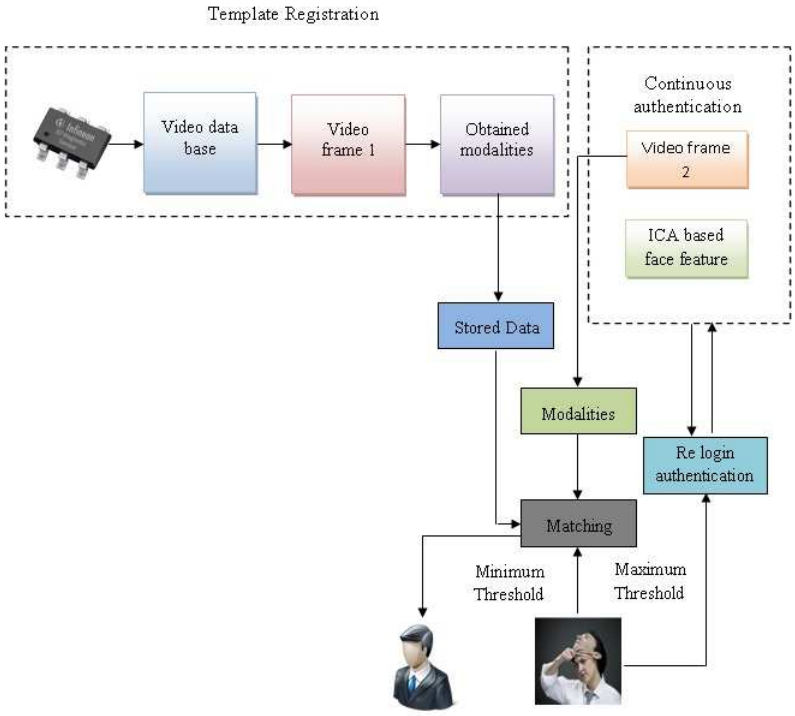

Figure 2: Schematic for continuous authentication system

modalities such as face, ornaments, dress colour, beard, scars and mustache for monitoring the logged in user in a continuous manner. Moreover, the login security of this system is augmented through the union of hard as well as soft biometric traits. Initially template registration process considers the database to obtain the above mention modalities are considered to store the data. Then continuous authentication process chooses the different video frame find the modalities with the dimension reducing ICA technique is used after this process finding maximum score value using optimization technique. This score value based to obtain the fusing score value to identify the user if minimum fusing score value means that is a original user and maximum fusing score means the user is imposter to re login the authentication process.

A genuine user will be the authentication result, if the fused score exceeds the predetermined threshold. Otherwise, the presence of imposter is evident. In the proposed system, a remedy is provided for the situation with imposter.

4.1 Initial Login Authentication

The user employs the conventional authentication system for entering the system. Then, the sensor focuses the user's body for making the registration of the different above mention modalities. During the period of training, the various poses of the user like turn head down, turn head to right, turn head to left, stretching the arms, quitting and leaning back in chair are caught due to the fact that the user may make movements or leave the spot.

Initial Authentication: Biometric face recognition authentication method can be used.

Face Detection: A user is typically looking in the frontal direction during the login session. This is a reasonable assumption because the user typically looks at the monitor at the login time as the user wants to be authenticated.

Body Localization: Location and size of the user's body with respect to his face are estimated.

4.2 Continuous Authentication

Continuous authentication starts after Initial login process. The system continuously authenticates the user by using the "soft face" and "clothing" enrollment templates registered in initial login authentication. The system tracks the face and the body separately based on the histograms registered in initial login process. Hard face recognition is not directly used in continuous authentication but it is stored for use in relogin authentication. In continuous authentication process, the template that is registered in the beginning and the second frame of the video are subjected to the matching process. The matching score value calculated in video frame 1 and video frame 2 in continuous authentication process help of optimization technique, this initial login and continuous authentication with modalities shown in Figure 3.

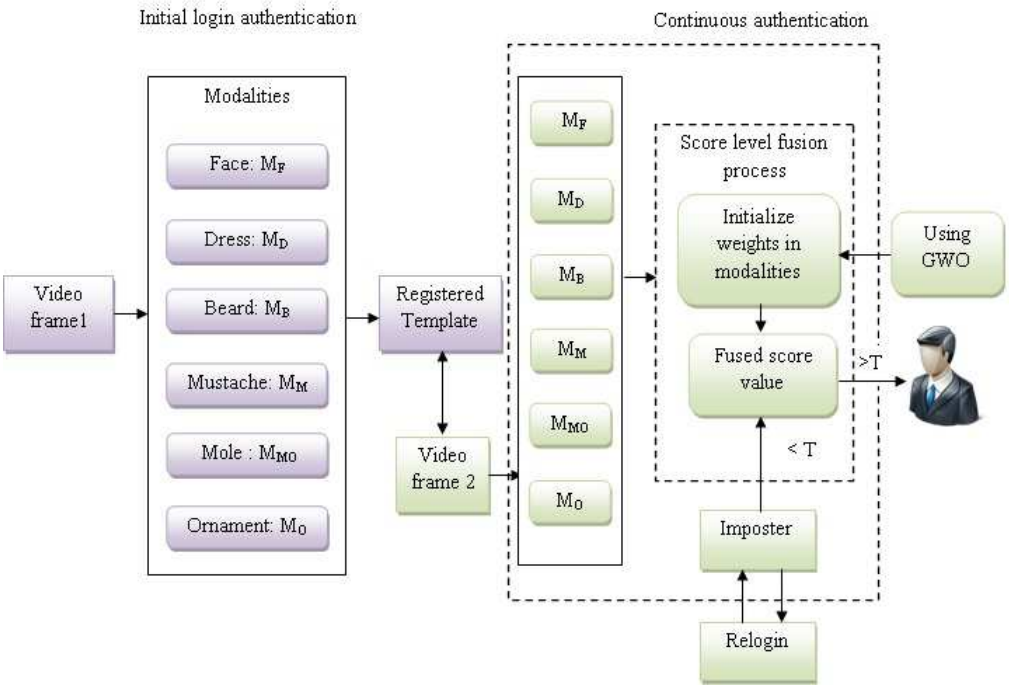

Figure 3: Diagram for Initial and continuous authentication

4.3 Enrollment Template

The system status enters enrollment template whenever the similarity falls below threshold. This step is introduced to reduce the false rejects caused by illumination changes. A pair of images, one just before and one immediately after the time when Similarity $\leq$ threshold is used for image subtraction; the number of pixels that show a large difference in brightness between the two images is counted.

4.4 Score Level Fusion Process

The score level fusion effectively matches scores output of the multiple biometric matchers by integrating them to produce a new match score. The score value calculation different modalities weights are considered to the score value evaluation process. Now the feature weights are optimally chosen with the help of GWO optimization technique the score value is attained. The gradual process of the score level fusion for the authentication process in different modalities considered to find the values. Score value calculation process optimal modalities weights are considered this optimization process is discuss below.

4.4.1 Grey Wolf Optimization Process

The grey wolves adequately frame a Canidae's piece family and are esteemed as the apex predators showing their position at the sustenance's food chain. They routinely show an inclination to make due as a group. The choices made by the alpha are passed on to the group. The Beta speaks to the second rank in the pecking order of the grey wolves. They are, basically, auxiliary wolves which adequately offer some assistance to the alpha in the choice making or comparable group functions. In the GWO technique the hunting (optimization) is guided by the α, β, δ and ω.

Initialization Process: In the district developing procedure, pick the weights of the different modalities such as face, ornaments, dress colour, beard, scars and mustache $W_i = W_1, W_2, ...W_n$ and algorithm parameters, for example, a, A,and C as coefficient vectors.

Fitness Evaluation: In video frame different modalities are considered to the score value calculation process the weights and random values are selected. This GWO Algorithm is being proposed here for accomplishing an enhancement in the performance of the score level fusion. In above equation W_i specifies weight and r_i are random values from [0 to 1].

Based on the fitness separate the solution: Now, we find the fitness separate solution (weight) based on the fitness value.Let the first best fitness solutions be α the second best fitness solutions β and the third best fitness solutions δ.

Update the position: We assume that the alpha (best candidate solution) beta and delta have the improved knowledge about the potential location of the prey in order to mathematically reproduce the hunting behavior of the grey wolves. As a result, we hoard the

first three best solutions attained so far and require the other search agents (including the omegas) to revise their positions according to the position of the best search agent. For revision, the novel solution $W(t+1)$ below mentioned formulas are employed.

$$\begin{aligned} D^{\alpha} &= [C_1.W^{\alpha} - W], \\ D^{\beta} &= [C_1.W^{\beta} - W], \\ D^{\delta} &= [C_1.W^{\delta} - W]. \end{aligned} \quad (1)$$

$$\begin{aligned} W_1 &= W_{\alpha} - A_1(D_{\alpha}, \\ W_2 &= W_{\beta} - A_2(D_{\beta}, \\ W_3 &= W_{\delta} - A_3(D_{\delta}. \end{aligned} \quad (2)$$

Algorithm 1 Pseudo code for GWO

Begin
Initialize the solution
$W_i = W_1, W_2, ...W_n$
Initialize a,A and C
Find the fitness for the initial solution

$$F_i = \sum_{i=1}^{n} W_i.r_i$$

Based on the fitness separate the solution
W_{α} = the best search solution
W_{β} = the second best search solution
W_{δ} = the third best search solution
Update the position of the current search solution
$W(t+1) = \frac{W_1+W_2+W_3}{3}$
Calculate the fitness for new search solution

$$F_i = Max \sum_{i=1}^{n} W_i.s_i$$

Store the best solution so far attained
Iteration=Iteration+1
Stop until optimal solution attained
End

To have hyper-spheres with different random radii the arbitrary parameters A and C help candidate solutions. Investigation and utilization are guaranteed by the adaptive values of A and α. The values of parameters A and α permit the GWO to smoothly transition them among the investigation and the utilization. With decreasing A, half of the iterations are dedicated to the investigation ($|A| < 1$) and the other half are devoted to the utilization. Encircling the behavior, the subsequent equations are employed in order to mathematically model.

$$D = |CW_{p(t)} - W(t)| \quad (3)$$

For find the coefficient vectors use Equation (3):

$$A = 2ar_1 = a, \qquad C = 2r_2 \quad (4)$$

Where t indicates the current iteration, A and C are coefficient vectors, $W_p$is the position vector of the prey T and indicates the position vector of a grey wolf. The components of α are linearly decreased from 2 to 0 over the course of iterations and r_1, r_2 are random vectors in [0, 1]. The GWO has only two main parameters to be adjusted (α and C). However, we have kept the GWO algorithm as simple as possible with the fewest operators to be adjusted. The maximum score value obtained in the process will be continued.

4.4.2 Matching Process

Matching is conducted with a weight preset as W. An optimized weighting strategy was used in an earlier phase for yielding a fused score of all the features. The fundamental structure of the matching process, which works in accordance to the preset threshold, is shown as follows:

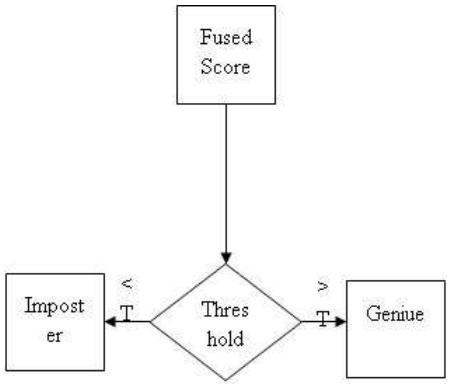

Figure 4: Matching process

The below Figure 4 states that a comparison is made between the fused modality score and the preset threshold. If the result of comparison is in such a way that the fused score exceeds the threshold level, the user is deemed as genuine. Else if the threshold is smaller than the fused score, the user is proved to be an imposter or a fake one. If a fake user is identified, our proposed methodology allows another process, known as Re-login authentication, to be carried out.

4.5 Independent Component Analysis (ICA)

ICA is a data analysis tool derived from the "source separation" signal processing techniques. The aim of source separation is to recover original signals S_i , from known observations O_j , where each observation is an (unknown) mixture of the original signals. If unsuccessful authentication occurs in any place of the authentication process, Re-login authentication is immediately conducted as the subsequent step in the proposed scheme. It is expected that, ICA source vectors being independent (instead of PCA eigenvectors being uncorrelated only), they will be closer to natural features of images, and thus more able to represent differences between faces.

4.5.1 Use of ICA in face feature authentication

ICA is an unsupervised technique which separates the independent sources from a mixture. The general model of ICA is

$$O = BS. \tag{5}$$

Where B represents unknown mixing matrix, S represents unknown source signal and O represents observed mixtures. In this case, it is assumed that the source signals are statistically independent and non-Gaussian and observed mixtures is the only information to have. In ideal condition, mixing matrix B can be inversed. If the estimation of separation matrix is accurate, then a good approximation of source signal will be obtained.

$$I = WO = WBS \quad and \quad W = B^{-1}. \tag{6}$$

Where I represents unknown mixing matrix.

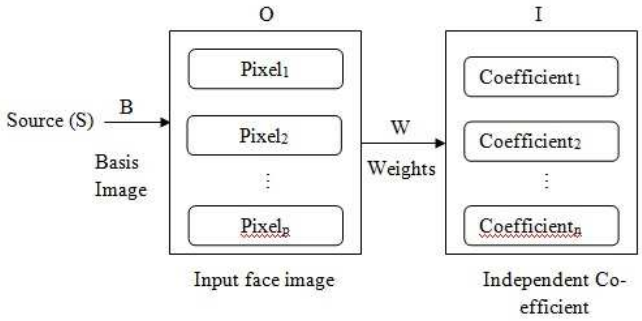

Figure 5: Image synthesis model

In Figure 5 shows the pixels are treated as variables and images are observations. This results the column of $B = W^{-1}$ as a set of basis image. Column of I contains a set of independent coefficient of basis images in A for reconstructing image in O. Therefore, I is a factorial code representation. In order to assess the sensitivity of ICA in terms of the dimension of the compressed and whitened space where it is implemented, we carried out a comparative assessment for different dimension whitened subspace.

This recognition phase computes the weights W_k for both the training as well as the test frame. The computation of the difference in weights allows finding the Euclidean distance. To achieve recognition, a threshold has to be predetermined. The expressions in the images would be identical, if the threshold and the Euclidean distance have the same value. The weight W_k is computed in accordance to the following equation.

$$W_k = I_k(B_i - \phi_k) \tag{7}$$

Where

$$I_k = \sum_{k=1}^{n} E_k.\phi_i$$

Further, I_k denotes the Eigen faces,E_k points to the Eigen vectors and ϕ_k represents the mean adjusted value.

Re-login step will be performed at the condition, when the authentication ends up in failure in the proposed continuous biometric system.

4.6 Relogin Authentication

In this process the system is locked and it tries to detect the user and re authenticates him automatically. If the system detects a user and re authenticates the user as genuine, the status moves to continuous authentication process. Here, the user is authenticated using both soft (colour histograms) and hard biometrics (face). The similarity score is used for relogin authentication. There will be a small discontinuity in the values of soft biometrics when the unauthenticated person tries to replace the student. When there is a discontinuity in the similarity scores based on the soft biometric, the system enters relogin authentication mode. In the relogin authentication mode, the user must provide valid soft and hard biometrics.

5 Result and Discussion

This section discusses about the results of the proposed method biometric authentication using ICA with GWO technique and has scrutinized their appearance in the working platform of MATLAB 2014 with the system configurations as i5 processor with 4GB RAM. This model consider the hard biometric is face and soft biometrics Ornaments, beard, mustache, dress color and mole different performance evaluation parameters are obtained.

5.1 Database Description

This work generates the synthetic database to the continuous authentication scheme. Each user was asked to perform the following set of actions while seated in front of the webcam. We collected videos of 20 subjects using the system shown in Figure 8. Every one user was asked to carry out the subsequent set of action while seated in front of the webcam. A few examples are illustrated in Figure 6.

5.2 Performance Evaluation Metrics

The effectiveness of proposed technique is analyzed by invoking some performance measures such as False Rejection Ratio (FRR), False Accept Ratio (FAR), Sensitivity, specificity and accuracy. The performance measures are explained below:

False Rejection Ratio: The system identifies imperfectly that a user is not in the camera's field of view although the user is yet in front of the camera. False discards lower the usability of the system.

$$FRR = \frac{Genuine\ \ scores\ \ falling\ \ below\ \ Threshold}{All\ \ Genuine\ \ Score} \tag{8}$$

Figure 6: Sample Video frames for authentication process

difference between the time of enrolment and the time of identification is mitigated. The suggested re-login authentication method is assessed by means of video clips where an authorized user logs in, the user leaves the work environment (without logging out) and next, another user (an impostor) emerges in the field of view of the webcam. Figure 6 shows the different position identification of the data.

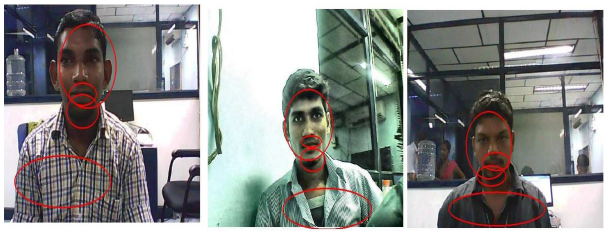

Figure 7: Face, clothing and mustache sample

False Accept Ratio: The system incorrectly identifies an imposter as the legitimate user. False admits lower the security of the system.

$$FRR = \frac{Imposter\ \ scores\ \ exceeding\ \ Threshold}{All\ \ Imposter\ \ Score} \tag{9}$$

Sensitivity: Sensitivity is a measure which determines the probability of the results that are true positive as 'that person has the authenticated person.

$$Sensitivity = \frac{NTP}{NTP + NFN} \tag{10}$$

where NTP denotes number of true positives; NFN denotes number of false negatives.

Specificity: Specificity is a measure which determines the probability of the results that are true negative as 'that person does not have the authenticated person.

$$Specificity = \frac{NTN}{NTN + NFN} \tag{11}$$

where NTN denotes number of true negatives; NFN denotes number of false negatives.

Accuracy: Accuracy is a measure which determines the probabilities that how may results are accurately authenticated.

$$Accuracy = \frac{(TP + TN)}{TP + TN + FP + FN}. \tag{12}$$

5.3 Registration of Biometric Data

The system registers face biometric data. This work using ICA with GWO optimization approach based face recognition. Because the system registers face biometric data every time a user logs in, the problem of the illumination

Face recognition is executed at regular intervals (e.g., once every 10 seconds). If it succeeds, Tlast, which represents the last time the face recognition was successful, is updated. Face recognition is used only for assisting the identification using colour histograms because the system cannot obtain the face information during different cases. The system enters the initial login authentication mode. On the other hand, if the user is absent for only a short time, it is more likely that he will be accepted given valid soft and hard biometric traits.

Figures 8 and 9 shows that the Graphical User Interface (GUI) for new user authentication and original; user authentication process. Initially load the video then obtain the score value and compare the matching score in original and new user authenticated the person, if the score value based obtain similarity in original user and imposter.

Figure 10 shows that the FRR and FAR in GWO and GSO techniques, the FRR rate is maximum value compared FAR. Performances of face and soft biometrics are evaluated using False Acceptance Rate FAR and the False Rejection Rate FRR. Test was conducted using different number of training files. FAR is the percentage of illegal users that are accepted as genuine. FRR is the percentage of legal user rejected as imposter. From the result, FAR and FRR is high for small number of trained samples. The proposed technique FRR is 0.82 its maximum value compared to GSO similar difference in FAR in authentication process.

5.4 Comparative Analysis in Performance Parameters

Here a comparison of authentication process in existing approach PCA with GSO and propose technique ICA with GWO techniques is compared. The parameters such as accuracy, sensitivity, specificity, FRR and FAR are compared.

Figure 8: New user login authentication

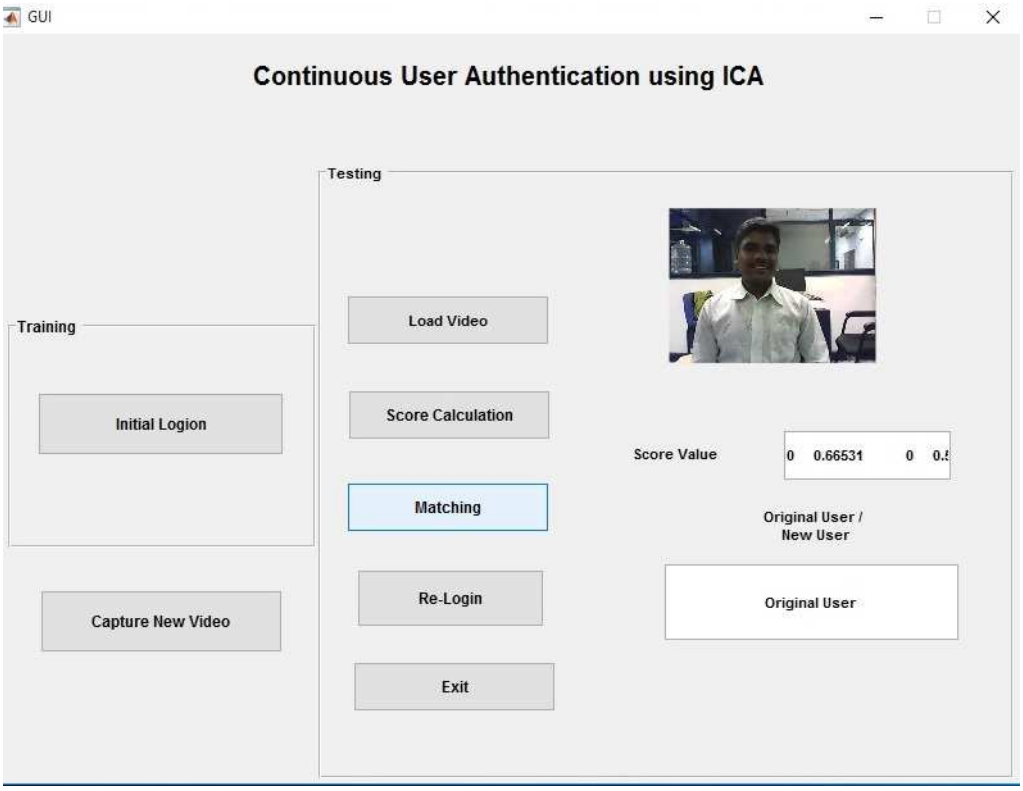

Figure 9: Original user login authentication

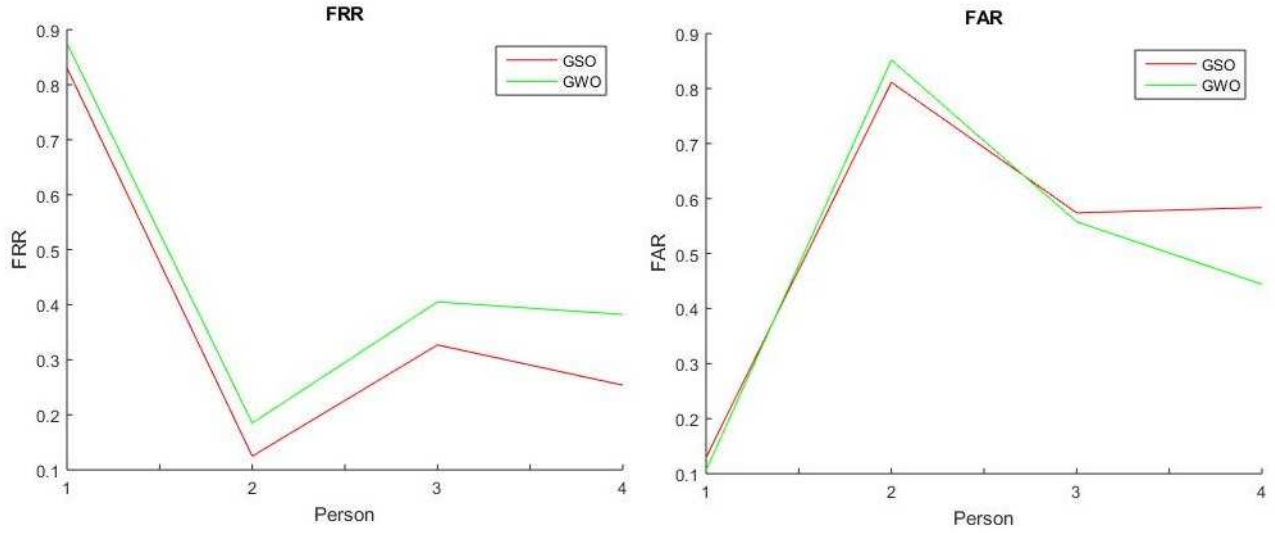

Figure 10: Comparison graph for FRR and FAR

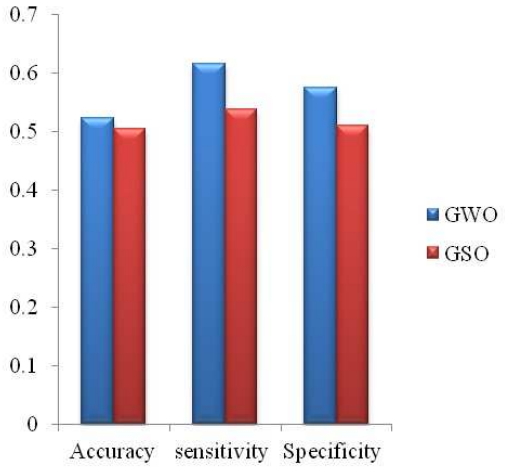

Figure 11: Comparative analysis graph

Figure 11 shows that the accuracy, sensitivity and specificity comparison GWO and GSO technique, here different four persons are considered to evaluate this performance. The maximum performance attained in ICA with GWO techniques, maximum accuracy is 0.5246 it's compared to GSO the difference is 0.256%. Likewise, the accuracy of the ICA with GWO being 6.05% , it is seen reduced by 0.16% and 1.15% respectively in the case of GSO. As a whole, the proposed method shows a significant hallmark of 0.75% when compared with the other methods in terms of the parameters specified in the bar graph. Similarly other parameters are smaller difference in authentication process.

6 Conclusion

This framework registers a new enrolment template every time the user logs in, which enables the system to effectively use soft biometric traits for continuous authentication; the proposed system uses face color information as well as clothing color (soft biometric) to continuously authenticate the user. This authentication process ICA with GWO produced the maximum accuracy value. The main purpose of this paper is to present a new e-learning model used for identification, authentication and tracking the student. The system is robust with respect to user's posture in front of the workstation. Experimental results demonstrate that the system is able to successfully authenticate the user continuously with high tolerance to the user's posture. In our ongoing work, we are considering introducing additional soft biometric traits. By applying these methods we enhance Continuous Authentication system and try to obtain better result other than state-of-art method.

References

[1] T. Ahmad, J. Hu, and S. Wang, "Pair-polar coordinate-based cancelable fingerprint templates", *Journal of Pattern Recognition*, vol. 44, pp. 2555-2564, 2011.

[2] M. Choras, and R. Kozi, "Contactless palmprint and knuckle biometrics for mobile devices", *Journal of Theoretical Advances*, vol. 7, no. 1, pp. 73-85, 2012.

[3] G. Gao, J. Yang, J. Qian, L. Zhang, "Integration of multiple orientation and texture information for finger-knuckle-print verification", *Journal of Neurocomputing*, vol. 135, pp. 180-191, 2014.

[4] P. Gupta, and P. Gupta, "An efficient slap fingerprint segmentation and hand classification algo-

rithm", *Journal of Neurocomputing*, vol. 8, pp. 1-14, 2014.

[5] P. Gupta, S. Srivastava and P. Gupta, "An accurate infrared hand geometry and vein pattern based authentication system", *Journal of Knowledge-Based Systems*, vol. 103, no. 9, pp. 143-155, 2016.

[6] L. Han, Q. Xie, and W. Liu, "An improved biometric based authentication scheme with user anonymity using elliptic curve cryptosystem," *International Journal of Network Security*, vol. 19, no. 3, pp. 469-478, 2017.

[7] C. Hegde, Phanindra, D. Shenoy, and Patnaik, "Human authentication using finger knuckle print", *Journal of Biometric Authentication*, vol. 3, pp. 1-8, 2011.

[8] O. Kaiwartya, M. Prasad, S. Prakash, *et al.*, "An investigation on biometric internet security," *International Journal of Network Security*, vol. 19, no. 2, pp. 167-176, 2017.

[9] Kekre, and Bharadi, "Finger-knuckle-print region of interest segmentation using gradient field orientation and coherence", *Journal of Emerging Trends in Engineering and Technology*, vol. 978, pp. 130-133, 2010.

[10] A. Kumar and D. Zhang, "Improving biometric authentication performance from the user quality", *Journal of Instrumentation and Measurement*, vol. 59, no. 3, pp. 730-735, 2010.

[11] J. Y. Lai, S. L. Wanga, A. W. C. Liew and X. J. Shi, "Visual speaker identification and authentication by joint spatiotemporal sparse coding and hierarchical pooling", *Journal of Information Sciences*, vol. 373, no. 8, pp. 219-232, 2016.

[12] C. T. Li, M. S. Hwang, "An online biometrics-based secret sharing scheme for multiparty cryptosystem using smart cards," *International Journal of Innovative Computing, Information and Control*, vol. 6, no. 5, pp. 2181-2188, May 2010.

[13] C. T. Li, M. S. Hwang, "An efficient biometrics-based remote user authentication scheme using smart cards", *Journal of Network and Computer Applications*, vol. 33, no. 1, pp. 1-5, 2010.

[14] C. H. Ling, C. C. Lee, C. C. Yang, and M. S. Hwang, "A secure and efficient one-time password authentication scheme for WSN", *International Journal of Network Security*, vol. 19, no. 2, pp. 177-181, 2017.

[15] A. Meraoumia, S. Chitroub, and A. Bouridane, "Palmprint and finger-knuckle-print for efficient person recognition based on Log-Gabor filter response", *Journal of Analog Integration Circuit Signal Processing*, vol. 69, no. 1, pp. 17-27, 2011.

[16] A. Prakash, "A biometric approach for continuous user authentication by fusing hard and soft traits", *International Journal of Network Security*, Vol.16, No.1, pp.65-70, Jan. 2014.

[17] A. Prakash, R. Dhanalakshmi, "Stride towards proposing multi-modal biometric authentication for online exam", *International Journal of Network Security*, vol. 18, no. 4, pp. 678-687, July 2016.

[18] Shankar, Sahoo, and Niranjan, "Using the digital signature of a fingerprint byan elliptic curve cryptosystem for enhanced authentication", *Journal of Information Security*, vol. 21, pp. 243-255, 2012.

[19] J. Tan, R. Li, Z. T. Jiang, *et al.*, "Synchronous front-face fluorescence spectroscopy for authentication of the adulteration of edible vegetable oil with refined used frying oil", *Journal of Food Chemistry*, vol. 217, no. 7, pp. 1-7, 2016.

[20] M. Tarek, O. Ouda, T. Hamza, "Pre-image resistant cancelable biometrics scheme using bidirectional memory model," *International Journal of Network Security*, vol. 19, no. 4, pp. 498-506, 2017.

[21] L. Zhang, L. Zhang, D. Zhang, and H. Zhu, "Online finger-knuckle-print verification for personal authentication", *Journal of Pattern Recognition*, vol. 43, no. 7, pp. 2560-2571, 2010.

Biography

Prakash Annamalai is working as Assistant Professor at Jerusalem College of Engineering, Chennai. He has received B.E and M.E degree in Computer Science and Engineering. He is currently pursuing Ph.D at Hindustan Institute of Technology and Science. His areas of research interests include Network Security and Image Processing.

Krishnaveni Raju is currently working as Professor at Hindustan Institute of Technology and Science, India. She is a PhD holder from Anna University. She completed her B.E degree from Bharadhiyar University and M.E degree from Madras University. She has published around 30 research papers in International Journals and International Conferences including Springer, IFIP, JCTN and many other referred journals. Her areas of research interests include Network Security, Biometrics Security and Web Security.

Dhanalakshmi Ranganayakulu a Ph.D holder from College of Engg., Guindy Anna University Chennai for the research activities in Information Security and Networking. She holds a B.E in Computer Science from Bharathidasan University and M.Tech in Advanced Computing from SASTRA University. She has vital research experience serving as a research Associate in the NTRO Sponsored Project Collaborated directed basic research on Smart and Secure Environment at Anna University under the consortium of IIT Madras. To her credit, she has nearly 25 research papers in International Conferences and International Journals including Elsevier, Springer, IFIP and IGI Global. Her fields of interest include Information Security, Data Mining, Knowledge and Semantic Networks, Intelligent Networks and Mobile Computing.

Feasibility of Eliminating IDPS Devices from a Web Server Farm

Sujatha Sivabalan, P. J. Radcliffe
(Corresponding author: Sujatha Sivabalan)

The Department of Electrical and Computer Engineering& RMIT University
Swanston Street, Melbourne, Australia
(Email: s3365148@student.rmit.edu.au)
(Received June 28, 2017; revised Aug. 27 & Sept. 25, 2017)

Abstract

Current web security systems need Intrusion Detection and Prevention Systems (IDPS), web proxies and firewalls to protect the websites from malicious network traffic. All these functions come at a cost for a web farm and add to power costs. Our previous work has concluded that the web server detection of application layer DDoS attacks is far more power efficient than an equivalent IDPS. This paper shows that all remaining IDPS functionality can be split between the firewall and the web server allowing the removal of the traditional IDPS and so substantially reducing the CPU load and total electrical power bill of a web farm.

Keywords: Application Layer (App-layer); Intrusion Detection and Prevention System (IDPS); Web Server

1 Introduction

The Web plays a vital role in modern life but web servers are under constant attack. Common attacks include buffer overflow based attacks, cross-site scripting, code injection, brute force attacks, and DoS attacks [11, 13]. A web security system is usually built around web proxies, firewalls and IDPS. The IDPS play a significant role in blocking attacks but they can be overwhelmed by high traffic levels and consume electrical power. In a modest system such as a host-based web server, the network traffic is received, analysed, and transmitted by these security devices before getting to the web server. Consider the above scenario; the packet needs to traverse the full TCP/IP stack three times before the web server. Consider a web server farm in an enterprise that operates many web servers with a large number of security devices. These security devices consume significant power due to this repetitive packet.

A possible solution is to remove any devices if the functionality can build into other existing devices. Former research work [14] analysed the power consumption of traditional IDPS as a separate device in a host based web server and compared this to a Two Dimensional Web page Daemon (TDWD) which implemented IDPS functionality within the web server in a host based system. The experimental results show that the traditional IDPS consumed significantly more power than the TDWD. Based on this novel foundation, this paper examines the possibility of detecting all types of attacks by distributing the IDPS functionality between a web server and other devices thus eliminating the IDPS box or IDPS software such as Bro or Snort. If this proves feasible, it will reduce equipment costs and the electricity usage in a web server farm.

This paper is organized as follows: Section 2 describes the different types of attacks handled by an IDPS and explores attack detection strategies for web servers and firewalls. Section 3 explains a novel security system that eliminates the IDPS. Section 4 considers practical implementation. Section 5 suggests some exciting future work based on the new architecture. Finally Section 6 provides a conclusion.

2 Literature Search

This first part of this section will examine IDPS functionality and discuss the types of attacks such a system can detect and block. The second part will then show that many authors have suggested how individual IDPS functions can be implemented either within a firewall or web server.

2.1 IDPS Functionality

An IDPS aims to detect and alert the system when suspicious traffic occurs and blocks the offending traffic [18]. IDPS detection methodologies on major attacks are discussed below.

Phishing [12] is a common problem in an email, where the embedded hyperlink in an apparently legitimate email

redirects to the fake website which aims to steal user secrets. Common aims of this attack includes financial gain, identity hiding mainly in the purchase of goods and for fame and notoriety. Phishing attack is possible through the websites blogs and on commercial websites. IDPS apply signature-based detection for phishing attacks. The author Khonji *et al.* [8] describes the Collaborative Intrusion Detection System (CIDS) where many number of IDS share phishing related data and each IDS will maintain a list of infected IP addresses, pattern matches to mitigate phishing attacks.

Several major websites including eBay, Google, and McAfee have been the targets of cross-site scripting [4], SQL injection exploits, or content based sniffing [6, 17]. An attacker injects malicious script in the web application thereby causing unintended script execution by the victims browsers. Once this attack is successful, the attacker can perform exploits such as account hijacking, cookie poisoning, DoS and web content manipulation. IDPS detection on these attacks are based on the signature rules such as pattern matching, whitelist techniques which compare inputs with the known good inputs, and model-based approaches to analyse the user behaviour [18].

Brute force attacks are an illegal attempt to websites by repeatedly trying username and password. Major victims are email and banking user. IDPS detect these types of attacks by pattern matching [18]. According to a review by Hydara *et al.* [6] this functionality can be provided by a web server.

Cookie poisoning is a fraudulent act on cookie data after accessing a website. This is a common attack on web applications, for example in an online shopping. An attacker can poison the cookie by neglecting the shipping fee or postal price using tools as Paros proxy [9] that results in financial loss to the owners. IDPS detect these types of attacks by state transition analysis or by model-based approaches.

Network layer DoS (Net-layer DoS) attacks include SYN attack, ICMP attack, and UDP attack [5]. These attacks aim to flood the server and make it unavailable to the legitimate requests. IDPS will detect and block flood attacks by using state transition analysis.

Well known commercial IDPS products such as those from Cisco IDS, Snort or Bro can detect these types of attacks with the help of attack signature matching in central databases. IDPS plays a major role in detecting and blocking attacks. Most inline IDPS sensors offer firewall capabilities to mitigate the suspicious network activity.

2.2 Detection Techniques by Web Server And Firewall

This subsection examines how web servers and firewalls [2] can take over responsibilities of an IDPS. There is a good body of research showing that web servers provide effective detection against individual forms of application layer (app-layer) attacks. For example WebIDS [7] from IBM Tivoli Risk Manager analyses the Web servers access log files to detect Web server attacks. Apache server modules such as mod_security, mod _evasive, and mod_rewrite [10] can be configured to defend against applayer attacks. Mod_security is as web application firewall designed for blocking applayer attacks. Mod_evasive is an Apache module used to detect DDoS attacks, the detection is based on number of single page access per unit time. Anton *et al.* [1] deployed a web server to detect cookie poisoning and SQL injection and stated that server side detection is more powerful for cross-site scripting than a firewall. Their approach based on checking the payload content such as response headers, < Meta tag > and the number of bytes. Web servers can filter for phishing attacks [8] with the help of blacklisting and whitelisting IP.

Web servers are unsuitable for Net-layer DoS attacks such as SYN, ICMP and UDP flood attacks. Haining *et al.* [5] showed that an advanced firewall is capable of resisting these types of flooding attacks. Gallagher [3] stated that instead of using an IDPS, firewalls could be configured to block Net-layer attacks and also Domain Name Service (DNS) and Network Time Protocol (NTP) reflection attacks.

Commercial web application scanners include AppScan, WebInspect, Hailstorm, Acunetix WVS. Open source web application scanners, include Paros and Pantera. These scanners examine the log files from the web server to detect problems. This is essentially an IDS function but not a real time thus making them less useful for directly blocking unwanted traffic.

The literature has outlined the functionality of IDPS and has shown that all individual app-layer IDPS functions can be moved into the web server, and that individual net-layer IDPS functions can be implemented in a firewall. There is no overarching reasoning as to why the IDPS is still required. There is no commentary on the possibility or advantages of completely eliminating an IDPS.

3 Novel Architecture

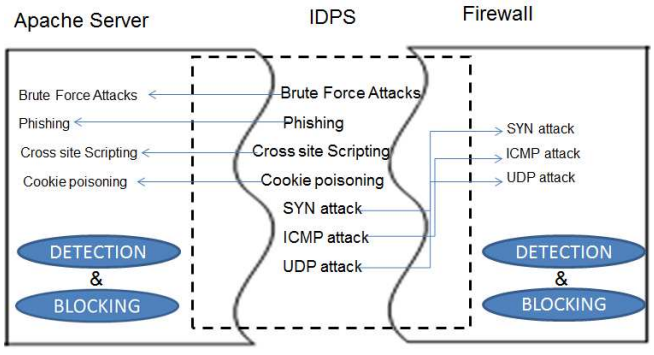

Figure 1: Attacks handled by web server and firewall

The literature search showed that many authors have implemented individual IDPS functions on the firewall or

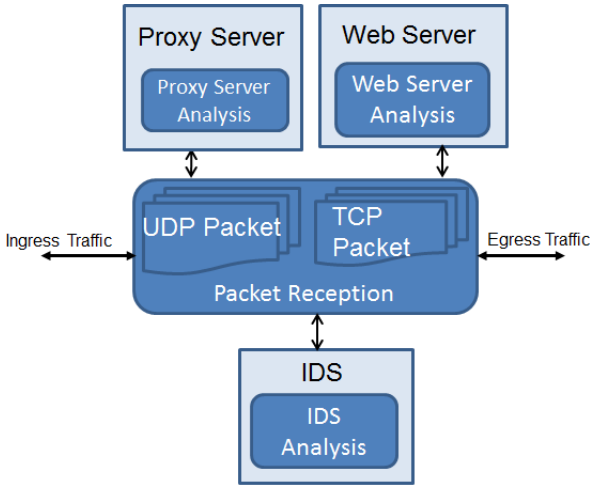

Figure 2: Novel architecture

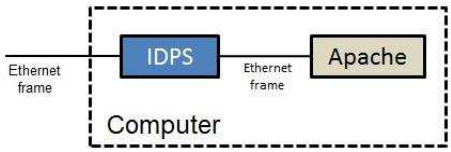

Figure 3: Traditional network design

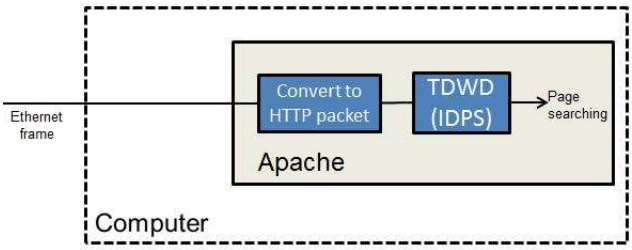

Figure 4: Novel design

web server. This section will show that the total IDPS function can successfully split between a web server and a firewall thus making it possible to eliminate the IDPS. Furthermore, it will show that this results in real power and CPU savings even in a cloud environment.

3.1 Eliminating IDPS

Figure 1 illustrates the attacks handled by IDPS can be splitted between a web server and a firewall. The Apache web server has powerful modules such as mod_status, mod_rewrite, mod_evasive, mod_clamav, and mod_proxy that can be customised using HTTP variables for effective detection of attacks.

The detection techniques used by web server modules are listed in Table 1. The Apache web server is capable of detecting all types of attacks except network layer attacks, which the firewall can handle. The first column lists the common types of attacks. The second column describes the user activities on each attack. The IDPS detection and blocking methods on each one is summarized in the third column.

The last column explores the web server techniques in handling those attacks. All attacks can be handled by the web server or firewall thus it is possible to eliminate the IDPS. In situations requiring very high security the redundancy of having an IDPS may be considered worthwhile but there are no attacks which and IDPS can handled that cannot be handled by the web server and firewall together.

3.2 Power Saving Opportunity

Figure 1 showed that much of the IDPS functionality can absorbed into the web server and so it is feasible that the security modules can work at the HTTP, TCP, or UDP level. This proposed a novel architecture is shown in Figure 2.

The system receives packets to the TCP, UDP, or HTTP level and these formed packets are shared between the applications thus the packets need not to travel up and down TCP/IP stack, from Ethernet frame to TCP/UDP and from TCP/UDP back down to Ethernet frame. Avoiding multiple packet reception and transmission on each devices results in far less CPU computing time and the use of far less electrical power. The next section examines the size of this electrical power saving.

4 Practical Implementation

The basic architecture of a traditional IDPS and web server is shown in Figure 3 where the ingress and egress traffic for each network device is at the level of the Ethernet frame. The IDPS used in our work was Snort, and another called Bro. Section 3.2 showed that this was computational inefficient and proposed a generic solution in Figure 2. Figure 4 shows how this was implement as a way to eliminate Bro or Snort when working with Apache. This novel design has been implemented in our previous work [14, 15, 16] and further extended in this paper. Our TDWD is built into the Apache web server and uses a two-dimensional linklist (client IP and time) with time based garbage collection. The following rules have been implemented in TDWD:

- DDoS attack: excessive similar page accesses per second, accessing pages at random, and repeatedly accessing the same page. Blocking rules can be evaluated by examining the linklist of accesses for each user IP.
- Brute Force attack: repeatedly accessing the administration page.
- Blocking: Blacklisting based on user IP.

Table 1: Attack methods handled by web server

Attacks	User Behavior	IDPS	Web server detection
Brute force attacks	Many failed login for a single account from same IP address or different IP addresses	• IDS signatures • Pattern matching • Block blacklist IP's • Account lockout • CAPTCHA	• Blacklist IP address using "HTTP USERAGENT" • Redirect CAPTCHA webpage
Phishing	Illegal URL and URI and Forged mail headers.	• Block blacklist IP's • Block illegal request • IDS signatures	• Blacklist IP address using "HTTP USERAGENT" • Block through Request using "THE_REQUEST" or "REQUEST_URI" • Block through the referrer using "HTTP_REFERER" • Get and Post scanning for blockable text eg bad web site.
Cross site scripting	Malicious script execution in the HTTP request.	Anomaly based detection	• Blacklist IP address using "HTTP USERAGENT" • Block through Request using "THE_REQUEST" or "REQUEST_URI" • Block through the referrer using "HTTP_REFERER" • Block through Query String. • Scan post and get for any script, reject if any found.
Cookie poisoning	Change in content of Cookie data.	• Block blacklist IP's • Block illegal request • IDS signatures	• Blacklist IP address using "HTTP USERAGENT" • Block through Request using "THE_REQUEST" or "REQUEST_URI" • Block through the referrer using "HTTP_REFERER" • Block through "HTTP_Cookie"

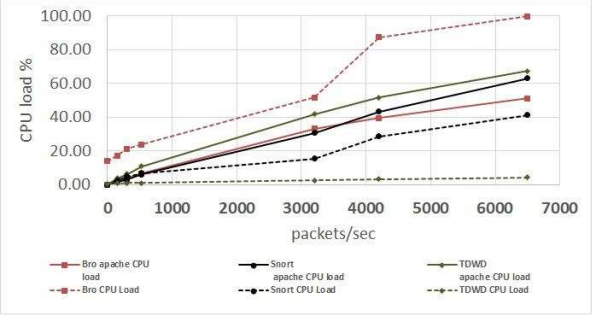

Figure 5: Bro, Snort and TDWD a comparison

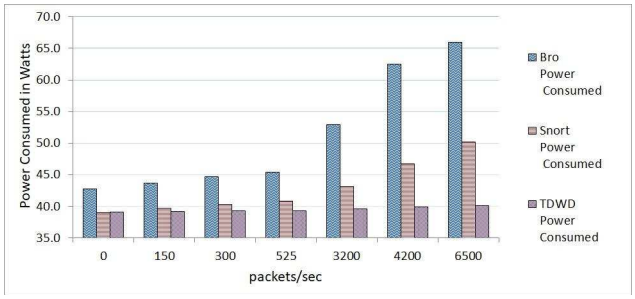

Figure 6: Power consumption of Bro, Snort and TDWD

In order to analyse the power savings of the TDWD approach three experimental setups were devised; the Bro IDPS running with Apache, the Snort IDPS running with Apache, and TDWD which integrates into Apache. In order to achieve a realistic comparison each IDPS devices was run by itself on the same machine with Apache active and servicing packets. The CPU usage is shown in Figure 5 where the dotted lines are the CPU load of the individual IDPS devices, and the solid lines show any variation in the Apache CPU load given the IDPS being used. Our previous research [16] shows that the CPU usage is

linearly related to the power consumption. Using a watt meter we successfully calibrated our computer against this model which enables the translation of the CPU utilization in Figure 5 to real watts in Figure 6. This shows that at the highest traffic load TDWD consumes 40 watts, Snort 50 watts, and Bro 65 watts.

Additionally the IDPS programs were configured to receive the network traffic but do no processing. The power consumption was proportional to the packets per second. When the IDPS analysis was added, the power consumption increased very marginally. The receive process, which is dominated by Ethernet frame to TCP/UDP translation, was responsible for the majority of the power consumption in the IDPS. To confirm this, the same type of packet analysis was programmed into TDWD, which runs inside Apache and inspects the already assembled HTTP packets. This added functionality marginally increased the power consumption of Apache but this was under 10% of the power used by Snort or Bro to achieve the same functionality.

The experimental work has clearly shown that IDPS functionality can be moved into a web server and that further more there is a significant saving in electrical power. This work points to a general methodology whereby network devices share data at the highest possible level (HTTP, TCP, or UDP) and do not waste CPU time and power in unnecessary conversions between Ethernet frames and higher levels. This approach could be useful in a wide variety of network designs.

5 Future Work

The TDWD is novel and powerful tool as the link list of user page requests with time stamping allows complex DDoS detection rules to be implemented as well as more basic rules such as blacklisting. There is considerable scope to develop rules well beyond the simple rules we have implemented. There is also scope for machine learning to analyse the link list to detect anomalies caused by attacks. The TDWD structure is suitable for other research projects, as it will work in a real webserver or in a cloud network. The focus of this work has been the reduction in electrical power usage by allowing networking devices to share data at the highest possible level, HTTP, TCP, or UDP. By reducing CPU utilization not only is power reduced, the speed of network operations should also be significantly improved. The ability to speed up network devices in this way and the actual time savings achievable are an interesting research area. Not only can networks be made greener, they can be made faster.

6 Conclusion

This paper has examined the literature and has concluded that an IDPS device can have all its functionality moved into the firewall or web server thus the IDPS may be removed. An IDPS may be kept for reasons of security redundancy but it is not required as a function. The IDPS functions that are moved into the web server can be implemented at the UDP, TCP, or HTTP level thus eliminating the repeated conversions between Ethernet frame and TCP/IP. This results in a useful saving of CPU capacity and electrical power. In proving that there was a saving of electrical power, a Two Dimensional Web page Daemon (TDWD) was developed to hold user access requests so they could be analysed for attacks, particularly DDoS attacks. This structure has proved to be very useful and may be the basis for developing novel intelligent attack detection and blocking algorithms.

References

[1] A. Barua, H. Shahriar, and M. Zulkernine, "Server Side Detection of Content Sniffing Attacks," in *IEEE 22nd International Symposium on in Software Reliability Engineering (ISSRE'11)*, pp. 20-29, 2011.

[2] A. Blyth, "An architecture for an XML enabled firewall," *International Journal of Network Security*, vol. 8, pp. 31-36, 2009.

[3] S. Gallagher, *Biggest DDoS Ever Aimed at Cloudflare Content Delivery Network*, May 24, 2016. (`http://arstechnica.com/security/2014/02/ biggest-ddos-ever-aimed-at-cloudflares-content-delivery-network/`)

[4] S. Goswami, N. Hoque, D. K. Bhattacharyya, and J. Kalita, "An unsupervised method for detection of XSS attack," *International Journal of Network Security*, vol. 19, pp. 761-775, 2017.

[5] W. Haining, Z. Danlu, and K. G. Shin, "Detecting SYN flooding attacks," in *IEEE Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM'02)*, pp. 1530-1539, 2002.

[6] I. Hydara, A. B. M. Sultan, H. Zulzalil, and N. Admodisastro, "Current state of research on cross-site scripting (XSS): A systematic literature review," *Information and Software Technology*, vol. 58, pp. 170-186, 2015.

[7] IBM, *Tivoli Risk Manager Analyses*, Jan. 8, 2018. (`https://publib.boulder.ibm.com/tividd/td/TRM/GC32\-1323\-00/en\_US/HTML/ad`)

[8] M. Khonji, Y. Iraqi, and A. Jones, "Phishing detection: A literature survey," *Communications Surveys & Tutorials*, vol. 15, pp. 2091-2121, 2013.

[9] J. Long, A. W. Bayles, J. C. Foster, C. Hurley, M. Petruzzi, N. Rathaus, et al., *Chapter 4 - Web Server & amp: Web Application Testing*, Penetration Tester's Open Source Toolkit, ed Burlington: Syngress, pp. 189-276, 2005.

[10] I. Muscat, *How To Mitigate Slow HTTP DoS Attacks in Apache HTTP Server*, May 5, 2015. (`https://www.acunetix.com/blog/articles/slow-http-dos-attacks-mitigate-apache-http-server/`)

[11] A. A. Orunsolu, A. S. Sodiya, A. T. Akinwale, B. I. Olajuwon, "An anti-phishing kit scheme for secure web transactions," *International Journal of Electronics and Information Engineering*, vol. 6, no. 2, pp. 72–86, 2017.

[12] A. San Martino and X. Perramon, "Phishing secrets: History, effects, countermeasures," *International Journal of Network Security*, vol. 11, pp. 163-171, 2010.

[13] J. J. Sheu, "Distinguishing medical web pages from pornographic ones: An efficient pornography websites filtering method," *International Journal of Network Security*, vol. 19, no. 5, pp. 839-850, 2017.

[14] S. Sivabalan and P. J. Radcliffe, "Real time calibration of DDoS blocking rules for web servers," *Computer Communication & Collaboration*, vol. 4, pp. 42-50, 2016.

[15] S. Sivabalan and P. J. Radcliffe, "A novel framework to detect and block DDoS attack at the application layer," in *IEEE TENCON Spring Conference*, pp. 578-582, 2013.

[16] S. Sivabalan and P. J. Radcliffe, "Power Efficient Secure Web Servers," *International Journal of Network Security*, vol. 20, no. 2, pp. 304-312, 2018.

[17] M. Stampar, "Inferential SQL injection attacks," *International Journal of Network Security*, vol. 18, pp. 316-325, 2016.

[18] T. Verwoerd and R. Hunt, "Intrusion detection techniques and approaches," *Computer Communications*, vol. 25, pp. 1356-1365, 2002.

Biography

Sujatha Sivabalan received a B.Eng from Bharathidasan University(India)in 2004 and M.Eng from Anna University (India) in 2007.Currently doing Ph.D. in School of Electrical and Computer Engineering, RMIT University, Australia. Her research interest includes network security, DDoS attack detection and blocking.

P. J. Radcliffe received a B.Eng from Melbourne University (Australia) in 1978 and worked for Ericsson Australia R&D for 7 years followed by consulting to various companies. He joined Royal Melbourne Institute of Technology (RMIT) and was awarded and M.Eng. in 1993 and a PhD in 2007. Main research interests include network protocols, Linux, and embedded systems. He received a national teaching award in 2011 and in 2012 received the RMIT early career researcher award.

A Novel Physical Channel Characteristics-based Channel Hopping Scheme for Jamming-resistant in Wireless Communication

Qiuhua Wang*[1], Hongxia Zhang[2], Qiuyun Lyu[1], Xiaojun Wang[1], And Jianrong Bao[3]
(Corresponding author: Qiuhua Wang)

School of Cyberspace, Hangzhou Dianzi University[1]
School of Communication Engineering, Hangzhou Dianzi University[2]
School of Information Engineering, Hangzhou Dianzi University[3]
No. 1158, 2nd Street, Jianggan District, Hangzhou 310018, P.R.China
(Email: wangqiuhua@hdu.edu.cn)
(Received Nov. 30, 2016; revised and accepted Mar. 11, 2017)

Abstract

Jamming is an effective denial-of-service (DoS) attack in wireless networks due to the open nature of radio propagation. In Jamming attack, the attacker purposely emits radio signals to corrupt the ongoing communication between the legitimate transmitter and receiver. Channel hopping is a feasible link-layer method for preventing jamming attack in wireless communications. In this paper, we propose a novel channel hopping scheme for jamming-resistant in wireless communication. In our proposed scheme, we explore the reciprocity, randomness and spatial uncorrelation of the wireless fading channel to generate random channel hopping sequences. We evaluate our channel hopping scheme through real-world experiments on 802.11a 5 GHz band. Experiment results show that our scheme is efficient and secure, and achieves higher channel agreement ratio with almost equally channel distribution.

Keywords: Channel Hopping; Channel Reciprocity; Jamming Attack; Quantization

1 Introduction

As wireless networks become increasingly popular, the security and reliability issues attract more and more attentions. Due to the broadcast and open nature of radio propagation, wireless networks are not only vulnerable to traditional attacks such as eavesdropping but also to jamming attacks [24]. Jamming is a very effective denial-of-service (DoS) attack, in which the attacker purposely emits radio signals to corrupt the ongoing communication between a pair of legitimate users [15, 28]. Jamming resistance is crucial for secure and reliable wireless communication.

The dominantly used approach to cope with jamming attacks is to employ physical layer techniques such as Direct Sequence Spread Spectrum (DSSS) [1] and Frequency Hopping Spread Spectrum (FHSS) [19]. These techniques use identical spreading codes or frequency hopping sequences known to both the sender and the receiver but unknown to jammers to achieve anti-jamming capability. To do this, both the sender and the receiver need to share secret keys (such as spreading codes in DSSS or frequency hopping sequences in FHSS) beforehand and keep them secret [3]. However, those spread spectrum techniques employ sophisticated physical-layer, which require more advanced and expensive transceivers and cannot be employed in most commodity wireless networks. Moreover, although the Frequency Hopping was available in the original 802.11 standard, it was not incorporated into the subsequent, more popular 802.11a, b and g protocols [16].

An alternative easy-performed method for anti-jamming is channel hopping (also known as channel surfing), in which legitimate transceivers quickly switch their communication channels to avoid jamming from attackers [8, 20, 23, 25].

The idea of channel hopping is motivated by frequency hopping. Channel hopping is similar to frequency hopping in that both of them change frequency during the communication. However, the difference between them is that, FHSS, unlike channel hopping, requires specialized antennas for transmitting and receiving signals. Channel hopping is a link-layer technology, it is much more feasible and easily used than FHSS and can be applied to the existing wireless devices without frequency hopping features [20]. Since there are multiple channels are available for next hopping, the key concern for channel hopping is to achieve the same channel selection between legitimate users.

Similar to FHSS, channel hopping also relies on a secret key shared by sender and receiver to control the channel selection. This secret key enables the communication parties to switch channels such that their transmission becomes unpredictable for a third party, thus reducing the probability of jamming. Without such a shared secret, it is impossible to establish effective anti-jamming communication between sender and receiver.

Until now, the requirement of shared keys has been fulfilled by out-of-band key pre-distribution on the devices. However, this approach suffers from scalability constraints in environments where a large number of users potentially take part in a pairwise communication, and may not even be feasible in highly dynamic network environments such as mobile ad hoc networks where two arbitrary parties usually do not have pre-shared secrets and they have to talk to each other beforehand to decide the channel switching sequence. Moreover, an attacker can compromise the pre-shared key and then jams the network. When those happen, the communication parties will have to agree on a new secret key in an ad-hoc manner using the wireless channel.

All these observations lead to the following challenge: How can two users that do not pre-share any secret key achieve the channel hopping agreement securely over a wireless channel in the presence of a jammer? Only when two legitimate transceivers select the same channel at each time slot can they successfully communicate.

Some research works resort to Diffie-Hellman key agreement algorithm or public key encryption (e.g RSA) to establish a shared secret key over an insecure channel. However, unfortunately, all these schemes share some common limitations: (1) They have to split each DH/RSA message into multiple packets at the sender and reassemble them into meaningful DH/RSA messages at the receiver due to the constraint of wireless network packet size. This takes a long time (and sometimes is impossible) for these schemes to finish a DH/RSA key establishment in presence of jammers [9]. (2) Such methods consume significant amount of computing resources and power which might not be available in certain scenarios (e.g., wireless sensor networks). (3) More importantly, since they are based on the hardness of a mathematical problem, they are only computational secure.

Recently, exploiting wireless channel characteristics (e.g reciprocity, randomness and spatial uncorrelation) to generate a shared secret key between two legitimate users has become a promising technique for its high reliability, easy implementation, and low energy consumption [26]. It provides an excellent approach to the problem of key-establishment and can even achieve information theoretical secrecy [14]. However, almost all of these schemes need extra information reconciliation to correct the quantized bit errors between two parties [21]. If we try to adopt these approaches for channel hopping purpose directly, the extra communication overheads are considerable because information reconciliation needs many rounds information exchange and should be performed each time the channel switches. As a result, it will take a long time (and sometimes it may be impossible) for these schemes to finish information reconciliation in the presence of jammers.

In this paper, we propose a novel channel hopping scheme based on wireless channel characteristics which is effective and energy-efficient. Our method makes use of the inherent reciprocity, randomness and spatial uncorrelation of wireless fading channel. In typical wireless network environments, the wireless channel between two users, Alice and Bob, is reciprocal and varies randomly over time and space. Alice and Bob can measure some wireless channel characteristics (such as received signal strength indicator (RSSI) [4, 6, 10, 11, 14, 17], amplitude [13, 22] and phase [5, 12]. The channel reciprocity theory demonstrates that bidirectional wireless channel characteristics should be identical between two transceivers within the channel coherence time. We can use these measurements as shared random secrets to achieve the channel selection agreement.

In our approach, the information reconciliation procedure is eliminated, which greatly reduces the communication overheads and time cost. Therefore, our method is more energy efficient. Our approach only needs one-time extra information exchange which happens in the quantization phase.

Furthermore, due to the spatial uncorrelation of wireless channel, as long as the jamming attacker, Eve, is more than a half-wave-length away from Alice and Bob, the channel measurements she obtains will be independent to that between the legitimate ones. This means that the attacker can obtain no information about the channel characteristics between legitimate communicators because she experiences independent fading [2] and thus cannot measure the same channel characteristics as Alice and Bob [14]. To this extent, our channel hopping method provides a strong security.

We also conduct real-testbed experiments to evaluate our approach. The results show that with least information exchange, we can achieve a channel agreement ratio higher than 95 % and even 100%.

The rest of this paper is organized as follows. Section 2 introduces the network and adversary model used in our proposed scheme. Section 3 provides the detailed description of our proposed channel hopping scheme. Section 4 presents the experiment results and performance analysis. Finally, we conclude the paper in Section 5.

2 Network and Attack Model

We now outline the basic wireless network and jamming attack model that we use throughout this paper.

2.1 Network Model

Here we consider an ubiqitous Alice-Bob-Eve wireless communication scenario in Figure 1, in which Alice, Bob

and Eve are geographically located at different positions. The legitimate users, Alice and Bob, want to transmit messages via wireless channel. An jamming-attacker, Eve, tries to jam the communication between Alice and Bob by sending random packets (or noise). Both Alice and Bob have multiple transducers that allow them to work on multiple channels. In our setting, Alice and Bob each sends data and ACK packets through the wireless channel from which they respectively measure the channel characteristic and construct the channel measurements, denoted by h_{ab} and h_{ba}. Due to the channel reciprocity, we have $h_{ab} \approx h_{ba}$ when they are conducted during the channel coherence time. Eve can estimate her channel to Alice or Bob, however, if Eve is more than λ / 2 (λ is the wavelength) away from Alice and Bob, she will experience independent channel variations, hence, her observations h_{ae} and h_{be} are sufficiently uncorrelated with h_{ab} and h_{ba} due to the spatial variations, e.g., $h_{ae} \neq h_{ab}$ and $h_{be} \neq h_{ba}$ [2].

If Alice and Bob communicate on a fixed channel, Eve, could identify the channel used for communication and then start to jam it indefinitely. Clearly, if the legitimates wish to continue communication, they must hop to a new channel. Let l represent the number of channels that can be utilized between legitimate users. For instance, l =12 in 802.11a when only non-overlapping channels are used for communication. The legitimate users may change channels over time. When such a channel change occurs, we say that the communication hops between channels.

2.2 Attack Model

Since we focus on message transmissions in the presence of a jammer, we only consider jamming attacks in this paper. Similar to the assumption in traditional channel hopping schemes [16, 20, 25], instead of considering a powerful attacker, we assume that Eve uses the same or similar hardware as legitimate users in terms of capability, energy capacity, and complexity, and the power-limited attacker can jam only one channel at a given time. One reason is that, to remain inconspicuous, the jammer would need to jam with conventional (802.11) hardware such as a single laptop with one or two wireless interfaces. In many cases, the attacker launches a jamming through a compromised node. Another reason for such assumption is that if a jammer is a high-power, broadband capable device, it can be easily detected by defenders since they violate the normal communication rules. Eve has the knowledge of the set of channels used by the sender and the receiver, and she chooses her jamming strategy depending upon the information that she obtains about the system.

3 Our Proposed Channel Hopping Scheme

Our method relies on the reciprocity of channel to achieve channel agreement, and spatial uncorrelation to prevent eavesdropping. Let $X^A = (x_1^A, x_2^A, ..., x_n^A)$ be the channel measurements recorded by Alice at time $t_1, t_2, ..., t_n$ respectively, and $X^B = (x_1^B, x_2^B, ..., x_n^B)$ be Bob′s channel measurements at time $t_{1'}, t_{2'}, ..., t_{n'}$, where $t_1 < t_{1'} < t_2 < t_{2'} < ..... < t_n < t_{n'}$, and x_i^u is the channel measurement value at time t_i or $t_{i'}$, $u = \{A, B\}$, $u = A$ represents Alice and $u = B$ represents Bob. According to the reciprocity of wireless channel, $x_i^A \approx x_i^B (1 \leq i \leq n)$, if they are obtained within the channel coherence time t_τ , i.e, $t_{i'} - t_i \ll t_\tau$. Since the channel variation is mainly caused by channel fading, it is random and unpredictable. Moreover, based on the location decorrelation property of the wireless channel, the attacker Eve cannot observe the same channel variations as the Alice-Bob channel if she is located several wavelengths away.

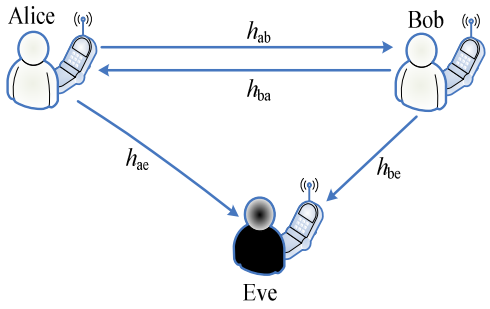

Figure 1: Wireless communication scenario.

3.1 Channel Hopping Protocol

Suppose that in the initial stage, Alice and Bob have obtained n channel measurements, $X^u = (x_1^u, x_2^u, ..., x_n^u), u = \{A, B\}$ prior to Eve′s arrival. Both Alice and Bob can use l channels, and their clocks are synchronized. Our protocol is described as follows.

1) Alice and Bob first quantize their channel measurements into binary bit sequences of length m respectively by performing channel quantization algorithm. Then they randomly permute the bit positions in their quantized sequences and obtain the random sequence $Q^u = (q_1^u, q_2^u, ..., q_m^u) \in \{0, 1\}^m$.

2) Alice and Bob divide their random sequence Q^u into blocks of length l denoted as

$$\begin{aligned} B_0^u &= (q_0^u, q_1^u, \cdots, q_{l-1}^u) \\ B_1^u &= (q_l^u, q_{l+1}^u, \cdots, q_{2l-1}^u) \\ \vdots &= \vdots \\ B_{\lfloor m/l \rfloor - 1}^u &= (q_{(\lfloor m/l \rfloor - 1)l}^u, q_{(\lfloor m/l \rfloor - 1)l+1}^u, \cdots, q_{\lfloor m/l \rfloor l-1}^u). \end{aligned}$$

Then Alice and Bob compute a random channel selection sequence $CS^u = (CS_1^u, CS_2^u, ..., CS_{\lfloor m/l \rfloor - 1}^u)$, respectively, where $CS_i^u = \{E_{B_i^u}(i) \bmod l\}$, i is the packet sequence number, $B_i^u = (q_{il}^u, q_{il+1}^u, ..., q_{(i+1)l-1}^u), (0 \leq i \leq \lfloor m/l \rfloor - 1)$, and $E_K(\cdot)$ is an encryption function.

3) Both Alice and Bob change their own channels according to the random channel selection sequence CS^u (which is unknown to anyone but the two parties involved).

4) Alice and Bob exchange their packet pair (DATA-ACK) with the sequence number i using channel CS_i^u. Once the communication begins, the channel characteristics are recorded and used to calculate the next round channel choice.

5) If Alice and Bob cannot achieve agreement on one channel (i.e. CS_i^A is not equal to CS_i^B), both of them go to the next channel represented by CS_{i+1}^u.

6) Once the current round channel selection sequence is used up, Alice and Bob jump to Step 1) to generate a new channel selection sequence based on the channel measurements obtained during the current round, and continue their next round communication according to the new computed channel selection sequence.

3.2 Quantization Algorithm

It is obvious that quantization is a crucial step in our proposed channel hopping scheme, and the choosing of the quantization algorithm has a great influence on the performance of our proposed protocol.

In the quantization stage, both Alice and Bob quantize their channel measurements into binary bit sequence based on particular thresholds. There are many proposals of channel quantization. The paper [27] summarizes some existing quantization methods and evaluates their performance. The difference in these quantization methods mainly results from their different choices of thresholds and the different number of thresholds they use. These quantization methods could generally be classified into two categories: Single-bit approaches and Multi-bit approaches [11]. Single-bit approaches quantize each channel measurement into at most one bit, while Multi-bit approaches quantize each channel measurement into multiple secret bits, k-bit ($k > 1$), but at a cost of higher bit error rate.

In tradition key generation based on the channel-characteristic, to achieve an identical shared key between two legitimate users, an information reconciliation protocol should be used to reconcile the bit errors. However, during information reconciliation phase, Bob and Alice must exchange reconciliation information on public channel several times, which is time and energy consumption. Even worse, information reconciliation leaks some information about the secret key which can be used by the attacker to guess portions of the extracted key, hence, a privacy amplification protocol will be further applied to solve this issue, which consumes more time and communication overheads.

Moreover, with the quantization bit error rate increasing, the subsequence information reconciliation will become more and more difficulty and the whole process of key generation may even be failure. That is because when the quantization bit error rate increases, the information reconciliation protocol has to be performed much more rounds to eliminate all errors, which will reveal more bits to the attacker. For example, when the bit error rate is 0.08 after the quantization phase, the Winnow information reconcile protocol should be performed 5 rounds to eliminate all errors with 57.13% information leaked. While when the bit error rate is 0.25, the Winnow information reconcile protocol should be performed 11 rounds with 96.77% information leaked [21]. So, the quantization approaches that exhibit high bit error rate are not useful in establishing a secret key.

In this paper, since we focus on designing a fast and energy efficient channel hopping protocol, we try to achieve channel agreement with high probability without using information reconcile and privacy amplification. This requires that the bit error rate of the quantization algorithm should be as low as possible. Our experimental result in Section 4.1 shows that with l=12, to achieve a channel agreement ratio higher than 95%, the bit error rate of the output of quantization should be lower than 0.4%. However according to our experiment results in Figure 2, none of multi-bit quantization approaches can achieve such low bit error rate. Therefore, in this paper, we choose to use two-threshold single-bit approach since it has lower bit error rate compared with Multi-bit approaches and other single-bit approaches [11].

The quantization algorithm used in this paper is described as follows.

1) Both Alice and Bob divide the channel measurement sequence into blocks of length j which is a configurable parameter.

2) For each block, they calculate two thresholds: the upper threshold q_+^u and the lower threshold q_-^u independently such that

$$q_+^u = \mu^u + \alpha \times \sigma^u. \quad (1)$$
$$q_-^u = \mu^u - \alpha \times \sigma^u. \quad (2)$$

where μ^u and σ^u represent the mean and the standard deviation of the measurement sequences in the ith block, and $0 < \alpha < 1$ is a parameter which can be tuned through experiments.

3) Each measurement value $x_i^u (1 \le i \le n)$ is mapped to a binary bit via a quantizer $Q^u(\cdot)$ as shown in Figure 3, such that measurements below q_-^u are encoded as bit 0; measurements above q_+^u are encoded as bit 1, while measurements within the interval $[q_-^u, q_+^u]$ are discarded.

$$Q^u(x_i^u) = \begin{cases} 1, \ if \ x_i^u > q_+^u \\ 0, \ if \ x_i^u < q_-^u \\ e, \ otherwise \end{cases} \quad (3)$$

where e is an undefined state. The superscript u stands for user and may refer to either Alice, in which

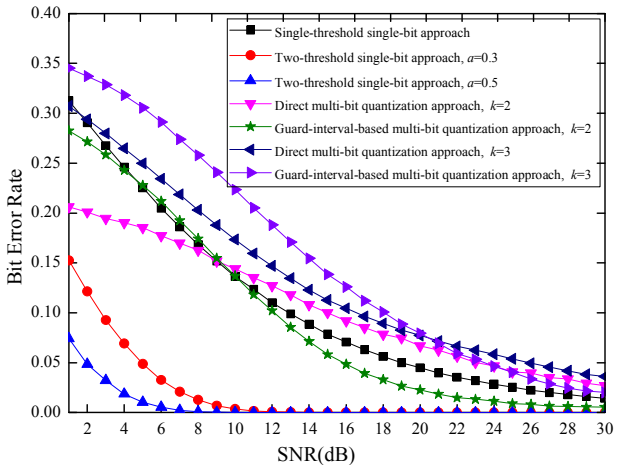

Figure 2: Bit error rate of different quantization approaches.

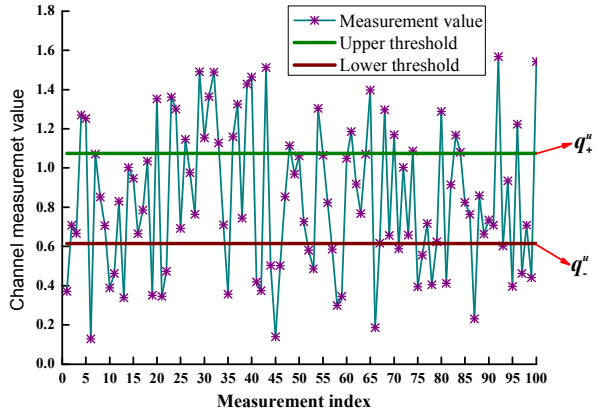

Figure 3: Bit error rate of different quantization approaches.

case the quantizer function is $Q^A(\cdot)$, or to Bob, for which the quantizer is $Q^B(\cdot)$.

4) Alice and Bob maintain a list of indexes of discarded values and exchange it with each other so that they only keep the ones that they both decide not to drop.

4 Experiment Results and Performance Analysis

In this section, we first describe our experiment settings, and then present the results and the performance of our scheme.

In our experiment, we make use of the most popular channel characteristic parameter, RSSI, as the indicator of the channel because its reading is readily available in the existing wireless infrastructures. Most of the current of-the-shelf wireless cards, without any modification, can measure it on a per packet/frame basis. RSSI can be read during the preamble stage of receiving an 802.11 frame. The variation over time of the RSS caused by motion and multipath fading, can be quantized and used for generating channel agreement sequences.

It should be noted that our approach is also applicable to any other parameters of channel characteristic, such as amplitude or phase, etc.

4.1 Experiment Setup

We conducted our experiments on three laptops (acting as Alice, Bob and Eve) equipped with in-built Intel PRO/Wireless 3945ABG network cards in a real indoor environment. Alice is configured as an access point (AP mode) and remains stationary, while Bob acts as a client (Station mode) and moves randomly at a speed of about 1m/s . Eve is configured to monitor mode and sits next to Alice, only about 30 centimeters away. Alice records the RSSI values of data packets from Bob, and Bob records the RSSI values of the corresponding MAC layer ACK packets from Alice. These data and ACK packets are all for data communication between Alice and Bob.

We perform experiments on 802.11a 5 GHz band because it has more non-overlapping channels (12 non-overlapping channels in the 802.11a) than 802.11b/g. We note that this experiment setting favors the jammer, as she only needs to scan 12 channels in order to detect the used channel between legitimate users. The jammer is guaranteed a direct hit once he locates the channel with traffic. In our experiments, the residence time that the legitimate communication stays fixed in a particular channel is 100 ms.

4.2 Channel Hopping Approach

As introduced in Section 2.1, after the legitimate users communicate on a single channel for a short period or once a communication channel is jammed, they must jump to another new channel to continue the communication. According to the jamming attack style, there are two most efficient hopping approach for jamming: the reactive hopping and the proactive hopping.

- Reactive hopping approach [24, 25]: the legitimate users hop to a new channel only after they have detected the presence of a jammer on the current channel they are using.

- Proactive hopping approach [7]: the legitimate users hop channels for every t seconds without attempting to detect the presence or the absence of the jammer on the current channel and hopping channel.

The advantage of the reactive hopping is the least hop number per unit time, while the proactive approach switches hop more often than is necessary. However, the advantage of the proactive hopping is that since channel hopping takes place for every t seconds it is difficult

for the jammer to identify the current channel. It is robust when appropriate t value is chosen [7]. Moreover, in the proactive hopping, the legitimate users dont need to detect the presence of a jammer. In fact, obtaining an accurate estimate of a channel's status in a short period of time is not easy. For example, the DOMINO system [18] as reported requires several seconds to make an accurate determination of a greedy station. In our experiment, we make use of the proactive hopping approach. Alice and Bob switch their channel every 100 ms.

It should be noted that when a pulsing, fast-switching attacker appears, channel hopping can work in conjunction with other approaches such as the packet fragmentation and the redundant encoding to defend against this type of jamming [23].

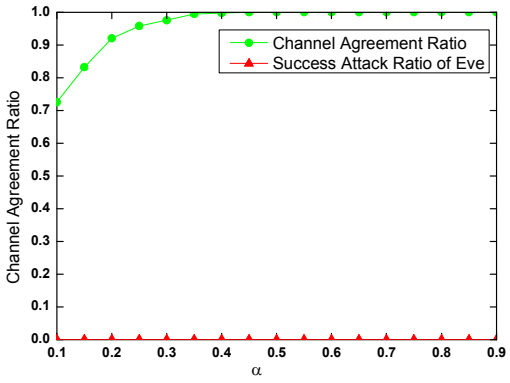

Figure 4: Channel agreement ratio and Eve's successful attack ratio as a function of quantization parameter α

4.3 Channel Hopping Agreement Ratio and Distribution of Channel Selection

We simulate the performance of our proposed channel hopping scheme with different quantization parameter α. The experimental results are shown in Figure 4. It is clear that larger value of α leads to a higher probability of channel agreement rate. When $\alpha = 0.3$, our channel hopping method can achieve a channel agreement ratio higher than 97%. When α =0.45, the channel agreement ratio of our method can even achieve 100%.

We also evaluate the channel selection distribution of our proposed scheme. The experiment results are draw in Figure 5. From Figure 5, we can see that the hopping probability on each channel is almost equally distributed.

Additionally, our approach eliminates the information reconciliation cost and only needs one time extra information exchange in the quantization phase. Therefore, our proposed channel hopping scheme is an effective one and is more energy efficient.

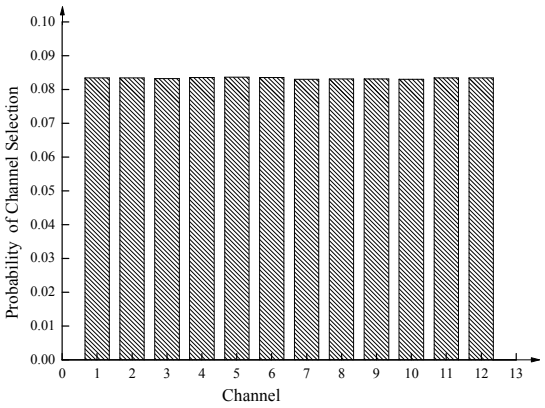

Figure 5: Distribution of channel selection

4.4 Security

As for the security, on the one hand, Eve′s channel measurements do not provide her any useful information about the measurement sequences X^A and X^B due to the spatial uncorrelation. On the other hand, the transmission of position indexes over the public channel in the quantization phase does not reveal any information about the quantized bits to Eve either. This is because they contain position indexes only, whereas the generated quantized bits depend upon the values of the channel measurement at those indexes. Further, this guarantees that Eve cannot use his observations to infer the values of the channel measurement of Alice or Bob at those indexes

Eve can perform two kinds of Jamming attacks: predictive jamming and reactive jamming [7, 3]. If Eve chooses to use the predictive jamming, she has two possible strategies. The first is random Jamming. Eve sends jamming signal on a random channel. In this case, the probability of Eve selecting the correct channel at random is $1/l$. In our experiments, $l = 12$, and Eve′s random attack successful probability is about $1/12 = 8.33\%$ (Of course, the more channels legitimate users have, the lower probability Eve can achieve). By itself, this gives a relatively high chance of success to the attacker. However, successful jamming one message gives no advantage for the next, since legitimate users will be on a different, unknown channel. If Eve is more intelligent, she may use the same method as Alice and Bob to select channel. Eve can also collect the RSSI sequences by eavesdropping the communication between Alice and Bob, and then computes the channel hop sequences following the same steps as our method. We tested the probability that Eve successfully computes the channel that Alice and Bob will hop to. The experimental results are also shown in Figure 4. From the results, Eves successful attack probability is almost 0, which is even much lower than the random attack. This demonstrates that our proposed scheme is also resistant to more sophisticated jammers. Our experiment results also verify that the channel fading is a random shared secret between Alice and Bob, and Eve almost obtains no useful information by eavesdropping.

On the other hand, if Eve chooses to use the reactive jamming attack, she needs to scan all channels first and

then launches attacks on the channel that legitimate users are using. The time to scan each channel and the radio start-up cost in a new channel of 802.11 devices is typically tens of milliseconds [24]. In our experiments, since Alice and Bob switch their channel every 100 ms, it is impossible for Eve to complete scanning before Alice and Bob switch their channels.

Hence, our proposed scheme is secure and causes no loss of secrecy.

5 Conclusions

In this paper, we focus on how to protect legitimate transmission from jamming attack by having the legitimate users hop among channels to hide the transmission from the jammer. We propose a novel channel hopping scheme for jamming-resistance based on wireless channel characteristics. We evaluate the proposed scheme through real-world experiments in terms of the channel agreement rate, distribution of channel selection and security. The experiment results show that our proposed scheme can work reliably with high efficiency and that it achieves higher channel agreement rate with almost equally channel distribution. Moreover, our scheme is light-weight and easy-implementable on the current wireless devices. In the future, we will focus on achieving channel hopping agreement without extra information exchange.

Acknowledgments

This work was done when Qiuhua Wang visited the Department of Electrical & Computer Engineering, Syracuse University. The authors would like to thank Dr. Yingbin Liang for her constructive comments and helpful suggestions. This work was partially supported by National Natural Science Foundation of China (No.61401128, No.61471152), Zhejiang Province Natural Science Foundation (No. LQ14F020010), Project of Zhejiang Provincial Key Enterprises Institute Construction and Project of Zhejiang Provincial Smart City regional synergy innovation center and the China Scholarship Council (CSC).

References

[1] A. Alagil, M. Alotaibi, and Y. Liu, "Randomized positioning dsss for anti-jamming wireless communications," in *International Conference on Computing, Networking and Communications (ICNC'16)*, pp. 1–6, 2016.

[2] G. D. Durgin, *Space-time Wireless Channels*, Prentice Hall PTR, New Jersey, USA, 2003.

[3] K. Grover, A. Lim, and Q. Yang, "Jamming and anti-jamming techniques in wireless networks: A survey," *International Journal of Ad-Hoc and Ubiquitous Computing*, vol. 17, no. 4, pp. 197–215, 2014.

[4] R. Guillaume, F. Winzer, and A. Czylwik, "Bringing phy-based key generation into the field: An evaluation for practical scenarios," in *The 82nd IEEE Vehicular Technology Conference (VTC Fall'15)*, pp. 1–5, 2015.

[5] J. Huang and T. Jiang, "Dynamic secret key generation exploiting ultra-wideband wireless channel characteristics," in *IEEE Wireless Communications and Networking Conference (WCNC'15)*, pp. 1701–1706, 2015.

[6] S. Jana, S. N. Premnath, M. Clark, S. K. Kasera, N. Patwari, and S. V. Krishnamurthy, "On the effectiveness of secret key extraction from wireless signal strength in real environments," in *International Conference on Mobile Computing and Networking (MOBICOM'09)*, pp. 321–332, 2009.

[7] K. Karunambiga, A. C. Sumathi, and M. Sundarambal, "Channel selection strategy for jamming-resistant reactive frequency hopping in cognitive wifi network," in *International Conference on Soft-Computing and Network Security (ICSNS'15)*, pp. 1–4, 2015.

[8] S. M. Khattab, D. Mosse, and R. G. Melhem, "Modeling of the channel-hopping anti-jamming defense in multi-radio wireless networks," in *International Conference on Mobile and Ubiquitous Systems: Computing, Networking, and Services (MOBIQUITOUS'08)*, pp. 1–5, 2008.

[9] A. Liu, P. Ning, H. Dai, and Y. Liu, "Usd-fh: Jamming-resistant wireless communication using frequency hopping with uncoordinated seed disclosure," in *IEEE International Conference on Mobile Adhoc and Sensor Systems (IEEE MASS'10)*, pp. 41–51, 2010.

[10] H. Liu, J. Yang, Y. Wang, and Y. Chen, "Collaborative secret key extraction leveraging received signal strength in mobile wireless networks," in *The 31st Annual IEEE International Conference on Computer Communications (IEEE INFOCOM'12)*, pp. 927–935, 2012.

[11] Y. Luo, L. Pu, Z. Peng, and Z. Shi, "Rss-based secret key generation in underwater acoustic networks: Advantages, challenges, and performance improvements," *IEEE Communications Magazine*, vol. 54, no. 2, pp. 32–38, 2016.

[12] M. G. Madiseh, S. He, M. L. McGuire, S. W. Neville, and X. Dong, "Verification of secret key generation from uwb channel observations," in *IEEE International Conference on Communications (ICC'09)*, pp. 593–597, 2009.

[13] S. Mathur, Miller R, A. Varshavsky, W. Trappe, and N. Mandayam, "Proximate: Proximity-based secure pairing using ambient wireless signals," in *The 9th ACM International Conference on Mobile Systems, Applications, and Services (MobiSys'11)*, pp. 211–224, 2011.

[14] S. Mathur, W. Trappe, and N. Mandayam, "Radio-telepathy: extracting a secret key from an unauthen-

ticated wireless channel," in *International Conference on Mobile Computing and Networking (MOBICOM'08)*, pp. 128–139, 2008.

[15] B. Mihajlov and M. Bogdanoski, "Analysis of the wsn mac protocols under jamming dos attack," *International Journal of Network Security*, vol. 16, no. 4, pp. 304–312, 2014.

[16] V. Navda, A. Bohra, S. Ganguly, and D. Rubenstein, "Using channel hopping to increase 802.11 resilience to jamming attacks," in *IEEE International Conference on Computer Communications (INFOCOM'07)*, pp. 2526–2530, 2007.

[17] S. N. Premnath, S. Jana, J. Croft, and P. L. Gowda, "Secret key extraction from wireless signal strength in real environments," *IEEE Transactions on Mobile Computing*, vol. 12, no. 5, pp. 917–930, 2013.

[18] M. Raya, J. Hubaux, and I. Aad, "Domino: A system to detect greedy behavior in IEEE 802.11 hot spots," in *ACM International Conference on Mobile Systems, Applications, and Services (MobiSys'04)*, pp. 84–97, 2004.

[19] Y. R. Tsai and J. F. Chang, "Using frequency hopping spread spectrum technique to combat multipath interference in a multiaccessing environment," *IEEE Transactions on Vehicular Technology*, vol. 43, no. 2, pp. 211–222, 1994.

[20] S. Vadlamani, B. Eksioglu, H. Medal, and A. Nandi, "Jamming attacks on wireless networks: A taxonomic survey," *International Journal of Production Economics*, vol. 172, pp. 76–94, 2016.

[21] Q. Wang, X. Wang, Q. Lv, X. Ye, Yi Luo, and L. You, "Analysis of the information theoretically secret key agreement by public discussion," *Security and Communication Networks*, vol. 8, no. 15, pp. 2507–2523, 2015.

[22] Q. Wang, K. Xu, and K. Ren, "Cooperative secret key generation from phase estimation in narrow band fading channels," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 9, pp. 1666–1674, 2012.

[23] A. D. Wood, J. A. Stankovic, and G. Zhou, "Deejam: Defeating energy-efficient jamming in IEEE 802.15.4-based wireless networks," in *IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks, SECON 2007, Merged with IEEE International Workshop on Wireless AdHoc and Sensor Networks*, pp. 60–69, 2007.

[24] W. Xu, W. Trappe, Y. Zhang, and T. Wood, "The feasibility of launching and detecting jamming attacks in wireless networks," in *Proceedings of The 6th ACM International Symposium on Mobile Ad Hoc Networking and Computing (ACM MobiHoc'05)*, pp. 46–57, 2005.

[25] W. Xu, T. Wood, W. Trappe, and Y. Zhang, "Channel surfing and spatial retreats: defenses against wireless denial of service," in *The Workshop on Wireless Security*, pp. 80–89, 2004.

[26] J. Zhang, T. Duong, A. Marshall, and R. Woods, "Key generation from wireless channels: A review," *IEEE Access*, vol. 4, no. 3, pp. 614–626, 2016.

[27] C. T. Zenger, J. Zimmer, and C. Paar, "Security analysis of quantization schemes for channel-based key extraction," in *Workshop on Wireless Communication Security at the Physical Layer*, pp. 267–272, 2015.

[28] Y. Zou, J. Zhu, X. Wang, and L. Hanzo, "A survey on wireless security: Technical challenges, recent advances, and future trends," *Proceedings of the IEEE*, vol. 104, no. 9, pp. 1727–1765, 2016.

Biography

Qiuhua Wang received her B.S. and M.S. degrees in communication engineering from Liaoning Technical University, Fuxin, China, in 2000 and 2003, respectively. She received her Ph.D. degree in Communications and Information systems from Zhejiang University, Hangzhou, China, in 2013. Now, she is an Associate Professor of the School of CyberSpace, Hangzhou Dianzi University. Her current research interests include information security, security issues in wireless networks, key management and physical layer security, etc.

Hongxia Zhang received her B.S. degree in Electronic and Information engineering from Shandong University of Technology, Zibo, China, in 2015. She is currently pursuing the M.S. degree in Communication Engineering. Her research interests include key generation and physical layer security.

Qiuyun Lyu received her B.S. and M.S. degrees in Computer Science and Technology from Chang'an University, Xi'an, China, in 2000 and 2003, respectively. Now, she is an Associate Professor of the School of CyberSpace, Hangzhou Dianzi University. Her current research interests include information security and privacy, security issues in wireless networks.

XiaojunWang received his B.S. and M.S. degrees in Communication and Information system from University of Electronic Science and Technology of China, Chengdu, china, in 1997 and 2000 respectively. Now, he is a teacher of the School of CyberSpace, Hangzhou Dianzi University. His research interests include information security, vulnerability analysis and software security.

Jianrong Bao received his B.S. degree in Polymer Materials & Eng., and the M.S.E.E. degree from Zhejiang University of Technology, Hangzhou, China, in 2000 and 2004, respectively. He received his Ph.D. E.E. degree from the Department of Electronic Engineering, Tsinghua University, Beijing, China, in 2009. He is with the school of Information Engineering, Hangzhou Dianzi University, Hangzhou, China. His research interests include space wireless communications, communication signal processing, information security & channel coding, etc.

DNA Cryptography for Secure Data Storage in Cloud

Sreeja Cherillath Sukumaran[1], Misbahuddin Mohammed[2]
(Corresponding author: Sreeja Cherillath Sukumaran)
Department of Computer Science, Christ University[1]
Hosur Road, Bangalore, Karnataka 560029, India
Computer Networks and Internet Engineering Division, Centre for Development of Advanced Computing (C-DAC)[2]
Electronics City, Bangalore, Karnataka 560100, India
(Email: sreejasukumaran@gmail.com)
(Received Dec. 01, 2016; revised and accepted Mar. 11, 2017)

Abstract

Cloud computing has revolutionized the way the data is stored, processed and made available. It has evolved in various forms of utility computing by sharing resources, infrastructures and data storage facilities and got wide acceptance because of its services and storage capacities. But security issues are a major concern in the cloud which is restricting its use among organizations which deals with sensitive data such as health care, Pharmaceuticals etc. and remain one of the greatest inhibitors for the adoption of Cloud computing if the security issues continue. For data protection, various techniques evolved through years for Ciphers, Cryptography, Steganography and recently DNA based encryption for security is the trend. DNA cryptography was a breakthrough in the field of security which uses bio-molecular concepts and gives us new hope of unbreakable algorithms but the concepts need to be exploited more especially in the cloud computing. This paper discusses cloud computing features, service models, security issues and proposes a DNA based encryption algorithm for securing data in cloud environment which will be cost effective and secure by using bio-computational techniques. The suggested algorithm uses indexing and DNA steganography techniques along with binary coding rules which make algorithm secure as it is an additional layer of biosecurity than conventional cryptographic techniques.

Keywords: Cloud Computing; Confidentiality; Data Security; DNA; DNA Encryption Techniques; Integrity

1 Introduction

Cloud computing received relevant diligence and has a vital role in the growth of information technology as it provides essential infrastructure, platform and software as services [13]. Cloud storage enables to handle a large volume of data which is crucial for business and individuals and a solution for storing exploding data. Cloud computing will become an essential part for everyone as the future is of Big data. The main features of the cloud include scalability, reliability and availability but same time information security, privacy and compliance are major concerns or issues which inhibit the growth and migration to the cloud by organizations especially those who deal with sensitive data. Security is always a major concern in Open System Architectures and considering the threats and vulnerabilities in the cloud various countermeasures has been proposed till now including cryptographic techniques but still there is a need for novel and cost effective techniques to thwart the attacks and biocomputing techniques are a solution to this as it provides novel and secure techniques which enhances the security by a bio-layer which is difficult to break.

2 Literature Review

This section discusses various aspects of Cloud Computing,data security issues and challenges in the Cloud, DNA computing, DNA, DNA Cryptography and a review of literature related to these concepts.

NIST defines [11] Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. The cloud model is composed of five essential characteristics which are on-demand self-service,broad network access,resource pooling,rapid elasticity and measured service, Cloud service models are also known as SPI model. They are Software as a Service (SaaS), Platform as a Service (PaaS) and Infrastructure as a Service (IaaS). Depending upon the customer requirements the Cloud ser-

vices can be deployed as Private cloud,Community cloud, Public cloud and hybrid cloud [8].

2.1 Security Issues in Cloud

Security issues in the cloud is a major concern and a barrier to its adoption by organizations. The virtualized nature of cloud storage makes the traditional mechanisms unsuitable for handling the security issues [5, 14, 18]. SaaS relieves the users from tasks like maintenance and installation of software. It has been used widely all around. Web 2.0, a key technology towards enabling SaaS and its usage is increasing drastically by user community. So, these environments really need a security.

Various security issues exist in cloud some of them are issues in the network level security, application level security but data security is a major concern in the cloud. It may be the security of data during data transfer, securing of data leaving a data- center to another data - center or the data at rest. Data security is most prominent considering the fact apart from business domains, organizations which deal with sensitive information such as health care is also using information technology for providing on time and better treatment for patients [10].

2.2 Data Storage in Cloud

Data security is a major concern in any technology but will become a critical aspect when both customer data and programs are residing at provider premises as in the case of Cloud. This becomes a major challenge in SaaS as the data is often processed in plain text and stored in the cloud and the SaaS users must rely on their providers for proper security [4, 20, 22]. In [17] reviews the data security and privacy issues in cloud computing with special emphasis on data confidentiality, Integrity, and availability. The data storage and confidentiality issues are the major barriers in cloud computing and resolving these issues by enhanced data protection and encryption techniques increases the user's trust which in turn removes the barriers of cloud adoption. Data Integrity and confidentiality are the most critical factors for users storing their confidential data in the cloud. Encryption techniques plays a major role in achieving confidentiality whereas digital signatures play a vital role to achieve data integrity.

Challenges and issues related to data storage remains a major threat for cloud adoption as the client will not have direct control of data stored in the Cloud. Cloud computing has greater security threats and vulnerable to attacks if integrity, confidentiality and privacy of the data is not ensured in the cloud environment. The authors [15] focuses on the security issues related to the confidentiality and privacy of the data stored in the cloud. The research related to data storage and security always emphasizes the need for novel techniques.

2.3 Aspects of Data Security

There are various aspects of data security that need to be taken care while migrating to a cloud environment which includes [3]

- Data-in-transit;
- Data-at-rest;
- Data Lineage;
- Data Remanence;
- Data Provenance.

Data-in-transit is also referred as data in motion and it has highest security risk in the data security aspects. The encryption technique used plays a major role along with a secure transmission protocol secures data to some extent.

Data-at-rest has security issues especially in a shared environment, so encrypting data using secure and strong encryption plays a major role. Data Lineage is critical for cloud computing as it is tracing the path of data which is very difficult in distributed environment.

Data Provenance refer to maintaining the integrity of data which is very important in cloud environment. Data Remanence refers to the data left out especially due to transfer of data or its removal. It causes minimal security but may turn critical for public cloud.

Considering the data security aspects, confidentiality and integrity are the major concerns for data residing in Cloud. For securing the data stored in the cloud the service providers use various Cryptographic techniques such as public key encryption, private key encryption and homomorphic encryption of which homomorphic technique is considered as best for data-at-rest. Security challenges in the cloud (SaaS) is similar to that of any web application technology but traditional security techniques are not enough to resist attacks and which indicates the need for innovative and secure technologies. DNA computing techniques can be exploited and applied for securing data in cloud environment as it provides bio-computational complexity and can be hybridized with conventional encryption techniques.

2.4 DNA Computing

Biocomputing is a trending concept which has application in cryptography to generate secure algorithms using the computational complexity of biomolecular concepts in addition to conventional cryptographic techniques. Biomolecular encryption techniques have given rise to a new branch of cryptography, DNA cryptography as the concepts mainly revolves around DNA and the central dogma of molecular biology. DNA, Deoxyribonucleic acid is a double-stranded helix of nucleotides with each nucleotide containing one of four bases A, G, C, T where A stands for adenine, G for guanine, C for cytosine and T for thymine respectively [2, 9].

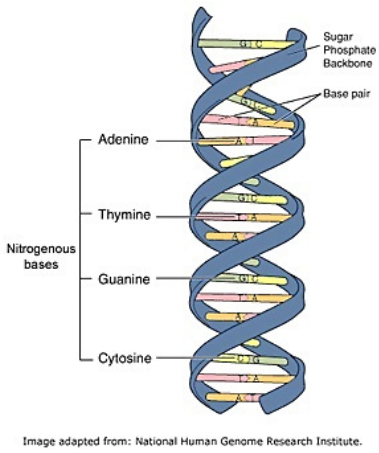

Figure 1: Helical structure of DNA

Figure 1 is the helical structure of DNA, which depicts nitrogenous bases, sugar phosphate backbone and base pair bonding [21]. Figure 2 represents the structure of DNA base pairs and the bonding between adenine and thymine, cytosine and guanine [19]. In eukaryotic genes, the coding regions referred as exons and noncoding regions are introns and the exons are interrupted by introns. Figure3 shows coding and non -coding regions in a segment of eukaryotic DNA [7].

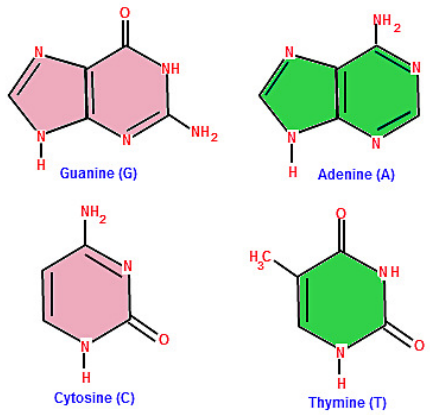

Figure 2: Structure of DNA base pairs

2.4.1 DNA Encryption Techniques for Information Security

Computational properties of DNA became a new branch of science and a research area for cryptographers from 1994 when Dr. Leonard M. Adleman used the computational properties of DNA to solve Hamiltonian path problem [1]. Research on DNA based encryption techniques can be broadly classified into three:

- DNA Cryptography;
- DNA Steganography;

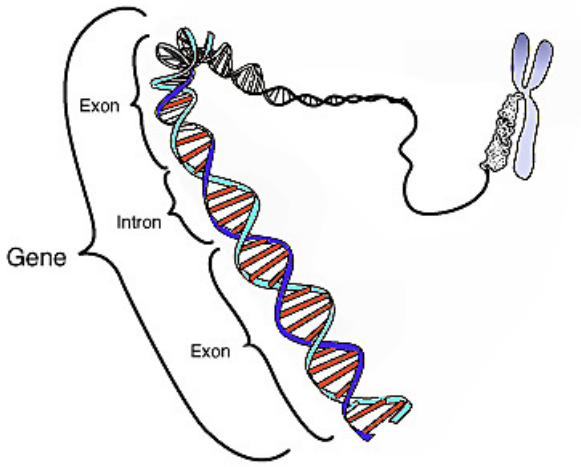

Figure 3: Structure coding region in a segment of eukaryotic DNA

- Pseudo DNA Cryptography.

DNA Cryptography is built on DNA based computations for encryption and various methodologies for both Symmetric and Asymmetric DNA cryptography has been proposed by researchers. The techniques based on biocomputations got wide acceptance due to secure nature of algorithms.

DNA Steganography became popular with a patented technology proposed by Carter Bancroft for hiding messages. The proposed technique involved concealing a DNA-encoded message within a genomic DNA sample followed by further concealment of the DNA sample to a microdot [6]. DNA Steganography is also considered as cryptographic technique even though there is no encryption.Pseudo DNA Cryptography got wide acceptance as it involves simulation by means of computations and arithmetic operations on Pseudo DNA instead of using real DNA with high-tech lab facilities.

DNA digital coding technology denotes the bases A, C, T, G as 00, 01,10,11 and the binary values can be interchanged with bases and this coding forms the basis of algorithms using Digital DNA [16]. DNA computing can be performed in two ways one by the means of biological operations using real DNA and the other technique involves simulation using Digital DNA and Pseudo DNA.

2.5 DNA and Cloud Computing

DNA has an excellent storage capacity of 10^6 TB in 1gm of DNA, which indicates a few grams of DNA may have the capacity to store whole data available in the world [12]. Cloud computing techniques also provide data storage so instead of using DNA for data storage exploiting DNA based encryption for data stored in the cloud will be a combination of trending techniques which ensures data security, Confidentiality, integrity and authentication are considered as most important aspects of in-

formation security whereas context is also an important aspect depending on it, the level required for data security will be different. Cryptographic techniques have a major role in securing data and depending on the context different encryption techniques can be used. The encryption techniques are broadly classified as symmetric key cryptography and asymmetric key cryptography depending on the context it can be used in conjunction with DNA computing.

3 Proposed Methodology

In this section, a methodology is proposed for securing data for cloud storage using DNA based encryption technique. In this method, DNA based coding, encryption technique, DNA Steganography and indexing method are proposed for securing data in the cloud.

The Proposed DNA Cryptography technique is different from that of the DNA cryptography which uses real DNA Sequences or Oligos, as the computations performed are using the digital DNA. A DNA sequence consists of four nucleic acid bases A (adenine), C (cytosine), G (guanine), T (thymine), where A and T are complementary, and G and C are complementary. Table 1 represents one of the gray coding method used for DNA bases.

Table 1: Binary coding based on DNA

Bases	Gray Coding
A	00
G	01
C	10
T	11

Figure4 represents architecture of the proposed DNA based encryption. Figure 5 represents the sequence diagram of the proposed system, where user encrypts the data and stores encrypted data into cloud. Decryption of the data is done after retrieving it from the cloud, which ensures confidentiality of the data as the encryption and decryption is performed at the client side. Figure6 represents architecture of proposed DNA based decryption

Algorithm 1 Data encryption algorithm

1: Begin
2: Input: *Data to be stored D,Random DNA R-DNA.*
3: Select the data D, which has to be stored securely in the cloud, Convert the data into binary, say BD'.
4: Convert the binary data into DNA sequence based on the DNA coding rule as per Table1,which generates a digital DNA→ D'DNA.
5: Create a random DNA strand by selecting DNA sequences from the digital databases→ R-DNA.
6: Select the R-DNA and index it. Select the coding and non- coding regions randomly or based on the index values.
7: Convert indexed R-DNA into short fragments based on the length of D'DNA base pair, and a key value based on D'DNA.
8: Remove the non-coding region and the generated DNA sequence is used as a cover for adding D'DNA.
9: Insert the D'DNA into non-coding regions of the generated R-DNA based on the index positions or random position depending on the indexing rule selected.
10: The resultant DNA sequence generated by DNA Steganography is converted into binary form using the selected binary rule.
11: Upload the encrypted data in the binary form and store it in the cloud.
12: Output: *Encrypted Data*
13: End

Algorithm 2 Data decryption algorithm

1: Begin
2: Input: *Encrypted Data.*
3: Extract the encrypted data from the cloud which is in binary form.
4: Apply the selected DNA binary coding rule to the data and get the data in the form of DNA sequence which is a combination of R-DNA and D'DNA.
5: Based on the index value position,select the coding and non-coding regions of the DNA.
6: Retrieve DNA fragments from the non-coding region.
7: Extract and separate D'DNA and R-DNA from the DNA Sequence.
8: Append the fragments of D'DNA and apply DNA coding rule to get the binary data.
9: Convert the binary to ASCII.
10: Convert ASCII to text.
11: Generates the original data.
12: Output: *Original Data and Random DNA*
13: End

4 Proof of Concept

Step 1. Let us consider the data to be stored in the cloud by the user is → MyConfidential Data.

Step 2. Convert the data to be stored into the Binary form.

MyConfidentialData → 01001101 01111001 01000011 01101111 01101110 01100110 01101001 01100100 01100101 01101110 01110100 01101001 01100001 01101100 01000100 01100001 01110100 01100001 → (1).

Step 3. Apply DNA binary Coding as per Table1 to binary data of Table2 or on (1) generates (2) in DNA form.A,G,C and T values can be altered and used as per user convenience depending on the selected

Table 2: Binary conversion of the data

Data to be stored in the Cloud	ASCII Value	Binary Value
My	77 121	01001101 01111001
Confidential	67 111 110 102 105 100 101 110 116 105 97 108	01000011 01101111 01101110 01100110 01101001 01100100 01100101 01101110 01110100 01101001 01100001 01101100
Data	68 97 116 97	01000100 01100001 01110100 01100001

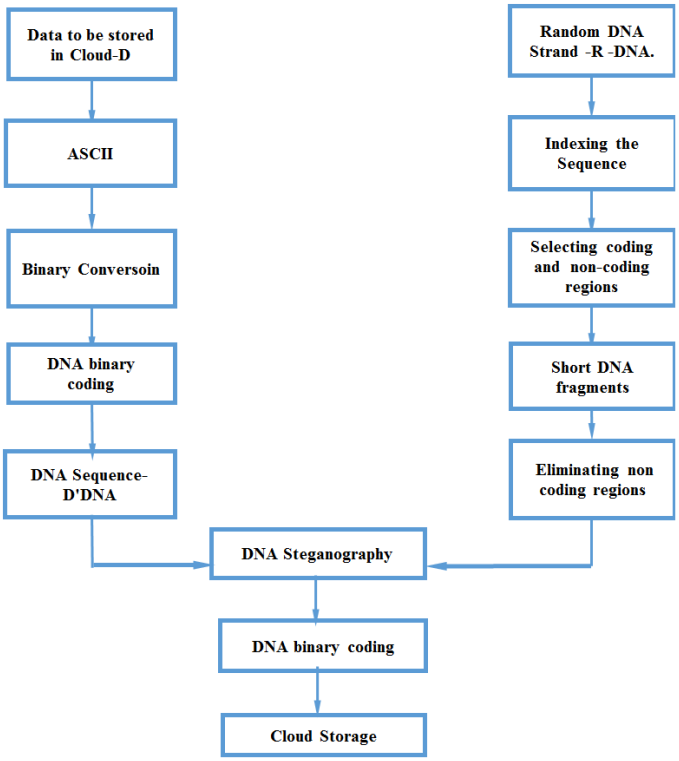

Figure 4: DNA encryption architecture

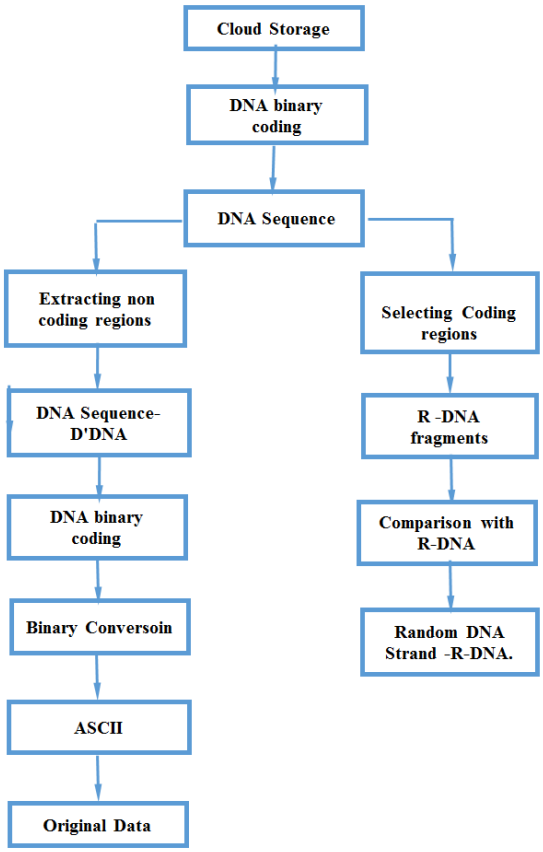

Figure 6: DNA decryption architecture

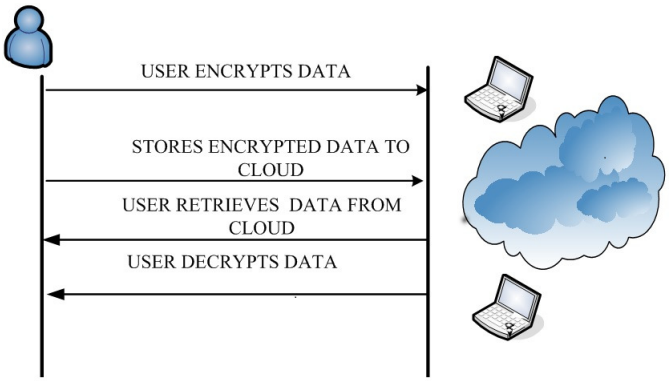

Figure 5: Sequence diagram of the proposed system

binary coding for the bases.

Table 3: Binary coding used in the proof

Bases	Gray Coding
A	00
G	01
C	10
T	11

GATG GTCG GAAT GCTT GCTC GCGC GCCG
GCGA GCGG GCTC GTGA GCCG GCAG GCTA

GAGA GCAG GTGA GCAG (D'DNA) → (2)

Step 4. Select Random DNA→ R-DNA. The random DNA is generated by downloading the sequence available from NCBI. The sequence used in this study is Reference Sequence,NC_006583.3 Canis lupus familiar is breed boxer chromosome 1, CanFam3.1, whole genome shotgun sequence is downloaded and a fragment of the sequence has been selected as R-DNA.A Sample of 32 bases is selected and indexed. Sample R-DNA selected is: ACTGCTGAGAGTTGAGCTCACCCTC AGTCCCTCACAGTTCCACACTGCCT

Step 5. Indexing the R-DNA.

$$A_1C_2T_3G_4C_5T_6G_7A_8G_9A_{10}G_{11}T_{12}$$

$$T_{13}G_{14}A_{15}G_{16}C_{17}T_{18}C_{19}A_{20}C_{21}C_{22}C_{23}$$

$$T_{24}C_{25}A_{26}G_{27}T_{28}C_{29}C_{30}C_{31}T_{32}$$

Step 6. Select the coding and non-coding regions randomly or based on the index positions.

Step 7. This sample is demonstrated using a random selection of coding and non-coding regions. Depending on the security of the data, complexity of the algorithm can be increased by defining index rules.

Step 8. Insert the D'DNA into non-coding regions of R-DNA. In this sample based on index values the coding and non-coding regions defined are:

$$A_1C_2T_3G_4C_5$$

→ Coding region

$$T_6G_7A_8G_9A_{10}G_{11}T_{12}T_{13}G_{14}A_{15}$$
$$G_{16}C_{17}T_{18}C_{19}A_{20}C_{21}C_{22}C_{23}$$

→ Non-Coding region

$$T_{24}C_{25}A_{26}G_{27}T_{28}C_{29}C_{30}C_{31}T_{32}$$

→ Coding region from the R-DNA

Step 9. Insert D'DNA into respective non-coding index positions.Random DNA selected act as cover medium to insert D'DNA and performs DNA steganography.

GATG GTCG GAAT GCTT GCTC GCGC GCCG GCGA GCGG GCTC GTGA GCCG GCAG GCTA GAGA GCAG GTGA GCAG → D'DNA Replacing Non-Coding region bases with 4bases (key value) of D'DNA per each base of non-coding region:

$$T_6G_7A_8G_9A_{10}G_{11}T_{12}T_{13}G_{14}A_{15}$$

$G_{16}C_{17}T_{18}C_{19}A_{20}C_{21}C_{22}C_{23}$ → Cover DNA(non-coding region)

$(GATG)_6(GTCG)_7(GAAT)_8(GCTT)_9(GCTC)_{10}$ $(GCGC)_{11}(GCCG)_{12}(GCGA)_{13}(GCAG)_{18}(GCTA)_{19}$ $(GAGA)_{20}(GCAG)_{21}(GTGA)_{22}(GCAG)_{23}$ → DNA Steganography.

Key value can be increased depending on the required security and length of the original data.

Step 10. Generate the DNA Sequence.

$A_1C_2T_3G_4C_5(GATG)_6(GTCG)_7(GAAT)_8(GCTT)_9$ $(GCTC)_{10}(GCGC)_{11}(GCCG)_{12}(GCGA)_{13}(GCGG)_{14}$ $(GCTC)_{15}(GTGA)_{16}(GCCG)_{17}(GCAG)_{18}(GCTA)_{19}$ $(GAGA)_{20}(GCAG)_{21}(GTGA)_{22}(GCAG)_{23}T_{24}C_{25}A_{26}$ $G_{27}T_{28}C_{29}C_{30}C_{31}T_{32}$ → In indexed form.

ACTGCGATGGTCGGAATGCTTGCTCGCGCG CCGGCGAGCGGGCTCGTGAGCCGGCAGGCT AGAGAGCAGGTGAGCAGTCAGTCCCTCACA GTTCCACACTGCCT → Stego DNA

Step 11. Convert the DNA Sequence into DNA based binary coding depending on the coding rule selected.

ACTGCGATGGTCGGAATGCTTGCTCGCGCG CCGGCGAGCGGGCTCGTGAGCCGGCAGGCT AGAGAGCAGGTGAGCAGTCAGTCCCTCACA GTTCCACACTGCCT → Encrypted Data in DNA form.

00101101100100110101111001010000110110111101 10111001100110011010010110010001100101011011 10011101000110100101100001011011000100010001 10000101110100011000011110000111101010111000 10000111111010001000101101101011 → Encrypted Data in binary form

Step 12. Store the data into cloud in the binary form or in integer form by converting the binary data to integers depending on the convenience of user.

5 Security Analysis

Security analysis of the proposed system is done for integrity and confidentiality which are considered as major threats for data stored in the Cloud environment.

5.1 Confidentiality of the Data

In the proposed algorithm to retrieve the encrypted data user must have the knowledge of selected random sequence, non-coding and coding regions, index value, binary coding rule used for the D'DNA, key value and DNA Sequence coding which increase the complexity of the algorithm.The knowledge of all these parameters are essential for recovering or decrypting the original data which ensures the data confidentiality.

Probability of finding the random DNA sequence $=(\frac{1}{163}) \times 10^8 \times \frac{selectingthefragmentbasesofR-DNA}{Totalnumberofbasesinselected Genome}$

The random sequence is selected from NCBI which has millions of seq, in the same way millions of sequences are available in other data bases that can be exploited

for creating R-DNA which also ensures security as it is difficult to find the sequence selected for the encryption.

$$\begin{aligned} & Complexity \\ = \; & (R-DNA) \times (BinarycodingRule) \times (index) \\ & \times (keyvalue(n)) \times (knowledgeofintrons(K_i)) \\ & \times (knowledgeofexons(K_e)) \\ = \; & (\frac{1}{163}) \times 10^8 \\ & \times \frac{selectingthefragmentbasesofR-DNA}{TotalnumberofbasesinselectedGenome} \\ & \times \frac{1}{24} \times (I_i^n) \times (n) \times (K_i) \times (K_e). \end{aligned}$$

The output of the proposed encryption algorithm can be saved even in the public cloud without any issues as it is difficult to break the system considering the computational complexity and the knowledge factors required.

5.2 Data Integrity Check Using SHA-2

The proposed system can be used as client-side encryption. To ensure the integrity of the data the user can generate a hash of the encrypted data and store it in a local repository. After retrieving data from the cloud each time user can compute the hash of data and can be compared with the hash value value stored in the local hash repository to ensure data integrity.

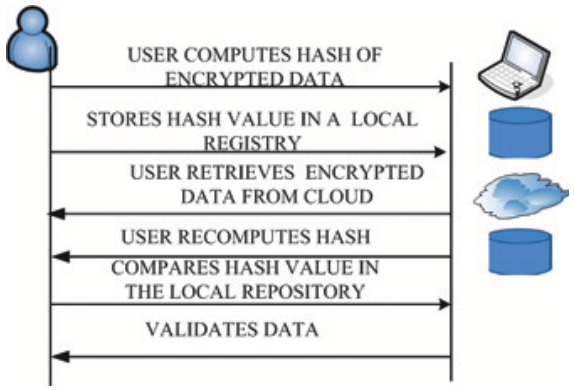

Figure 7: Sequence diagram for data integrity check

The proposed technique is feasible and easy to use and ensures the data integrity using hashing technique. Figure 7 is depicting the sequence diagram of the proposed system for data integrity check. SHA-256 is the suggested one as it is collision resistant. The hash value computed for the encrypted data using proposed methodology is df1140d3a68f05e6bccb16df7fa10f1b404d487994f5d441a9d8621549e0c685.

6 Conclusions and Future Work

Cloud computing is not widely accepted or restricted in few domains because of the security issues related to data storage. In this paper, a DNA based encryption technique is proposed for storing data securely in the cloud especially in the public cloud where data storage is a major concern and for SaaS users where security is a major concern. The technique will provide enhanced security as it adds the computational complexity by using biocomputing techniques in addition to Cryptography and user can check the integrity of the data without relying on the third party.

The proposed DNA Cryptography is a novel encryption technique for secure storage of data in the cloud environment, the method is still primitive, but using DNA cryptography for cloud has great Scope considering the importance of cloud storage in the Industries and day to day life. Everywhere data is bombarding in the form of images, videos and other digital forms. So, a platform for storage is very essential and DNA encryption is a trending concept which is going to dominate the security world in the future.

Table 4: Hash value generated for data integrity check

Encrypted Data	SHA-2
0010110110010	df1140d3a
011010111100	68f05e6bcc
1010000110110	b16df7fa1
11110110111001	0f1b404d48
10011001101001	49e0c6850
01100100011001	
01011011100111	
010001101001011	
000010110110001	
000100011000010	
111010001100001	
1110000111101010	
111000100001111	
110100010001011	
01101011	

References

[1] L. M. Adleman, "Molecular computation of solutions to combinatorial problems," *Nature*, vol. 369, pp. 40, 1994.

[2] E. S. Babu, C. N. Raju, and M. H. K. Prasad, "Inspired pseudo biotic DNA based cryptographic mechanism against adaptive cryptographic attacks," *International Journal of Network Security*, vol. 18, no. 2, pp. 291-303, 2016.

[3] R. Bhadauria and S. Sanyal, "Survey on security issues in cloud computing and associated mitigation techniques," *arXiv preprint arXiv:1204.0764*, 2012.

[4] K. Brindha and N. Jeyanthi, "Securing portable document format file using extended visual cryptography to protect cloud data storage," *International*

Journal of Network Security, vol. 19, no. 5, pp. 684-693, 2017.

[5] Z. Cao, C. Mao, L. Liu, "Analysis of one secure anti-collusion data sharing scheme for dynamic groups in the cloud," *International Journal of Electronics and Information Engineering*, vol. 5, no. 2, pp. 68–72, 2016.

[6] C. T. Clelland, V. Risca, and C. Bancroft, "Hiding messages in dna microdots," *Nature*, vol. 399, no. 6736, pp. 533–534, 1999.

[7] Cold Spring Harbor Laboratory, *Exons and Introns*, Oct. 23, 2016. (http https://www.dnalc.org/view/15549-transcrition-translation-exons-and-introns.html)

[8] M. Hogan, F. Liu, A. Sokol, and J. Tong, "NIST cloud computing standards roadmap," *NIST Special Publication*, vol. 35, 2011.

[9] X. Li, C. Zhou, N. Xu, "A secure and efficient image encryption algorithm based on DNA coding and spatiotemporal chaos," *International Journal of Network Security*, vol. 20, no. 1, pp. 110-120, 2018.

[10] L. Liu, W. Kong, Z. Cao, J. Wang, "Analysis of one certificateless encryption for secure data sharing in public clouds," *International Journal of Electronics and Information Engineering*, vol. 6, no. 2, pp. 110-115, 2017.

[11] P. Mell, T. Grance *et al.*, "The NIST definition of cloud computing," 2011.

[12] M. Misbahuddin and C. Sreeja, "A secure image-based authentication scheme employing dna crypto and steganography," in *Proceedings of the Third International Symposium on Women in Computing and Informatics*, pp. 595–601, 2015.

[13] A. Mosa, H. M. El-Bakry, S. M. Abd El-Razek, S. Q. Hasan, "A proposed E-government framework based on cloud service architecture," *International Journal of Electronics and Information Engineering*, vol. 5, no. 2, pp. 93–104, 2016.

[14] D. Patel, "Accountability in cloud computing by means of chain of trust dipen contractor," *International Journal of Network Security*, vol. 19, no. 2, pp. 251-259, 2017.

[15] B. T. Rao and N. vurukonda, "A study on data storage security issues in cloud computing," *Procedia Computer Science*, vol. 92, pp. 128–135, 2016.

[16] C. Sreeja, M. Misbahuddin, and N. Mohammed Hashim, "Dna for information security: A survey on dna computing and a pseudo dna method based on central dogma of molecular biology," in *International Conference on Computer and Communications Technologies (ICCCT'14)*, pp. 1–6, 2014.

[17] Y. Sun, J. Zhang, Y. Xiong, and G. Zhu, "Data security and privacy in cloud computing," *International Journal of Distributed Sensor Networks*, 2014.

[18] W. L. Tai, Y. F. Chang, "Comments on a secure authentication scheme for IoT and cloud servers," *International Journal of Network Security*, vol. 19, no. 4, pp. 648-651, 2017.

[19] Tutorvista, *Deoxyribonucleic Acid*, Nov. 23, 2016. (http://chemistry.tutorvista.com/biochemistry/deoxyribonucleic-acid.html)

[20] N. Vaanchig, H. Xiong, W. Chen, Z. Qin, "Achieving collaborative cloud data storage by key-escrow-free multi-authority CP-ABE scheme with dual-revocation," *International Journal of Network Security*, vol. 20, no. 1, pp. 95-109, 2018.

[21] Virtual Genetics Education Centre, *DNA, Genes and Chromosomes*, Nov. 22, 2016. (http://www2.le.ac.uk/departments/genetics/vgec/highereducation/topics/dnageneschromosomes)

[22] Z. Wang, Y. Lu, G. Sun, "A policy-based de-duplication mechanism for securing cloud storage," *International Journal of Electronics and Information Engineering*, vol. 2, no. 2, pp. 70–79, 2015.

Biography

Sreeja C.S. did her B.Sc. (Industrial Chemistry) from University of Calicut -Kerala, MCA from Bharathiar University-Tamilnadu and M.Phil in computer science from Christ University-Bangalore. She is currently pursuing her PhD in Christ University-Bangalore under the guidance of Dr. Mohammed Misbahuddin. Her area of interests in research includes Information Security, Biocomputing, DNA Cryptography, Steganography, Authentication Public Key Cryptography and Cloud Security.

Dr. Mohammed Misbahuddin did his B. Tech (CSE) from Gulbarga University, M. Tech (S/w Engg.) from JNTU-Anantapur and PhD (CSE) in Network Security from JNTU Hyderabad. He is currently working as Principal Technical Officer (Scientist D) in Centre for Development of Advanced Computing (C-DAC), E-City, Bangalore, where he is a key member of projects in the areas of PKI and e-Authentication. He is the Co-Investigator of a National Project named e-Pramaan A National e-Authentication Service along with Aadhaar. He has 15 years of experience in Research, Training and Project Management. He has applied 3 patents with IPO in the area of Secure and Usable Authentication. He has been in various Programme committees of IEEE /ACM conferences and is a reviewer for two International Journals. His area of interest is Network Security and Cryptography especially Secure and Usable Authentication, Public Key Cryptography, Risk based Engines, Cloud Security and DNA Cryptography.

A New SPN Type Architecture to Strengthen Block Cipher Against Fault Attack

Gitika Maity[1], Jaydeb Bhaumik[2], and Anjan Kundu[3]
(Corresponding author: Gitika Maity)
Computer Science and Engineering, Haldia Institute of Technology[1]
I.C.A.R.E Complex, H.I.T Campus, Hatiberia, PO-HIT, Midnapore, West Bengal 721657, India
(Email: Gitika.Maity@gmail.com)
Department of Electronics & Communications Engineering, Haldia Institute of Technology[2]
Institute of Radio Physics and Electronics, University of Calcutta[3]
(Received Dec. 2, 2016; revised and accepted Apr. 11 & May 14, 2017)

Abstract

In recent years, Differential Fault Analysis (DFA) has been proven as the most efficient technique to attack any block cipher by introducing a computational error. In this paper, a new Substitution Permutation Network (SPN) type architecture is proposed which has better resistance against DFA as compared to Advanced Encryption Standard (AES). The proposed architecture is similar to AES except round key mixing function. Here, round key is mixed with round output, using nonlinear vectorial Boolean function called 'Nmix'. Using 4 faulty-fault free ciphertext pairs, 32 bits of 10^{th} round key is retrieved by injecting a random byte fault at the input of 9^{th} round. The computational complexity will be in the order of 2^{36} to obtain 128 bits 10^{th} round key. Total 16 numbers of faulty and fault free ciphertext pairs are required. Similarly, when a fault is injected at the input of 8^{th} round, then the 10^{th} round key is obtained with computational complexity of 2^{53} and 20 numbers of faulty-fault free ciphertext pairs are required.

Keywords: Block Cipher; Fault Attack; Nonlinear Boolean Function; Substitution and Permutation Network

1 Introduction

Cryptography is an important mathematical tool which is used to provide security in several systems like e-Commerce, RFID, sensor network, mobile phones, smart cards, personal digital assistants (PDAs) etc. Cryptographic algorithms are mainly used to satisfy a subset of four cryptographic properties namely confidentiality, message integrity, authentication and non repudiation. For high speed applications algorithms are usually implemented in hardware. But when implemented in ASIC or FPGA, mathematical security of cryptographic algorithms are not sufficient and hence susceptible to fault attack. The fault is introduced by attacker during the execution of cryptographic algorithm to derive the secret key. This type of fault attack was introduced by Boneh, DeMillo and Lipton [8, 9]. Subsequently Differential Fault Analysis (DFA) on secret key cryptosystem has been discussed by Biham *et al.* in [6]. They shows lower complexity compared to simplified fault attack.

The US National Institute of Standard and Technology (NIST) selected Rijndael as the Advanced Encryption Standard (AES) in 2000 [11]. This algorithm has been adopted as a world wide standard for symmetric key encryption. Till date, fault based attack on advanced encryption algorithm has lowest computational complexity compared to all other attacks. Presently very low cost methods are used for fault injection such as variation of supply voltages, clock glitches, temperature variation, UV light radiations etc. Optimal fault injection method has been reported in [21]. How to find key from algebraic equation, is discussed in [15]. Fault attack by inducing byte level fault at the input of 9^{th} round of AES has been reported in [16], where 250 faulty ciphertexts are needed to recover the key. DFA against AES analyzed by Dusart *et al.* in [7]. They show that by injecting fault at byte level in the 8^{th} round and 9^{th} round, the attacker can derive the key using 40 ciphertext pairs. A survey on fault attack against AES and their counter measures are discussed in [2]. Several counter measures are proposed to resist fault attack on hardware implementation of block cipher AES. Counter measure techniques are hardware redundancy, time redundancy, information redundancy and hybrid redundancy. Differential Fault Analysis on ultra-lightweight cipher PRESENT, has been delineated in [10]. To recover the secret key, it takes 2 faulty encryptions and an exhaustive search of 2^{16}. An improved fault attack against Eta Pairing is described in [13]. An improved fault attack against Miller's algorithm has been

presented in [14]. Differential power attack resistant Rijndael circuit is presented in [1]. In [20], it is shown that, by introducing fault at byte level in the 8^{th} round and 9^{th} round of 128 bits AES algorithm, attacker can easily recover the total key using two faulty ciphertexts. A fault based attack on a modified version of AES called MDS-AES has been reported in [12], where one pair of faulty-fault free ciphertext are used to derive 10^{th} round key with a computational complexity of 2^{16}. A complete differential fault analysis against LS-designs and on other families of SPN-based block ciphers have been shown in [17]. They have also validated DFA using a practical example of hardware implementation of SCREAM running on an FPGA. In [24, 19], the block cipher key were deduced by inducing a single random byte fault at the input of the eighth round of the AES algorithm. By exploiting the key-scheduling algorithm, DFA on AES reported in [23, 22]. It takes two faulty ciphertexts and a brute force search of 48 and 40 bits respectively.

In this paper, a modified SPN-type architecture has been proposed to strengthen it against fault attack without affecting area and time significantly. Here XOR operation in AddRoundKey step is replaced by a Boolean nonlinear Nmix function [4]. Effectiveness of the proposed architecture is then analysed against fault attack, by introducing a random byte fault at the input of 9^{th} round and 8^{th} round. The attacker has to search for 2^{36} times to obtain the desired 128 bit key. Also 16 numbers of faulty-fault free ciphertext pairs are necessary, which is much greater than the complexity of attacking original AES. When a random byte fault is introduced at the the input of 8^{th} round, then to recover 32 bits key it takes 5 faulty ciphertext pairs. So, the attacker has to search for approximately 2^{53} times to obtain the 128 bits key and 20 faulty-fault free ciphertext pairs are necessary.

This paper is organized as follows. Following the introduction, a description of proposed SPN type block cipher algorithm is given in Section 2. Fault analysis on proposed SPN-type architecture when fault is injected at the input of 9^{th} and 8^{th} round have been discussed in Sections 3 and 4 respectively. Comparison with existing works is discussed in Section 5. Finally the paper is concluded in Section 6.

2 Description of SPN Type Block Cipher Algorithm

The description of AES-Rijndael algorithm has been provided in detail by Daemen *et al.* in [11]. The proposed SPN type architecture is a modified version of AES-Rijndael algorithm. In the proposed architecture, SubByte, ShiftRow and MixColumn operations are exactly same as in AES. Only round key mixing operation of AES has been modified where XOR operation is replaced by Nmix. In this algorithm, key size and block size are 128 bits. Similar to AES, the 128 bits message block is arranged as a 4×4 array of bytes. The elements of the matrix are represented by variables s_{ij} where $0 \leq i \leq 3$ and $0 \leq j \leq 3$ where i, j denoting the row and column indexes of the state matrix. Number of round is 10 and

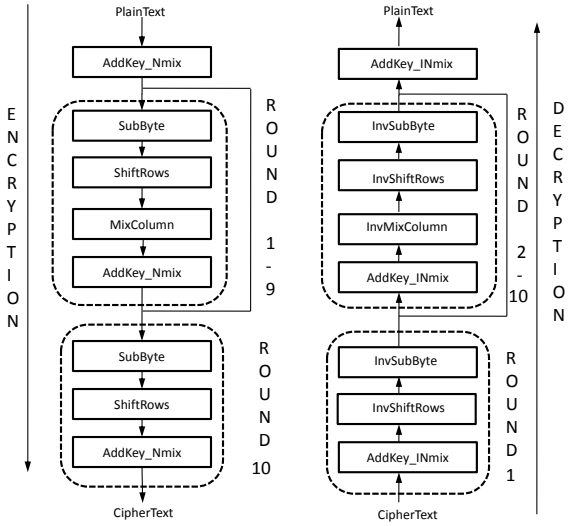

Figure 1: Block diagram of SPN type block cipher algorithm

in each round key is generated from cipher key by key expansion algorithm. At the end of 10^{th} round ciphertext is generated. Similar to AES-128, round 1-9 consists of SubByte, ShiftRow, MixColumn and Round Key Mixing. Round 10 consists of following 3 operations: SubByte, ShiftRows and Round Key Mixing. Basics of each functional block is as follows

SubBytes: This is a non linear substitution step where each byte is replaced by a new byte.

ShiftRows: In this step 2^{nd}, 3^{rd} and 4^{th} rows are circular shifted left by 1, 2 and 3 bytes respectively. First row remains unchanged.

MixColumns: Here the four bytes of each column of the state matrix are multiplied by the following matrix.

$$\begin{pmatrix} 02 & 03 & 01 & 01 \\ 01 & 02 & 03 & 01 \\ 01 & 01 & 02 & 03 \\ 03 & 01 & 01 & 02 \end{pmatrix}$$

Round Key Mixing: In each round the corresponding byte of state matrix are mixed with the generated roundkey. A non linear function Nmix is used here in AddRoundKey step of encryption. For decryption, the Inverse Nmix (INmix) is used. Details of Nmix and INmix has been discussed in [4]. In [5], Nmix function is used to design an integrated scheme for error correction and message authentication. In [3] nonlinear mixing function Nmix has been used to design the block cipher HDNM8. For the sake of completeness, an overview of Nmix and INmix is given in the following subsections.

2.1 Nonlinear Mixing (Nmix) Function

It operates on two n-bit variables $X = (x_{n-1}x_{n-2}....x_0)$ and $K = (k_{n-1}k_{n-2}....k_0)$ and produces an n-bit output $Y = (y_{n-1}y_{n-2}....y_0)$ where the each output bit y_i and carry bit c_i are defined as

$$\begin{aligned} y_i &= x_i \oplus k_i \oplus c_{i-1} \\ c_i &= \bigoplus_{j=0}^{i} x_j k_j \oplus x_{i-1}x_i \oplus k_{i-1}k_i \end{aligned} \quad (1)$$

where $0 \leq i \leq n-1$, $c_{-1} = 0$, $x_{-1} = 0$, $k_{-1} = 0$ and c_i is the carry propagated from bit position i^{th} to $(i+1)^{th}$. The end around carry c_{n-1} is ignored. Each output y_i is balanced for all $0 \leq i \leq n-1$.

In case of function Nmix, output XOR difference is not equal to input difference (XOR) when single input changes i.e $Nmix(A,K) \oplus Nmix(B,K) \neq A \oplus B$ where A, B and K are three n-bit variables. This property of Nmix function is utilized to strengthen proposed SPN-type architecture against fault attack.

2.2 Inverse Nonlinear Mixing (INmix) Function

INmix takes two n-bit variables $Y = (y_{n-1}y_{n-2}....y_0)$ and $K = (k_{n-1}k_{n-2}....k_0)$ as inputs and produces an $n-bit$ output $X = (x_{n-1}x_{n-2}....x_0)$, where each output bit x_i and carry d_i are defined as

$$\begin{aligned} x_i &= y_i \oplus k_i \oplus d_{i-1} \\ d_i &= \bigoplus_{j=0}^{i} x_j k_j \oplus x_{i-1}x_i \oplus k_{i-1}k_i \end{aligned} \quad (2)$$

where $0 \leq i \leq n-1$, $d_{-1} = 0$, $x_{-1} = 0$, $k_{-1} = 0$ and d_i is the carry propagated from bit position i^{th} to $(i+1)^{th}$. The end around carry d_{n-1} is ignored.

3 Fault Attack on Ninth Round of Proposed SPN-type Architecture

In this section, a single random non zero byte fault is induced in the first byte of 9^{th} round input. Propagation of fault is shown in Figure 2. After SubBytes operation, injected fault f has been change to $f^{'}$. Fault remains in the same position after ShiftRows and after MixColumns, it is distributed within 4 bytes of 1st column. If the attacker wants to retrieve 128 bit key then 4 byte faults have to be injected at 1^{st}, 5^{th}, 9^{th} and 13^{th} bytes respectively. Assume a fault is injected at the first byte of 9^{th} round input and let the expressions of $CT1$ and $CT2$ are

$$CT_1 = \begin{pmatrix} y_0 & y_4 & y_8 & y_{12} \\ y_1 & y_5 & y_9 & y_{13} \\ y_2 & y_6 & y_{10} & y_{14} \\ y_3 & y_7 & y_{11} & y_{15} \end{pmatrix}$$

$$CT_2 = \begin{pmatrix} (y_0 + F_1) & y_4 & y_8 & y_{12} \\ y_1 & y_5 & y_9 & (y_{13} + F2) \\ y_2 & y_6 & (y_{10} + F_3) & y_{14} \\ y_3 & (y_7 + F_4) & y_{11} & y_{15} \end{pmatrix}$$

The associated keys K_9 and K_{10} of 9^{th} and 10^{th} round are considered as

$$K_9 = \begin{pmatrix} k_{90} & k_{94} & k_{98} & k_{912} \\ k_{91} & k_{95} & k_{99} & k_{913} \\ k_{92} & k_{96} & k_{910} & k_{914} \\ k_{93} & k_{97} & k_{911} & k_{915} \end{pmatrix}$$

$$K_{10} = \begin{pmatrix} k_{100} & k_{104} & k_{108} & k_{1012} \\ k_{101} & k_{105} & k_{109} & k_{1013} \\ k_{102} & k_{106} & k_{1010} & k_{1014} \\ k_{103} & k_{107} & k_{1011} & k_{1015} \end{pmatrix}$$

From the fault pattern shown in Figure 2, following equations are constructed

$$\begin{aligned} &[(INmix(ISB(INmix(y_0, k_{100})), k_{90})) \oplus \\ &(INmix(ISB(INmix(y_0 + F_1), k_{100})), k_{90})] = \\ &2[(INmix(ISB(INmix(y_{13}, k_{1013})), k_{91})) \oplus \\ &(INmix(ISB(INmix(y_{13} + F_2), k_{1013})), k_{91})] \end{aligned} \quad (3)$$

$$\begin{aligned} &(INmix(ISB(INmix(y_{13}, k_{1013})), k_{91})) \oplus \\ &(INmix(ISB(INmix(y_{13} + F_2), k_{1013})), k_{91}) = \\ &(INmix(ISB(INmix(y_{10}, k_{1010})), k_{92})) \oplus \\ &(INmix(ISB(INmix(y_{10} + F_3), k_{1010})), k_{92}) \end{aligned} \quad (4)$$

$$\begin{aligned} &[(INmix(ISB(INmix(y_7, k_{107})), k_{93})) \oplus \\ &(INmix(ISB(INmix(y_7 + F_4), k_{107})), k_{93})] = \\ &3[(INmix(ISB(INmix(y_{13}, k_{1013})), k_{92})) \oplus \\ &(INmix(ISB(INmix(y_{13} + F_2), k_{1013})), k_{92})] \end{aligned} \quad (5)$$

In Equations (3), (4) and (5) the keys k_{100}, k_{107}, k_{1010}, k_{1013} are the 10^{th} round keys and k_{90}, k_{91}, k_{92} and k_{93} are the 9^{th} round keys. From Equations (3), (4) and (5) 4 bytes of 10^{th} round keys can be recovered. Similarly, if an attacker injects a non zero random fault at 5^{th} byte, then another 32 bits key is obtained. Assuming $CT3$ be a fault free ciphertext and $CT4$ the corresponding faulty ciphertext and expressions of $CT3$ and $CT4$ are

$$CT_3 = \begin{pmatrix} y_0 & y_4 & y_8 & y_{12} \\ y_1 & y_5 & y_9 & y_{13} \\ y_2 & y_6 & y_{10} & y_{14} \\ y_3 & y_7 & y_{11} & y_{15} \end{pmatrix}$$

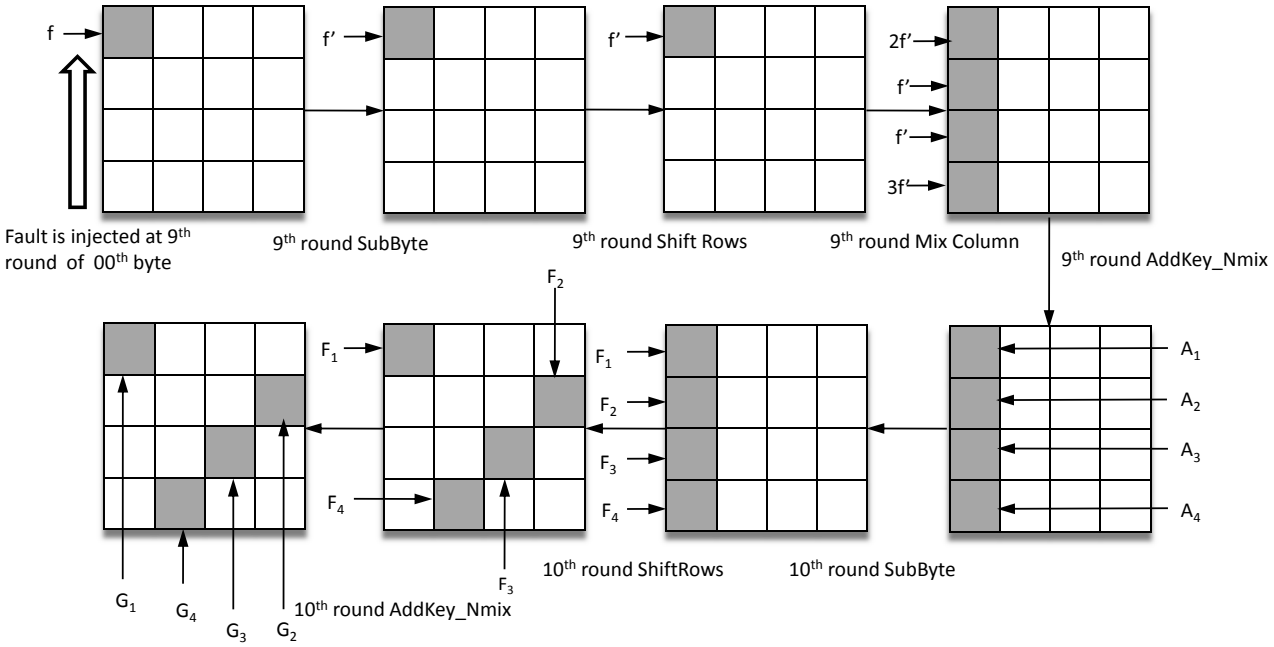

Figure 2: Fault propagation when fault is injected at the 1^{st} byte of 9^{th} round input

$$CT_4 = \begin{pmatrix} y_0 & (y_4 & y_8 & y_{12} \\ & +G_1) & & \\ (y_1 & y_5 & y_9 & y_{13} \\ +G_2) & & & \\ y_2 & y_6 & y_{10} & (y_{14} \\ & & & +G_3) \\ y_3 & y_7 & (y_{11} & y_{15} \\ & & +G_4) & \end{pmatrix}$$

Following equations are formulated using CT_3 and CT_4.

$$\begin{aligned}(INmix(ISB(INmix(y_4, k_{104})), k_{94})) \oplus \\ (INmix(ISB(INmix(y_4 + G_1), k_{104})), k_{94})] = \\ 3[(INmix(ISB(INmix(y_{14}, k_{1014})), k_{96})) \oplus \\ (INmix(ISB(INmix(y_{14} + G_3), k_{1014})), k_{96})]\end{aligned} \tag{6}$$

$$\begin{aligned}(INmix(ISB(INmix(y_{14}, k_{1014})), k_{96})) \oplus \\ (INmix(ISB(INmix(y_{14} + G_3), k_{1014})), k_{96}) = \\ (INmix(ISB(INmix(y_{11}, k_{1011})), k_{97})) \oplus \\ (INmix(ISB(INmix(y_{11} + G_4), k_{1011})), k_{97})\end{aligned} \tag{7}$$

$$\begin{aligned}[(INmix(ISB(INmix(y_1, k_{101})), k_{95})) \oplus \\ (INmix(ISB(INmix(y_1 + G_2), k_{101})), k_{95})] = \\ 2[(INmix(ISB(INmix(y_{14}, k_{1014})), k_{96})) \oplus \\ (INmix(ISB(INmix(y_{14} + G_3), k_{1014})), k_{96})]\end{aligned} \tag{8}$$

From Equations (6), (7) and (8) another 4 bytes of 10^{th} round keys k_{101}, k_{104}, k_{1011}, k_{1014} can be recovered. Similarly if the attacker inject a non zero random fault at 9^{th} byte then another 32 bits key is obtained. From the fault propagation, following equations are constructed.

$$\begin{aligned}[(INmix(ISB(INmix(y_5, k_{105})), k_{99})) \oplus \\ (INmix(ISB(INmix(y_5 + H_2), k_{105})), k_{99})] = \\ 3[(INmix(ISB(INmix(y_8, k_{108})), k_{98})) \oplus \\ (INmix(ISB(INmix(y_8 + H_1), k_{108})), k_{98})]\end{aligned} \tag{9}$$

$$\begin{aligned}(INmix(ISB(INmix(y_{15}, k_{1015})), k_{911})) \oplus \\ (INmix(ISB(INmix(y_{15} + H_4), k_{1015})), k_{911}) = \\ (INmix(ISB(INmix(y_8, k_{108})), k_{98})) \oplus \\ (INmix(ISB(INmix(y_8 + H_1), k_{108})), k_{98})\end{aligned} \tag{10}$$

$$\begin{aligned}[(INmix(ISB(INmix(y_2, k_{102})), k_{910})) \oplus \\ (INmix(ISB(INmix(y_2 + H_3), k_{102})), k_{910})] = \\ 2[(INmix(ISB(INmix(y_8, k_{108})), k_{98})) \oplus \\ (INmix(ISB(INmix(y_8 + H_1), k_{108})), k_{98})]\end{aligned} \tag{11}$$

From Equation (9), (10) and (11) another 4 bytes 10^{th} round keys k_{102}, k_{105}, k_{108}, k_{1015} can be recovered. Similarly if the attacker inject a non zero random fault at 13^{th} byte then another 32 bits key can be obtained.

Following equations are constructed employing fault propagation diagram

$$[(INmix(ISB(INmix(y_6, k_{106})), k_{914})) \oplus$$

$$(INmix(ISB(INmix(y_6 + I_3), k_{106})), k_{914})] = 3[(INmix(ISB(INmix(y_{12}, k_{1012})), k_{912})) \oplus (INmix(ISB(INmix(y_{12} + I_1), k_{1012})), k_{912})] \quad (12)$$

$$(INmix(ISB(INmix(y_{12}, k_{1012})), k_{912})) \oplus (INmix(ISB(INmix(y_{12} + I_1), k_{1012})), k_{912}) = (INmix(ISB(INmix(y_9, k_{109})), k_{913})) \oplus (INmix(ISB(INmix(y_9 + I_2), k_{109})), k_{913}) \quad (13)$$

$$[(INmix(ISB(INmix(y_3, k_{103})), k_{915})) \oplus (INmix(ISB(INmix(y_3 + I_4), k_{103})), k_{915})] = 2[(INmix(ISB(INmix(y_{12}, k_{1012})), k_{912})) \oplus (INmix(ISB(INmix(y_{12} + I_1), k_{1012})), k_{912})] \quad (14)$$

Similarly, another 4 bytes of 10^{th} round keys k_{103}, k_{106}, k_{109}, k_{1012} can be recovered by the attacker employing Equations (12), (13) and (14).

3.1 Working Example

An example is provided in this subsection. Here a fault is injected at 1^{st} byte of 9^{th} round input. Assume $PT1$ is a given plaintext

$$PT_1 = \begin{pmatrix} 2f & cb & c7 & 9e \\ 28 & a0 & 81 & 23 \\ 8e & 9f & bd & 5b \\ 28 & 3e & e4 & 4b \end{pmatrix}$$

and the cipher key K_0 is as follows

$$K_0 = \begin{pmatrix} c7 & bd & d7 & be \\ 9b & d9 & 9b & 9c \\ da & cd & 6c & fa \\ bc & 28 & f8 & 9c \end{pmatrix}$$

The 9^{th} round key is obtained by employing AES key expansion algorithm is as follows

$$K_9 = \begin{pmatrix} af & 35 & 24 & 90 \\ f0 & fa & 45 & 99 \\ a7 & 87 & 34 & 33 \\ d8 & 17 & be & 59 \end{pmatrix}$$

and the 10^{th} round key is as follows

$$K_{10} = \begin{pmatrix} 77 & 42 & 66 & f6 \\ 33 & c9 & 8c & 15 \\ 6c & eb & df & ec \\ b8 & af & 11 & 48 \end{pmatrix}$$

Corresponding fault free ciphertext is as follows

$$CT_1 = \begin{pmatrix} ca & ea & 6f & 6d \\ fe & cb & 3f & 16 \\ 27 & 89 & 26 & 6d \\ 8a & 62 & 2e & d0 \end{pmatrix}$$

The corresponding faulty ciphertext after injecting fault at 1^{st} byte of 9^{th} round is as follows

$$CT_1' = \begin{pmatrix} \mathbf{71} & ea & 6f & 6d \\ fe & cb & 3f & \mathbf{a6} \\ bb & 71 & \mathbf{8e} & 11 \\ 8a & \mathbf{bd} & 2e & d0 \end{pmatrix}$$

Bolded bytes show how the faults have been propagated in the ciphertext.

Let another plaintext be

$$PT_2 = \begin{pmatrix} a8 & f4 & bc & 6c \\ cd & c3 & 76 & aa \\ e7 & 80 & af & d5 \\ 0b & a1 & 9a & e1 \end{pmatrix}$$

The corresponding ciphertext is for the same cipher key is

$$CT_2 = \begin{pmatrix} eb & 8d & 5a & 15 \\ c6 & cd & a4 & ba \\ 27 & 89 & 26 & 6d \\ af & ae & b3 & 1b \end{pmatrix}$$

The corresponding faulty ciphertext after injecting fault at 1^{st} byte of 9^{th} round input is as follows

$$CT_2' = \begin{pmatrix} \mathbf{75} & 8d & 5a & 15 \\ c6 & cd & a4 & \mathbf{2e} \\ 27 & 89 & \mathbf{0a} & 6d \\ af & \mathbf{86} & b3 & 1b \end{pmatrix}$$

First by using equation 3 and a pair of faulty-fault free ciphertext, set of $k_{90}, k_{91}, k_{100}, k_{1013}$ values are obtained. Then by using another faulty-fault free ciphertext pair another set of values of k_{90}, k_{91}, k_{100} and k_{1013} are obtained which are satisfying Equation (3). Intersection of these two sets produces a reduced set of values k_{90}, k_{91}, k_{100} and k_{1013}. A third set of $k_{90}, k_{91}, k_{100}, k_{1013}$ values are obtained from another pair of faulty-fault free ciphertext and reduced key set and then by intersecting more reduce key set is obtained. And finally using 4^{th} pair of ciphertext, a set of $k_{90}, k_{91}, k_{100}, k_{1013}$ values are obtained and set of values obtained in previous step are intersected to obtain correct 10^{th} round 16 bits key.

In this way, employing Equations (4) and (5) and 4 faulty-fault free ciphertext pairs similar analysis is done and finally, four bytes $k_{100}, k_{107}, K_{1010}$ and K_{1013} of 10^{th} round keys are obtained correctly.

To recover the set of $k_{100}, k_{107}, k_{1010}, k_{1013}$ keys, it needs computational complexity of 4×2^{32} i.e. 2^{34}. Computational complexity of 16×2^{32} i.e. 2^{36} is required to obtain 128-bits key.

4 Fault Attack on Eighth Round of Proposed SPN Architecture

In proposed SPN architecture, a non-zero fault has been induced at the input of 8^{th} round. After 8^{th} round Mix-Column step, the fault is distributed into 4 bytes. Again

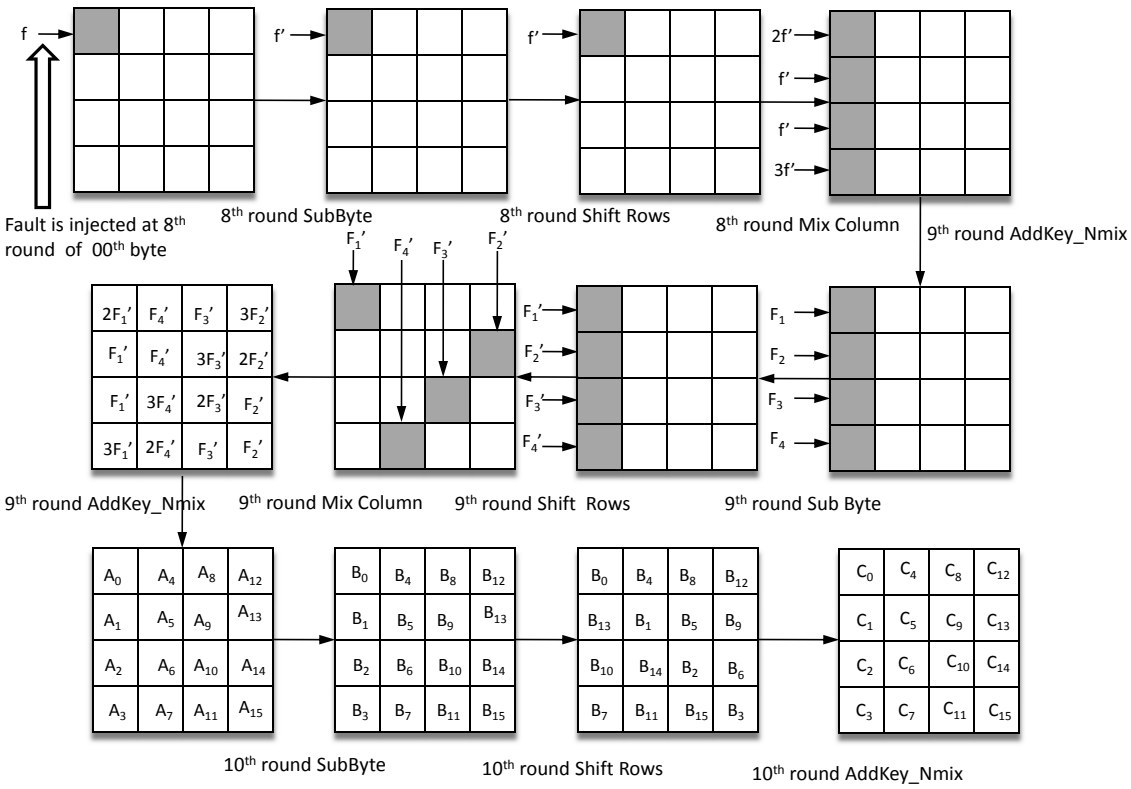

Figure 3: Fault propagation when fault is injected at the input of 8^{th} round

after MixColumn step of 9^{th} round the fault is spreaded throughout all the bytes of state matrix as shown in Fig.3.

Assume $CT1$ is a fault free ciphertext

$$CT_1 = \begin{pmatrix} x_0 & x_4 & x_8 & x_{12} \\ x_1 & x_5 & x_9 & x_{13} \\ x_2 & x_6 & x_{10} & x_{14} \\ x_3 & x_7 & x_{11} & x_{15} \end{pmatrix}$$

If $CT2$ is the corresponding faulty ciphertext, then it can be expressed in following matrix form

$$CT_2 = \begin{pmatrix} x_0 + F_0 & x_4 + F_4 & x_8 + F_8 & x_{12} + F12 \\ x_1 + F_1 & x_5 + F_5 & x_9 + F_9 & x_{13} + F13 \\ x_2 + F_2 & x_6 + F_6 & x_{10} + F_{10} & x_{14} + F14 \\ x_3 + F_3 & x_7 + F_7 & x_{11} + F_{11} & x_{15} + F15 \end{pmatrix}$$

Let the associated round keys are K_8, K_9 and K_{10} in 8^{th}, 9^{th} and 10^{th} rounds respectively

$$K_8 = \begin{pmatrix} k_{80} & k_{84} & k_{88} & k_{812} \\ k_{81} & k_{85} & k_{89} & k_{813} \\ k_{82} & k_{86} & k_{810} & k_{814} \\ k_{83} & k_{87} & k_{811} & k_{815} \end{pmatrix}$$

$$K_9 = \begin{pmatrix} k_{90} & k_{94} & k_{98} & k_{912} \\ k_{91} & k_{95} & k_{99} & k_{913} \\ k_{92} & k_{96} & k_{910} & k_{914} \\ k_{93} & k_{97} & k_{911} & k_{915} \end{pmatrix}$$

$$K_{10} = \begin{pmatrix} k_{100} & k_{104} & k_{108} & k_{1012} \\ k_{101} & k_{105} & k_{109} & k_{1013} \\ k_{102} & k_{106} & k_{1010} & k_{1014} \\ k_{103} & k_{107} & k_{1011} & k_{1015} \end{pmatrix}$$

From the fault pattern, following equations are constructed

$$[INmix(a, k_{80}) \oplus (INmix(b, k_{80}))] = 2[INmix(c, k_{81}) \oplus INmix(d, k_{81})] \quad (15)$$

Where,
$a = (ISB((INmix(ISB(INmix(x_0, k_{100})), k_{90}))),$
$b = (ISB((INmix(ISB(INmix(x_0 + F_0), k_{100})), k_{90}))$
$c = (ISB((INmix(ISB(INmix(x_{13}, k_{1013})), k_{91})$ and
$d = ISB((INmix(ISB(INmix(x_{13} + F_{13}), k_{1013})), k_{91}))$

$$(INmix(e, k_{81}) \oplus (INmix(f, k_{81}) = (INmix(g, K_{82}) \oplus (INmix(h, K_{82}) \quad (16)$$

Where,
$e = ISB(INmix(ISB(INmix(x_{13}, k_{1013})), k_{91})),$
$f = ISB(INmix(ISB(INmix(x_{13} + F_{13}), k_{1013})), k_{91}),$
$g = ISB(INmix(ISB(INmix(x_{10}, k_{1010})), k_{92}))$ and
$h = ISB(INmix(ISB(INmix(x_{10} + F_{10}), k_{1010})), k_{92}).$

$$[(INmix(i, k_{83}) \oplus (INmix(j, k_{83})] =$$

$$3[(INmix(k, K_{81}) \oplus (INmix(l, K_{81})] \tag{17}$$

Where,
$i = ISB(INmix(ISB(INmix(x_7, k_{107})), k_{93}))$,
$j = ISB(INmix(ISB(INmix(x_7 + F_7), k_{107})), k_{93})$,
$k = ISB(INmix(ISB(INmix(x_{13}, k_{1013})), k_{91}))$ and
$l = ISB(INmix(ISB(INmix(x_{13} + F_{13}), k_{1013})), k_{91})$.
Five pairs of fault free-faulty ciphertexts are needed to recover 32 bits of 10^{th} round key. To recover 32 bits key, computational complexity of 5×2^{48} i.e. 2^{51} is required. From the fault distribution, similarly another set of 9 equations can be constructed and from these equations, rest of the keys can be recovered. To recover 128 bits 10^{th} round keys, total 20 faulty-fault free ciphertext pairs are necessary and computational complexity is 20×2^{48} i.e. 2^{53}.

5 Comparison with Existing Works

In this section, a comparison is provided in terms of fault model, fault location, number of faulty encryptions and computational complexity, of our work and with the works reported in [7, 15, 16, 19, 20]. Existing related works either based on byte level or bit level fault model. Our work is based on byte level fault model. From Table 1, it is observed that in proposed architecture, 16 and 20 faulty-fault free ciphertext pairs are necessary to mount fault attack by injecting fault at the input of 9^{th} and 8^{th} round respectively. Also fault attack complexity in proposed scheme is relatively higher than that of AES [19]. Fault attack on AES [19] requires minimum 2 faulty-fault free ciphertext pairs with complexity 2^{32}. Fault attack on MDS-AES [12] needs 2 faulty cipher text pairs with brute-force search of complexity 2^{16}. Whereas to mount fault attack on proposed SPN-type architecture minimum 16 faulty-fault free ciphertext pairs are necessary with complexity 2^{36}.

6 Conclusion

In this paper, a new SPN-type architecture has been proposed to improve the security of block cipher against fault attack. Here, instead of linear round key mixing function, first time effect of nonlinear round key mixing function is used and analysed, to protect fault attack. Proposed architecture also provides better security against fault attack compared to AES. To derive 128 bits 10^{th} round key it needs computational complexity of 2^{36} and 16 faulty-fault free ciphertext pairs, when fault is injected at input of 9^{th} round. When a fault is introduced at input of 8^{th} round then it needs computational complexity of 2^{53} and 20 faulty-fault free ciphertext pairs, to recover 128 bits of 10^{th} round key.

References

[1] M. Alam, S. Ghosh, D. R. Choudhury, and I. Sengupta, "First-order DPA vulnerability of Rijndael: Security and area-delay optimization trade-off," *International Journal of Network Security*, vol. 15, no. 3, pp. 219-230, 2013.

[2] S. Ali, X. Guo, R. Karri, and D. Mukhopadhyay, "Fault attacks on AES and their countermeasures," in *Secure System Design and Trustable Computing*, pp. 163-208, 2016.

[3] J. Bhaumik, D. R. Chowdhury, "HDNM8: A round-8 high diffusion block cipher with nonlinear mixing function," *Springer Proceedings in Mathematics and Statistics*, vol. 91, pp. 41-55, 2013.

[4] J. Bhaumik, D. R. Chowdhury, "NMIX: An ideal candidate for key mixing," in *Proceedings of the International Conference on Security and Cryptography*, pp. 285-288, 2009.

[5] J. Bhaumik, D. R. Chowdhury, "An integrated ECC-MAC based on RS code," *Transactions on Computational Science*, LNCS 5430, pp. 117-135, 2009.

[6] E. Biham, A. Shamir, "Differential fault analysis of secret key cryptosystems," in *Annual International Cryptology Conference (CRYPTO'97)*, LNCS 1294, pp. 513-525, 1997.

[7] J. Blomer, J. P. Seifert, "Fault based cryptanalysis of the advanced encryption standard (AES)," in *International Conference on Financial Cryptography (FC'03)*, pp. 162-181, 2003.

[8] D. Boneh, R. A. DeMillo, R. J. Lipton, "On the importance of eliminating errors in cryptographic computations," *Journal of Cryptology*, vol. 12, pp. 241-246, 2001.

[9] D. Boneh, R. A. DeMillo, R. J. Lipton, "On the importance of checking cryptographic protocols for faults," in *International Conference on the Theory and Applications of Cryptographic Techniques (EUROCRYPT'97)*, LNCS 1233, pp. 37-51, 1997.

[10] J. Breier, W. He, "Multiple fault attack on PRESENT with a hardware trojan implementation in FPGA," in *International Workshop on Secure Internet of Things (SIoT'15)*, pp. 58-64, 2015.

[11] J. Daemen, V. Rijmen, *The Design of Rijndael*, Springer, Heidelberg 2002.

[12] S. Das, J. Bhaumik, "A fault based attack on MDS-AES," *International Journal of Network Security*, vol. 16, no. 3, pp. 193-198, 2014.

[13] Y. Dou, J. Weng, Y. Wei, and C. Ma, "Improved fault attack against Eta pairing," *International Journal of Network Security*, vol. 16, no. 1, pp. 71-77, 2014.

[14] Y. Dou, J. Weng, Y. Wei, and C. Ma, "Fault attack against Miller's algorithm for even embedding degree," *International Journal of Network Security*, vol. 16, no. 3, pp. 199-207, 2014.

[15] P. Dusart, G. Letourneux and O. Vivolo, *Differential Fault Analysis on A.E.S.*, 2002. (`http://eprint.iacr.org/2003/010`)

Table 1: Comparison of Fault attack on AES with our proposed SPN type architecture accomplishing properties of the encryption function

Reference	Fault Model	Algorithm	Fault Location	No. of Faulty Encryptions	Complexity
[7]	Force 1 bit to 0	AES	Chosen	128	
[7]	Implementation Dependent	AES	Chosen	256	
[16]	Switch 1 bit	AES	Any bit of chosen bytes	50	
[16]	Disturb 1 byte	AES	Anywhere among 4 bytes	250	
[15]	Disturb 1 byte	AES	Anywhere between last two MixColumn	40	
[20]	Disturb 1 byte	AES	Anywhere between 7^{th} round and 7^{th} round MixColumn	2	
[19]	Disturb 1 byte	AES	Anywhere between 7^{th} round and 7^{th} round MixColumn	2	2^{32}
[12]	Disturb 1 byte	MDS-AES	Input of 9^{th} round	2	2^{16}
This paper	Disturb 1 byte	Proposed SPN	Input of 9^{th} round	16	2^{36}
This paper	Disturb 1 byte	Proposed SPN	Input of 8^{th} round	20	2^{53}

[16] C. Giraud, "DFA on AES," *Cryptology ePrint Archive*, Report 2003/008.

[17] B. Lac, A. Canteaut, J. J. A. Fournier, *DFA on LS-designs with a Practical Implementation on SCREAM (extended version)*, Dec. 25, 2017. (`http://eprint.iacr.org/2017/076.pdf`)

[18] A. Mirsaid, T. Gulom, "The encryption algorithm AES-RFWKPES32-4," *International Journal of Electronics and Information Engineering*, vol. 5, no. 1, pp. 20-29, 2016.

[19] D. Mukhopadhyay, "An improved fault based attack of the advanced encryption standard," in *Advances in Cryptography (AFRICACRYPT'09)*, LNCS 5580, pp. 421-434, 2009.

[20] G. Piret, J. J. Quisquater, "A differential fault attack technique against SPN structures, with application to the AES and Khazad," in *International Workshop on Cryptographic Hardware and Embedded Systems (CHES'03)*, pp. 77-88, 2003.

[21] S. Skorobogatov, R. Anderson, "Optical fault induction attacks," in *International Workshop on Cryptographic Hardware and Embedded Systems (CHES'02)*, pp. 2-12, 2003.

[22] J. Takahashi, T. Fukunaga, *Differential Fault Analysis on the AES Key Schedule*, 2007. (`http://eprint.iacr.org/2007/480`)

[23] J. Takahashi, T. Fukunaga, K. Yamakoshi, "DFA mechanism on the AES schedule," in *Proceedings of 4th International Workshop on Fault Detection and Tolerance in Cryptography*, pp. 27-41, 2007.

[24] M. Tunstall, D. Mukhopadhyay, S. Ali, "Differential fault analysis of the advanced encryption standard using a single fault," *Information Security Theory and Practice*, vol. 6633, pp. 224-233, 2011.

Biography

Gitika Maity is currently working as an Asistant Professor in the Department of Computer Science and Engineering, Haldia Institute of Technology, Haldia, India. She received her B. Tech. and M. Tech. Degrees from West Bengal University of Technoloy, India. Her research interests include Cryptography, Cellular Automata and Digital VLSI Design.

Jaydeb Bhaumik is currently working as a Professor and Head in the Department of Electronics and Communication Engineering, Haldia Institute of Technology, Haldia, India. He received his B. Tech. and M. Tech. degrees in Radio Physics and Electronics from University of Calcutta in 1996 and 1999 respectively and PhD degree from Indian Institute of Technology Kharagpur in 2010. His research interests include Cryptography, Cellular Automata, Error Correcting Codes and Digital VLSI Design. He is a member of IEEE and life member of Cryptology Research Society of India since 2008. He has published more than 40 research papers in international journals and conferences.

Anjan Kumar Kundu received his B. Tech.and M. Tech. degree, in Radio Physics and Electronics from the University of Calcutta, India. He has more than ten years of experience in the field of teaching and research. He joined the Radio Physics and Electronics in 2005 as Lecturer and serving the department till date. His research interest includes Microwave Tomography, Cryptography and RFID, Electromagnetics and Microwave engineering. He has published several papers in national and international journals and conferences.

Provably Secure and Repeatable Authenticated Privacy-Protection Scheme Using Chaotic Maps with Distributed Architecture

Hongfeng Zhu, Junlin Liu

(Corresponding author: Hongfeng Zhu)

Software College, Shenyang Normal University
No.253, HuangHe Bei Street, HuangGu District, Shenyang, P.C 110034 China
(Email:zhuhongfeng1978@163.com; 272297257@qq.com)

(Received Nov. 5, 2016; revised May 17, 2017 and accepted July 13, 2017)

Abstract

Nowadays, the distributed password-authenticated key agreement schemes become more and more popular. Compare with the three traditional architectures (client/server, two clients/server and multi-server), the distributed architecture can solve problems of single-point of security, single-point of efficiency and single-point of failure. Moreover, it has the characteristics of scalability, flexibility and fairness. In the paper, we proposed a new Provably Secure and Distributed Privacy-Protection scheme using chaotic maps. The proposed scheme firstly achieves mutual authenticated among three nodes in three rounds with privacy protection, and at the same time, the unregistered server can store a temporary authenticator for a while for improving the efficiency. Security of the scheme is based on chaotic maps hard problems and a secure one way hash function. Compared with the related literatures recently, our proposed scheme can not only own high efficiency and unique functionality, but is also robust to various attacks and achieves perfect forward secrecy. Finally, we give the security proof and the efficiency analysis of our proposed scheme.

Keywords: Chaotic Maps Keywords; Distributed Architecture; Key Agreement; Privacy-protection

1 Introduction

Nowadays, more and more people want to enjoy surfing on Internet and meanwhile care about their privacy. The most popular technology is authenticated key exchange (AKE) [5, 12, 13] which can establish an authenticated and confidential communication channel. Many papers adopt multi-server architecture (MSA) [7, 14] to reduce the numbers of users' registration, and the literature [14] can achieve Privacy-Protection and without using symmetric cryptography which can lower the calculated amount. For seeking universal computing environment, Zhu [15] proposed an AKE protocol in different realm, which can make two-party in two-realm negotiate a session key in the standard model. Naturally, the group key agreement scheme with privacy preserving can be proposed in [10, 11, 16]. But the multi-server architecture makes the registration center become the focus of hacker. Furthermore, single-point of efficiency and single-point of failure trouble the registration center all the time.

An excellent architecture can make some hard problems become better easily. For example, distributed architecture can solve centralized architecture problems. Zhu [9] firstly proposed a new distributed architecture which called Multiple Servers to Server Architecture (MSTSA). The paper [9] proposed the first provably secure and flexible password-authenticated key agreement scheme based on chaotic maps [1, 3] with MSTSA in random oracle model [2]. Then, Zhu gives another password-authenticated key agreement scheme [13] with MSTSA which security is proved in standard model. But abovementioned two distributed schemes using chaotic maps have two main problems: without privacy protection and have many communicated rounds [4]. Therefore, the paper proposes a new distributed scheme to solve the two main problems. We adopt chaotic maps because it has numerous advantages, such as extremely sensitive to initial parameters, unpredictability, boundeness, etc. Meanwhile, chaotic sequence generated by chaotic system has the properties of non-periodicity and pseudo-randomness.

The main contributions are shown as below: (1) The paper presents a new password authenticated key exchange scheme with privacy protection towards Multiple Servers to Server Architecture. (2) The proposed scheme achieves mutual authenticated among three nodes in three rounds with privacy protection. (3) The scheme can make the unregistered server store a temporary authenticator for a while for avoiding the registered server involved over and over again. (4) The proposed scheme is mainly based

on chaotic maps without using modular exponentiation and scalar multiplication on an elliptic curve. In Security aspect, the protocol can resist all common attacks, such as impersonation attacks, man-in-the-middle attacks, etc. (5) About functionality, the protocol also has achieved some well-known properties, such as perfect forward secrecy and execution efficiency.

The rest of the paper is organized as follows: Some preliminaries are given in Section 2. Next, a distributed privacy-protection scheme is described in Section 3. Then, the security analysis and efficiency analysis are given in Section 4 and Section 5. This paper is finally concluded in Section 6.

2 Preliminaries

2.1 Multi-server Architecture

In the multi-server environment [7], each user must perform authentication procedure to login the server for a transaction. If the user is in a single authentication architecture, then the user must register at various servers and memorize the corresponding identifications and passwords, which could not be convenient for a user. In order to make the registration to various servers easier for users, each user must register with the registration center to obtain a secure account. Then the user uses the secure account to perform the login and authentication procedures with various servers.

2.2 Multiple Servers to Server Architecture

In the proposed multiple servers to server communication architecture, the registration center is not fixed. In other words, any server can work as a registration center. However in multi-server authentication architecture, the single registration center will face to single-point of security, single-point of efficiency and single-point of failure problems. The proposed architecture can solve the problems under multi-server environment with only one registration center architecture, that means "once security register for all registration" [9].

2.3 Chebyshev Chaotic Maps

Zhang [8] proved that semi-group property holds for Chebyshev polynomials defined on interval $(-\infty, +\infty)$. The enhanced Chebyshev polynomials are used in the proposed protocol:

$$T_n(x) = (2xT_{n-1}(x) - T_{n-2}(x))(\mathrm{mod}N),$$

where $n \geq 2$,$x \in (-\infty, +\infty)$,and N is a large prime number. Obviously,

$$T_{rs}(x) = T_r(T_s(x)) = T_s(T_r(x))$$

Definition 1. *(Enhanced Chebyshev polynomials) The enhanced Chebyshev maps of degree $n(n \in N)$ are defined as: $T_n(x) = (2xT_{n-1}(x) - T_{n-2}(x))(\mathrm{mod}p)$, where $n \geq 2$,$x \in (-\infty, +\infty)$, and p is a large prime number. Obviously, $T_{rs}(x) = T_r(T_s(x)) = T_s(T_r(x))$.*

Definition 2. *(DLP, Discrete Logarithm Problem) Given an integer a,find the integer r, such that $T_r(x) = a$.*

Definition 3. *(CDH, Computational Diffie-Hellman Problem) Given an integer x, and the values of $T_r(x), T_s(x)$, what is the value of $T_{rs}(x)$?*

It is widely believed that there is no polynomial time algorithm to solve DLP, CDH with a non-negligible probability.

2.4 Threat Model

The threat model should be adopted the widely accepted security assumptions about password based authentication schemes [2].

1) The $user_i$ holds the uniformly distributed low-entropy password from the small dictionary. The server keeps the private key. At the time of registration, the server sends the personalized security parameters to the $user_i$ by secure channel and the $user_i$ should keep the personalized security parameters safe.

2) An adversary and a $user_i$ interact by executing oracle queries that enables an adversary to perform various attacks on authentication protocols.

3) The communication channel is controlled by the adversary who has the capacity to intercept, modify, delete, resend and reroute the eavesdropped messages.

In the password authenticated protocol Π, each participant is either a user $u_i \in U$ or a trusted server S interact number of times. Only polynomial number of queries occurs between adversary and the participant interaction. This enables an adversary to simulate a real attack on the authentication protocol. The possible oracle queries are as follows:

Execute (Π_U^i, Π_S^j)**:** This query models passive attacks against the protocol which is used to simulate the eavesdropping honest execution of the protocol. It prompts an execution of the protocol between the user's instances Π_U^i and server's instances Π_S^j that outputs the exchanged messages during honest protocol execution to A.

Send (Π_U^i, m)**:** This query sends a message m to an instance Π_U^i, enabling adversary A for active attacks against the protocol. On receiving m, the instance Π_U^i continues according to the protocol specification. The message output by Π_U^i, if any, is returned to A.

Reveal (Π_U^i)**:** This query captures the notion of known key security. The instance Π_U^i, upon receiving the query and if it has accepted, provides the session key, back to A.

Corrupt (Π_U^i, m)**:** These queries together capture the notion of two-factor security. The former returns the password of U_i while the latter returns the information stored in the smart card of U_i.

Test (Π_U^i)**:** This query is used for determining whether the protocol achieves authenticated key exchange or not. If Π_U^i has accepted, then a random bit $b \in \{0,1\}$ chosen by the oracle, A is given either the real session key if $b = 1$, otherwise, a random key drawn from the session key space.

We say that an instance Π_U^i is said to be open if a query Reveal (Π_U^i) has been made by adversary, and unopened if it is not opened. We say that an instance Π_U^i has accepted if it goes into an accept mode after receiving the last expected protocol message.

Definition 4. *Two instances Π_U^i and Π_S^i are said to be partnered if the following conditions hold: (1) Both Π_U^i and Π_S^i accept; (2) Both Π_U^i and Π_S^i share the same session identifications(sid); (3) The partner identification for Π_U^i and Π_S^i and vice-versa.*

Definition 5. *We say an instance Π_U^i is considered fresh if the following conditions are met: (1) It has accepted; (2) Both Π_U^i and its partner Π_S^i are unopened; (3) They are both instances of honest clients.*

Definition 6. *Consider an execution of the authentication protocol Π by an adversary A, in which the latter is given access to the Execute, Send, and Test oracles and asks at most single Test query to a fresh instance of an honest client. Let b' be his output, if $b' = b$, where b is the hidden bit selected by the Test oracle. Let D be user's password dictionary with size $|D|$. Then, the advantage of A in violating the semantic security of the protocol Π is defined more precisely as follows:*

$$Adv_{\Pi,D}(A) = [2\Pr[b' = b] - 1]$$

The password authentication protocol is semantically secure if the advantage $Adv_{\Pi,D}(A)$ is only negligibly larger than $O(q_s)/|D|$, where q_s is the number of active sessions.

3 The Proposed Privacy Protection Scheme with Multiple Servers to Server Architecture

3.1 User Registration Phase

The concrete notations used hereafter are: ID_{S_i} means identity of the ith server; ID_A means the identity of Alice; a, a_1, r_a, r_i are all nonces; $(x, T_{k_i}(x))$, the public key based on Chebyshev chaotic maps of the ith server; k_i, the secret key based on Chebyshev chaotic maps of the ith server; H, A secure one-way hash function. H: $\{0,1\}^* \rightarrow \{0,1\}^l$ for a constant l; $\|$ means concatenation operation.

Figure 1 illustrates the user registration phase.

Step 1. When a user wants to be a new legal user, she chooses her identity ID_A, a random number r_a, and computes $H(r_a||PW)$. Then Alice submits $ID_A, H(r_a||PW)$ to the **RC** via a secure channel.

Step 2. Upon receiving $ID_A, H(r_a||PW)$ from Alice, the **RC** computes $B = H(ID_A||k_i) \oplus H(r_a||PW)$, where k_i is the secret key of S_i. Then Alice stores $\{ID_A, r_a, B\}$ in a secure way.

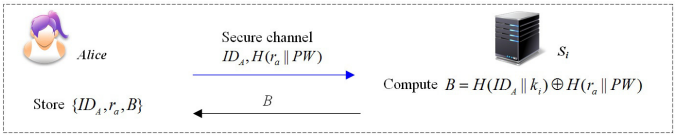

Figure 1: a premium user registration phase

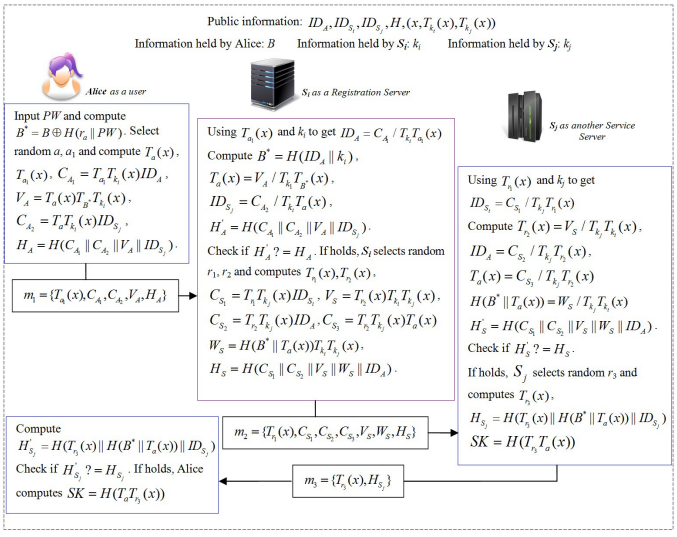

Figure 2: Authenticated key agreement phase for MSTSA with privacy protection

3.2 Authenticated Key Agreement Phase for MSTSA with Privacy Protection

Figure 2 illustrates the process of authenticated key agreement phase.

Step 1. If Alice wishes to consult some personal issues establish with S_j in a secure way, she will input password and compute $B_A^* = B_A \oplus H(r_a||PW)$, and then choose two random integer numbers a, a_1 and compute $T_a(x)$, $T_{a_1}(x)$, $C_{A_1} = T_{a_1} T_{k_i}(x) ID_A$, $V_A = T_a(x) T_{B^*} T_{k_i}(x)$, $C_{A_2} = T_a T_{k_i}(x) ID_{S_j}$, $H_A =$

$H(C_{A_1}||C_{A_2}||V_A||ID_{S_j})$. After that, Alice sends $m_1 = \{T_{a_1}(x), C_{A_1}, C_{A_2}, V_A, H_A\}$ to S_i which she has registered.

Step 2. After receiving the message $m_1 = \{T_{a_1}(x)$, C_{A_1}, C_{A_2}, V_A, $H_A\}$ from Alice, S_i will use $T_{a_1}(x)$ and the secret key k_i to get $ID_A = C_{A_1}/T_{k_i}T_{a_1}(x)$. Then S_i computes $B^* = H(ID_A \,||\, k_i)$, $T_a(x) = V_A/T_{k_1}T_{B^*}(x)$, $ID_{S_j} = C_{A_2}/T_{k_i}T_a(x)$, $H'_A = H(C_{A_1} \,||\, C_{A_2} \,||\, V_A \,||\, ID_{S_j})$. Check if H'_A is equal to H_A. If holds, that means Alice is the real and legal user. Next, S_i selects random r_1, r_2 and computes $T_{r_1}(x), T_{r_2}(x)$, $C_{S_1} = T_{r_1}T_{k_j}(x)ID_{S_i}$, $V_S = T_{r_2}(x)T_{k_i}T_{k_j}(x)$, $C_{S_2} = T_{r_2}T_{k_j}(x)ID_A$, $C_{S_3} = T_{r_2}T_{k_j}(x)T_a(x)$, $W_S = H(B^* \,||\, T_a(x))T_{k_i}T_{k_j}(x)$, $H_S = H(C_{S_1} \,||\, C_{S_2} \,||\, V_S \,||\, W_S \,||\, ID_A)$. After that, S_i sends $m_2 = \{T_{r_1}(x)$, C_{S_1}, C_{S_2}, C_{S_3}, V_S, W_S, $H_S\}$ to server S_j which Alice wants to get service.

Step 3. After receiving the message $m_2 = \{T_{r_1}(x)$, C_{S_1}, C_{S_2}, C_{S_3}, V_S, W_S, $H_S\}$ from S_i, S_j uses $T_{r_1}(x)$ and the secret key k_j to get $ID_{S_i} = C_{S_1}/T_{k_j}T_{r_1}(x)$. Then S_j computes $T_{r_2}(x) = V_S/T_{k_j}T_{k_i}(x)$, $ID_A = C_{S_2}/T_{k_j}T_{r_2}(x)$, $T_a(x) = C_{S_3}/T_{k_j}T_{r_2}(x)$, $H(B^* \,||\, T_a(x)) = W_S/T_{k_j}T_{k_i}(x)$, $H'_S = H(C_{S_1} \,||\, C_{S_2} \,||\, V_S \,||\, W_S \,||\, ID_A)$. Check if H'_S is equal to H_S. If holds, that means S_i is the real and legal server. Next, S_j selects random r_3 and computes $T_{r_3}(x)$, $H_{S_j} = H(T_{r_3}(x)||H(B^*||T_a(x))||ID_{S_j})$, $SK = H(T_{r_3}T_a(x))$. Finally, S_j sends $m_3 = \{T_{r_3}(x), H_{S_j}\}$ to Alice.

Step 4. After receiving the message $m_3 = \{T_{r_3}(x), H_{S_j}\}$, Alice computes $H'_{S_j} = H(T_{r_3}(x) \,||\, H(B^* \,||\, T_a(x)) \,||\, ID_{S_j})$ using local information. Then, Alice checks if H'_{S_j} is equal to H_{S_j}. If holds, that means S_i has helped Alice to authenticate S_j, because S_j owns the authenticator $H(B^*||T_a(x))$ which only Alice and S_i can compute B^*. Finally, Alice computes the session key $SK = H(T_aT_{r_3}(x))$.

If any authenticated process does not pass, the protocol will be terminated immediately.

Remark 1. *$H(B^*||T_a(x))$ is the temporary authenticator which can be used for a certain time. So, Alice and S_j can use $H(B^*||T_a(x))$ to construct some other session keys, such as $H(H(B^*||T_a(x)))$, $H(H(B^*||T_a(x))||T_{r_3}(x))$ and so on, without S_i involved.*

3.3 Password Changing Phase

Figure **??** illustrates the password changing phase.

Step 1. When a user wants to change her password, she chooses a new password, two random numbers r'_a, a, and computes $B^* = B \oplus H(r_a||PW)$, $T_a(x)$, $K_{A-S_i} = T_aT_k(x)$, $H_A = H(B^*||ID_{S_i}||T_a(x)||C_1||C_2)$, $C_1 = ID_A \times K_{A-S_i}$ and $C_2 = H(r'_a||PW') \times K_{A-S_i}$. Then Alice sends $m_1 = \{T_a(x), C_1, C_2, H_A\}$.

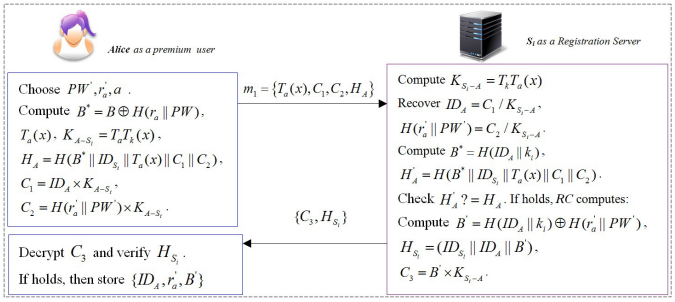

Figure 3: Password changing phase

Step 2. Upon receiving $m_1 = \{T_a(x)$, C_1, C_2, $H_A\}$ from Alice, S_i computes $K_{S_i-A} = T_kT_a(x)$ and recovers $ID_A = C_1/K_{S_i-A}$, $H(r'_a||PW') = C_2/K_{S_i-A}$. Next S_i computes $B^* = H(ID_A||k_i)$ and $H'_A = H(B^*||ID_{S_i}||T_a(x)||C_1||C_2)$. Then S_i checks $H'_A = H_A$ or not. If holds, S_i computes $B' = H(ID_A||k_i) \oplus H(r'_a||PW')$, $H_{S_i} = (ID_{S_i}||ID_A||B')$ and $C_3 = B' \times K_{S_i-A}$, where k_i is the secret key of S_i. Finally S_i sends $\{C_3, H_{S_i}\}$ to Alice.

Step 3. Upon receiving $\{C_3$, $H_{S_i}\}$, Alice uses K_{A-S_i} to decrypt C_3 to get B'. Then Alice computes locally $H'_{S_i} = (ID_{S_i}||ID_A||B')$ to compare with H_{S_i}.If they are equal, Alice stores $\{ID_A, r'_a, B'\}$ in a secure way.

4 Security Analysis

4.1 The Provable Security of the Proposed Scheme [2]

First of all, we transform the process of our proposed scheme with privacy protection in MSTSA to the following two simulation Algorithms.

Theorem 1. *Let D be a uniformly distributed dictionary of possible passwords with size D, Let P be the improved authentication protocol described in Algorithm 1 and 2. Let A be an adversary against the semantic security within a time bound t. Suppose that CDH assumption holds, then,*

$$Adv_{\Pi,D}(A) \leq \frac{2q_h^2}{2^{l+1}} + \frac{(q_s+q_e)^2}{p^2} + 2q_h Adv_G^{cdh}(A) + \frac{2q_h}{p} + \frac{q_s}{D}$$

where $Adv_G^{cdh}(A)$ is the success probability of A of solving the chaotic maps-based computational Diffie-Hellman problem. q_s is the number of Send queries, q_e is the number of Execute queries and q_h is the number of random oracle queries.

Proof. This proof defines a sequence of hybrid games, starting at the real attack and ending up in game where the adversary has no advantage. For each game $G_i (0 \leq i \leq 4)$, we define an event $Succ_i$ corresponding to the event in which the adversary correctly guesses the bit b in the test-query.

Algorithm 1 Simulation of send query

On a query $send(\Pi_U^i, start)$, assume that U_i is in correct state, then we proceed as follows:
Choose two numbers $a, a_1 \in_R Z_p^*$, compute $\{T_{a_1}(x), C_A, V_A, H_A\}$. This query returns $\{T_{a_1}(x), C_A, V_A, H_A\}$ as answer.
On a query $send(S_i, \{T_{a_1}(x), C_A, V_A, H_A\})$, assume that Si is in correct state, we continue as follows:
Compute $ID_A = C_{A_1}/T_{k_i}T_{a_1}(x)$, $B^* = H(ID_A||k_i)$, $T_a(x) = V_A/T_{k_1}T_{B^*}(x)$, $ID_{S_j} = C_{A_2}/T_{k_i}T_a(x)$ and $H'_A = H(C_{A_1}||C_{A_2}||V_A||ID_{S_j})$.
if $H'_A \neq H_A$ **then**
  Reject the message.
  else S_i choose two numbers $r_1, r_2 \in_R Z_p^*$ and computes $\{T_{r_1}(x), C_{S_1}, C_{S_2}, V_S, W_S, H_A\}$. This query returns $\{T_{r_1}(x), C_{S_1}, C_{S_2}, V_S, W_S, H_A\}$ as answer.
  On a query $send(S_j, \{T_{r_1}(x), C_{S_1}, C_{S_2}, V_S, W_S, H_A\})$, assume that Sj is in correct state, we continue as follows:
  Compute $T_{r_2}(x) = V_S/T_{k_j}T_{k_i}(x)$, $ID_A = C_{S_2}/T_{k_j}T_{r_2}(x)$, $H(B^*||T_a(x)) = W_S/T_{k_j}T_{k_i}(x)$ and $H'_S = H(C_{S_1}||C_{S_2}||V_S||W_S||ID_A)$
  if $H'_S \neq H_S$ **then**
    Reject the message.
    else S_j chooses a number $r_3 \in_R Z_p^*$ and computes $H_{S_j} = H(T_{r_3}(x)||H(B^*||T_a(x))||ID_{S_j})$ and $SK = T_{r_3}T_a(x)$. The query $\{T_{r_3}(x), H_{S_j}\}$ returns as answer.
  end if
end if
On a query $send\{T_{r_3}(x), H_{S_j}\}$, assume that U_i is in correct state, then we proceed as follows:
U_i computes $H'_{S_j} = H(T_{r_3}(x)||H(B^*||T_a(x))||ID_{S_j})$.
if $H'_{S_j} \neq H_{S_j}$ **then**
  Reject the message.
  else compute $SK = T_aT_{r_3}(x)$ and the user U_i instance accepts.
end if

Algorithm 2 Simulation of Execute query

On a query Reveal (Π_U^i), we proceed as follows:
if The instance Π_U^i is accepted **then**
  This query answered the session key.
end if

Game G_0**:** This game correspond to the real attack in the random oracle model. In this game, all the instances of U_A and U_B are modeled as the real execution in the random oracle. By definition of event $Succ_i$ in which the adversary correctly guesses the bit b involved in the Test-query, we have

$$Adv_{\Pi,D}(A) = 2|\Pr[Succ_0] - \frac{1}{2}| \quad (1)$$

Game G_1**:** This game is identical to the game G_0, except that we simulate the hash oracles h by maintaining the hash lists $List_h$ with entries of the form (Inp, Out).On hash query for which there exists a record (Inp, Out) in the hash list, return Out. Otherwise, randomly choose $Out \in \{0,1\}$, send it to A and store the new tuple (Inp, Out) into the hash list. The Execute, Reveal, Send, Corrupt, and Test oracles are also simulated as in the real attack where the simulation of the different polynomial number of queries asked by A. From the viewpoint of A, we identify that the game is perfectly indistinguishable from the real attack. Thus, we have

$$\Pr[Succ_1] = \Pr[Succ_0] \quad (2)$$

Game G_2**:** In this game, the simulation of all the oracles is identical to game G_1 except that the game is terminated if the collision occurs in the simulation of the partial transcripts $\{T_{a_1}(x), C_A, V_A, H_A\}$, $\{T_{r_1}(x), C_{S_1}, C_{S_2}, V_S, W_S, H_S\}$ or $\{T_{r_3}(x), H_{S_j}\}$ and on hash values. According to the birthday paradox, the probability of collisions of the simulation of hash oracles is at most $q_h^2/2^{l+1}$. Similarly, the probability of collisions in the transcripts simulations is at most $\frac{(q_h+q_e)^2}{2p^2}$. Since a, a_1, r_i were selected uniformly at random. Thus, we have

$$\Pr[Succ_2] - \Pr[Succ_1] = \frac{q_h^2}{2^{l+1}} + \frac{(q_h+q_e)^2}{2p^2} \quad (3)$$

Game G_3**:** In this game, the session key is guessed without asking the corresponding oracle h so that it become independent of password and ephemeral keys a, r_3 which are protected by the chaotic maps-based computational Diffie-Hellman problem. We change the way with earlier game unless A queries h on the common value $SK = H(T_aT_{r3}(x))$. Thus, $Adv_G^{cdh}(A) \geq \frac{1}{q_h}|\Pr[Succ_3] - \Pr[Succ_2]| - \frac{1}{p}$,that is, the difference between the game G_3 and the game G_2 is as follows:

$$Pr[Succ_3] - \Pr[Succ_2]| \leq q_h Adv_G^{cdh}(A) + \frac{q_h}{p} \quad (4)$$

Game G_4**:** This game is similar to the game G_3 except that in Test query, the game is aborted if A asks a hash function query with $SK = H(T_aT_{r3}(x))$. A gets the session key SK by hash function query with probability at most $\frac{q_h^2}{2^{l+1}}$. Hence, we have

$$|\Pr[Succ_4] - \Pr[Succ_3]| \leq \frac{q_h^2}{2^{l+1}} \quad (5)$$

If A does not make any h query with the correct input, it will not have any advantage in distinguishing the real session key from the random once. Moreover, if the corrupt query Corrupt $(U, 2)$ is made that means the password-corrupt query Corrupt $(U, 1)$ is not made, and the password is used once in local computer to authenticate user for getting some

important information and no more used in the process of the protocol Π. Thus, the probability of A made on-line password guessing attack is at most $\frac{q_s}{D}$. Furthermore, the probability of A made off-line password guessing attack is 0, because even if A gets the secret information $\{ID_A, r_a, B\}$, he has no any compared value to authenticate the guessing password is right or not. Combining the Equations (1) - (5) one gets the announced result as:

$$Adv_{\Pi,D}(A) \leq \frac{2q_h^2}{2^{l+1}} + \frac{(q_s+q_e)^2}{p^2} + 2q_h Adv_G^{cdh}(A) + \frac{2q_h}{p} + \frac{q_s}{D}$$

□

4.2 Further Security Discussion

Proposition 1. *The proposed scheme could resist password guessing attack.*

Proof. In this attack, an adversary may try to guess a legal user password PW using the transmitted messages. Password guessing attack can only crack a function with one low entropy variable (password), so if we at least insert one large random variable which can resist this attack. In our protocol, the adversary only can launch the on-line password guessing attack, because there are no any of the transmitted messages including password as the input value. Even if the adversary gets the secret information $\{ID_A, r_a, B\}$, he has no any compared value to authenticate the guessing password is right or not without the server help. In other words, the adversary cannot construct the form $function(*||PW') = y$, where * is any known message, and only the server can compute the value y. On the other side, about on-line password guessing attack, because the maximum number of allowed invalid attempts about guessing password is only a few times, then the account will be locked by the registration server. □

Proposition 2. *The proposed scheme could support mutual authentication.*

Proof. In our scheme, the Registration Server S_i verifies the authenticity of user A's request by verifying the condition $H'_A = H_A$? during the proposed phase. To compute $B^* = H(ID_A||k_i)$, the password is needed. Therefore, an adversary cannot forge the message. Additionally, C_{A_1}, C_{A_2}, V_A includes large random numbers a and a1, the adversary cannot replay the old message. This shows that S_i can correctly verify the message source.

For Alice authenticating the server S_j, it can be divided three steps: Firstly, S_i transfers the authenticator $H(B^*||T_a(x))$ which can only be decrypted by S_j using his own secret key k_j. Secondly, only S_i or S_j can compute $T_{k_i}T_{k_j}(x)$, so S_j authenticates S_i by verifying the condition $H'_S = H_S$? Finally, Alice authenticates the server S_j by verifying the condition $H'_{S_j} = H_{S_j}$? S_j computes H_{S_j} only by the helping of S_i, and while S_i and S_j have achieved mutual authentication.

Hence, mutual authentication can successfully achieve in our scheme. □

Proposition 3. *The proposed scheme could support Privacy-Protection.*

Proof. Alice's identity is anonymity for outsiders because ID_A is covered by $C_{A_1} = T_{a_1}T_{k_i}(x)ID_A$, and then only the Registration Server S_i can use his secret key to recover the ID_A. It is the same way for covered the identity of S_j. Due to PKC-based about our scheme, the ID_A and ID_{S_j} must be emerged to S_i, or it cannot construct the authenticator of the user and send the covered authenticator to S_j.

For the second message $m_2 = \{T_{r_1}(x), C_{S_1}, C_{S_2}, C_{S_3}, V_S, W_S, H_S\}$, we construct $C_{S_2} = T_{r_2}T_{k_j}(x)ID_A$ to covered Alice's identity, and $C_{S_1} = T_{r_1}T_{k_j}(x)ID_{S_i}$, $C_{S_3} = T_{r_2}T_{k_j}(x)T_a(x)$ to covered S_i's identity and $T_a(x)$. The encrypted message $C_{S_1}, C_{S_2}, C_{S_3}$ are generated from r_1, r_2 which are different in each session and are only known by S_i. S_j can decrypt $C_{S_1}, C_{S_2}, C_{S_3}$ using $T_{r_1}(x)$ and his own secret key which is secure under the CMBDLP and CMBDHP assumptions, and furthermore getting all the information $ID_{S_i} = C_{S_1}/T_{k_j}T_{r_1}(x)$, $ID_A = C_{S_2}/T_{k_j}T_{r_2}(x)$ and $T_a(x) = C_{S_3}/T_{k_j}T_{r_2}(x)$. Additionally, since the values r_1, r_2 of the random elements are very large, attackers cannot directly guess the value r_1, r_2 of the random elements to generate $T_{r_1}(x), T_{r_2}(x)$.

For S_j, because it has know all the necessary information including ID_{S_i}, ID_A, $T_a(x)$ and the covered authenticator $H(B^*||T_a(x))$, S_j only need send the authentication of integrity message $m_3 = \{T_{r_3}(x), H_{S_j}\}$ which has no any of information about identities of the involved three nodes.

Therefore, the proposed scheme provides privacy protection. □

Proposition 4. *The proposed scheme could resist stolen verifier attack.*

Proof. In the proposed scheme, any party stores nothing about the legal users' information. All the en/decrypted messages can be deal with the user's password which is stored in the user's brain, or the secret keys which are covered strictly, so the proposed scheme withstands the stolen verifier attack. □

Proposition 5. *The proposed scheme could withstand replay and man-in-the-middle attacks.*

Proof. The verification messages include the temporary random numbers a, a_1, r_a, r_i. More important thing is that all the temporary random numbers are protected by CDH problem in chaotic maps which only can be uncovered by the legal users (using secret keys or password). So our proposed scheme resists the replay and man-in-the-middle attacks. □

Table 1: Security of our proposed protocol

Category		Eun-Jun Yoon's Scheme [14] (2013)	Zhu's Scheme [5] (2015)	Zhu's Scheme [1] (2016)	Our scheme
Architecture		Multi-server (Centralized)	MSTSA (Distributed)	MSTSA (Distributed)	MSTSA (Distributed)
Architecture properties and functionality	Single-point of security	N/A	Provided	Provided	Provided
	Single-point of efficiency	N/A	Provided	Provided	Provided
	Single-point of failure	N/A	Provided	Provided	Provided
	Symmetry	N/A	Provided	Provided	Provided
	Transparency	**	***	***	***
	Simplicity	*	***	***	***
	Expandability	**	***	***	***
	No timestamp	Provided	Provided	Provided	Provided
	Secure password update	Provided	Provided	Provided	Provided
	Repeatable Authenticated	N/A	N/A	N/A	Provided
Security requirements	Privacy-Protection	N/A	N/A	N/A	Provided
	Mutual authentication	Provided	Provided	Provided	Provided
	Guessing attacks	Provided	Provided	Provided	Provided
	Man-in-the-middle attack	Provided	Provided	Provided	Provided
	Replay attack	Provided	Provided	Provided	Provided
	Key freshness property	Provided	Provided	Provided	Provided
	Perfect forward secrecy	Provided	Provided	Provided	Provided
	Data integrity	Provided	Provided	Provided	Provided
	Impersonation attack	Provided	Provided	Provided	Provided
	Known key secrecy property	Provided	Provided	Provided	Provided
	Stolen verifier attack	Provided	Provided	Provided	Provided
Security Model		Heuristic method	Random Oracle	Standard model	Random Oracle
Required components		**Hardware, software, biometric and password**	**Software and password**	**Software and password**	**Software and password**

N/A: not available or not support. *: provided but in low level.
: provided but in middle level. *: provided but in high level.

Proposition 6. *The proposed scheme could resist user impersonation attack.*

Proof. In such an attack, an adversary may try to masquerade as a legitimate user Alice to cheat another legitimate user. For any adversary, there are two ways to carry this attack:

- The adversary may try to launching the replay attack. However, the proposed scheme resists the replay attack.
- The adversary may try to generate a valid authenticated message $\{T_{a_1}(x), C_{A_1}, C_{A_2}, V_A, H_A\}$ for two random values a, a_1. However, the adversary cannot compute $\{C_{A_1}, C_{A_2}, V_A\}$ as computation of $\{C_{A_1}, C_{A_2}, V_A\}$ requires PW which is only known to legal users.

This shows that the proposed scheme resist user impersonation attack. □

Proposition 7. *The proposed scheme could have Key freshness property.*

Proof. Note that in our scheme, each established session key $SK = H(T_a T_{r_3}(x))$ includes random values a and r_3. The unique key construction for each session shows that proposed scheme supports the key freshness property. □

Proposition 8. *The proposed scheme could have known key secrecy property.*

Proof. In our scheme, if a previously established session key $SK = H(T_a T_{r_3}(x))$ is compromised, the compromised session key reveals no information about other session keys due to following reasons:

- Each session key is hashed with one-way hash function. Therefore, no information can be retrieved from the session key.
- Each session key includes two nonces, which ensures different key for each session.

Since no information about other established group session keys from the compromised session key is extracted, our proposed scheme achieves the known key secrecy property. □

Proposition 9. *The proposed scheme could have forward secrecy.*

Proof. Forward secrecy states that compromise of a legal user's long-term secret key does not become the reason to compromise of the established session keys. In our proposed scheme, the session key has not included the user's long-term secret key: Password. This shows that our scheme preserves the forward secrecy property. □

Proposition 10. *The proposed scheme could have perfect forward secrecy.*

Proof. A scheme is said to support perfect forward secrecy, if the adversary cannot compute the established session key, using compromised secret key k of any server. The proposed scheme achieves perfect forward secrecy. In our proposed scheme, the session key has not included the server's long-term secret key k_i, k_j because the session key is $SK = H(T_a T_{r_3}(x))$. This shows that our scheme provides the perfect forward secrecy property. □

From the Table 1, we can see that the proposed scheme is more secure and has much functionality compared with the recent related scheme.

5 Efficiency Analysis

Compared to RSA and ECC, Chebyshev polynomial computation problem offers smaller key sizes, faster computation, as well as memory, energy and bandwidth savings. In our proposed protocol, no time-consuming modular exponentiation and scalar multiplication on elliptic curves are needed. However, Wang [6] proposed several methods to solve the Chebyshev polynomial computation problem. For convenience, some notations are defined as follows. The computational cost of XOR operation could be ignored when compared with other operations. Table 2 shows performance comparisons between our proposed scheme and the literature of [7] in multi-server architecture and [13, 9] in MSTSA. Therefore, as in Table 2 the concrete comparison data as follows:

Table 2: Efficiency of our proposed scheme

Phase		Eun-Jun Yoon's Scheme [7] (2013)	Zhu's Scheme [9] (2015)	Zhu's Scheme [13] (2016)	Our scheme
A	Total	$1T_{hash}$	Not discussed	$1T_{hash}$	$2T_{hash}$
B	Total	$1T_{hash}$	No need	No need	No need
C	User	$5T_{hash}$ + $2T_{Ecc}$	$4T_{hash}$ +$2T_{sym}$+ $2T_{CH}$	$2T_F$ +$2T_{CH}$	$4T_{hash}$ + $4T_{CH}$
	S_i as a ***RC***/***RC***	$7T_{hash}$	$4T_{hash}$ +$4T_{sym}$+ $2T_{CH}$	$2T_F$ +$4T_{CH}$	$4T_{hash}$ + $5T_{CH}$
	S_j	$5T_{hash}$ +$2T_{Ecc}$	$3T_{hash}$ +$2T_{sym}$+ $2T_{CH}$	$2T_F$ +$2T_{CH}$	$4T_{hash}$ + $5T_{CH}$
	Total	$9T_{hash}$ + $3T_{Exp}$	$11T_{hash}$+ $8T_{sym}$+ $6T_{CH}$	$6T_F$ +$8T_{CH}$	$12T_{hash}$ + $14T_{CH}$
D	Total	$2T_{hash}$	$6T_{hash}$ +$4T_{sym}$+ $2T_{CH}$	$1T_{hash}$ +$2T_F$ +$2T_{CH}$	$7T_{hash}$ + $2T_{CH}$
E	Total	No need	$6T_{hash}$+ $4T_{sym}$+ $2T_{CH}$	No need	No need
F		5	5	4	3

A: User registration **B**: Server registration

C: authentication phase **D**: Password change phase

E: Shared key update among servers phase

F: Rounds of Authentication phase

T_{hash}: The time for executing the hash function;

T_F: The time for executing the pseudo-random function;

- T_{sym}: The time for executing the symmetric key cryptography;

T_{XOR}: The time for executing the XOR operation;

- T_{Exp}: The time for a modular exponentiation computation;
- T_{Ecc}: The time for executing the ECC multiplications

(ECC: Elliptic curve cryptosystem)

- T_{CH}: The time for executing the $T_n(x)$ mod p in Chebyshev polynomial.

6 Conclusion

We only use chaotic maps and a secure one-way hash function to construct a distributed password authenticated key scheme which provides a provable privacy protection towards Multiple Servers to Server Architecture. Our proposed scheme only needs three rounds can catch mutual authenticated with privacy protection among three parties in MSTSA, and the unregistered server can store a temporary authenticator for a certain time without the registered server involved. The above-mentioned innovation points can improve the efficiency of protocol immensely. Based on our discussion we proposed a suitable protocol that covers those goals and offered an efficient protocol that formally meets the proposed security definition. Finally, after comparing with related literatures respectively, we found our proposed scheme has satisfactory security, efficiency and functionality. Therefore, our protocol is more suitable for practical applications.

7 Acknowledgement

This work is supported by the Liaoning Provincial Natural Science Foundation of China (Grant No. 201602680).

References

[1] M. S. Baptista, "Cryptography with chaos," *Physics Letters A*, vol. 240, no. 1, pp. 50-54, 1998.

[2] S. H. Islam, "Provably secure dynamic identity-based three-factor password authentication scheme using extended chaotic maps," *Nonlinear Dynamics*, vol. 78, no. 3, pp. 2261-2276, 2014.

[3] T. F. Lee, "Enhancing the security of password authenticated key agreement protocols based on chaotic maps," *Information Sciences*, vol. 290, pp. 63-71, 2015.

[4] Y. Sun, H. Zhu, and X. Feng, "A novel and concise multi-receiver protocol based on chaotic maps

with privacy protection," *International Journal of Network Security*, vol. 19, No. 3, pp. 371-382, 2017.

[5] H. J. Wang, H. Zhang, J. X. Li and C. Xu, "A (3,3) visual cryptography scheme for authentication," *Journal of Shenyang Normal University (Natural Science Edition)*, vol. 31, no. 3 pp. 397-400, 2013.

[6] X. Wang, J. Zhao, "An improved key agreement protocol based on chaos," *Communications in Nonlinear Science and Numerical Simulation*, vol. 15, pp. 4052-4057, 2010.

[7] E. J. Yoon, K. Y. Yoo, "Robust biometrics-based multi-server authentication with key agreement scheme for smart cards on elliptic curve cryptosystem," *Journal of Supercomputing*, vol. 63, pp. 235-255, 2013.

[8] L. Zhang, "Cryptanalysis of the public key encryption based on multiple chaotic systems," *Chaos Solitons Fractals*, vol. 37, no. 3, pp. 669-674, 2008.

[9] H. F. Zhu, "Flexible and password-authenticated key agreement scheme based on chaotic maps for multiple servers to server architecture," *Wireless Personal Communications*, vol. 82 no. 3, pp. 1697-1718, 2015.

[10] H. F. Zhu, "Secure chaotic maps-based group key agreement scheme with privacy preserving," *International Journal of Network Security*, vol. 18, no. 6, pp. 1001-1009, 2016.

[11] H. F. Zhu, R. Wang, "A survey to design privacy preserving protocol using chaos cryptography," *International Journal of Network Security*, vol. 20, no. 2, pp. 313-322, 2018.

[12] H. Zhu, Y. Zhang, "An improved two-party password-authenticated key agreement protocol with privacy protection based on chaotic maps," *International Journal of Network Security*, vol. 19, No. 4, pp. 487-497, 2017.

[13] H. F. Zhu, Y. F. Zhang, Y. Xia and H. Y. Li, "Password-authenticated key exchange scheme using chaotic maps towards a new architecture in standard model," *International Journal of Network Security*, vol.18, no. 2, pp. 326-334, 2016.

[14] H. F. Zhu, Y. F. Zhang, and Y. Sun, "Provably secure multi-server privacy-protection system based on Chebyshev chaotic maps without using symmetric cryptography," *International Journal of Network Security*, vol. 18, no. 5, pp. 803-815, 2016.

[15] H. F. Zhu, Y. F. Zhang, and Y. Zhang, "A provably password authenticated key exchange scheme based on chaotic maps in different realm," *International Journal of Network Security*, vol. 18, no. 4, pp.688-698, 2016.

[16] H. Zhu, Y. Zhang, Y. Zhang and H. Li, "A novel and provable authenticated key agreement protocol with privacy protection based on chaotic maps towards mobile network," *International Journal of Network Security*, vol. 18, No. 1, pp. 116-123, 2016.

Biography

Hongfeng Zhu obtained his Ph.D. degree in Information Science and Engineering from Northeastern University. Hongfeng Zhu is a full professor of the Kexin software college at Shenyang Normal University. He is also a master's supervisor. He has research interests in wireless networks, mobile computing, cloud computing, social networks, network security and quantum cryptography. Dr. Zhu had published more than 50 international journal papers (SCI or EI journals) on the above research fields.

Junlin Liu graduated with a Bachelor of Engineering from Shenyang Normal University in 2017. In her college, after completing the learning task, she interests in exploring her professional knowledge. During graduate, under the guidance of his master instructor, she researches information security theory and technology.

Cubic Medium Field Equation Public Key Cryptosystem

Gang Lu[1], Linyuan Xue[1], Xuyun Nie[1,2,3], Zhiguang Qin[1,3], and Bo Liu[1]
(Corresponding author: Xuyun Nie)
School of Information and Software Engineering, University of Electronic Science and Technology of China[1]
4 Jianshe North Rd 2nd Section, Chenghua Qu, Chengdu 610054, China
(Email: xynie@uestc.edu.cn)
State Key Laboratory of Information Security, Institute of Information Engineering, Beijing 100093, China[2]
Network and Data Security Key Laboratory of Sichuan Province, Chengdu 610054, China[3]
(Received Dec. 14, 2016; revised and accepted Apr. 11 & May 2, 2017)

Abstract

Medium Field Equation (MFE) multivariate public key cryptosystems were broken by High Order Linearization Equation (HOLE) attack. In order to avoid HOLE attack, we proposed an improvement of MFE, Cubic MFE public key encryption scheme. In our construction, multiplications of three second order matrices were used to get a set of cubic polynomials in the central map. Through theoretical analysis and computer experiments, the Cubic MFE is shown to be secure against HOLE attack and other existing attacks.

Keywords: Cubic Polynomial; High Order Linearization Equation; Medium Field Equation; Multivariate Public Key Cryptosystem

1 Introduction

In 1994, Peter Shor [18] showed that some number theoretic hard problems such as Integer Factorization and Discrete Log Problem can be solved in polynomial time on quantum computer. Once the quantum computer is practical, cryptographic algorithms based on the hard problems above will be no longer secure.

Multivariate Public Key Cryptography (MPKC) could be seen as one of promising candidates to resist quantum algorithm attack [17]. The security of the MPKC relies on the difficulty of solving a system of nonlinear multivariate polynomial equations on a finite field, which is an NP-hard problem in worst case. In general, the public key of MPKC is composed of three maps, two affine maps and a map called central map which is the key point of designing an MPKC.

In the past few decades, a lot of multivariate cryptosystems have been proposed, but many of them were broken. C^* [10] is considered as the first MPKC, which was broken by Patarin[15] with linearization equation attack. Then, Patarin extended the idea of C^* and proposed Hidden Field Equation (HFE) scheme [16]. Ding *et al.* [6] showed that inverting HFE is quasi-polynomial if the size of the field and the degree of the HFE polynomials are fixed. After that, many MPKC encryption schemes have been proposed, such as TTM [11], MFE [21], Square [4] and ABC [19]. Most instances of TTM were broken because there are some linearization equations satisfied by their public key. Square and ABC were broken by differential attack [2, 12] . Cubic ABC [8] then was proposed to resist differential attack. This scheme is still secure by now.

Medium Field Equation (MFE) [21] was proposed by Wang *et al.* in 2006. The inventors of MFE used products of second order matrices to derive quadratic polynomials in its central map. It can be avoid the Paratin relations or linearization equations. But the original MFE was broken by High Order Linearization Equation (HOLE) attack [7] in 2007. In order to resist existing attack, many modifications of MFE were proposed [9, 20, 22] etc. But all of them are insecure [3, 14, 23]. Nie *et al.* [13] pointed out that it is impossible to derive secure MFE by changing the form of second order matrices with their transpose and adjoint.

Although MFE is insecure, the idea of its construction is elegant. And MFE is very efficient. We want to modify its central map to propose a security MFE scheme.

In this paper, we propose a Cubic MFE encryption scheme to avoid HOLE attack. Firstly, we introduce an extra second order matrix in the central map and use products of three second order matrices to get cubic polynomials; secondly, we add three equations in the central map to ensure the successful decryption. Through theoretical analysis and computer experiments, we show that our Cubic MFE scheme can be secure against HOLE attack. Furthermore, the Cubic MFE can resist direct attacks for some chosen parameters. At last, we present efficiency comparison with Cubic ABC and implementation for practical parameters.

This paper is organized as follows. We briefly introduce the original MFE scheme and its cryptanalysis in Section 2. In Section 3, we present our Cubic MFE. And security analysis will be presented in Section 4. In Section 5, we give practical parameters and efficiency comparison. Finally, we conclude this paper in Section 6.

2 Preliminaries

In this section, we will introduce the MFE public key cryptosystem and the previous attack on MFE.

2.1 MFE Public Key Cryptosystem

We use the same notations as in [21]. Let $\mathbb{K}$ be a finite field of characteristic 2 and $\mathbb{L}$ be its degree r extension field. In MFE, we always identify $\mathbb{L}$ with $\mathbb{K}^r$ by a $\mathbb{K}$-linear isomorphism $\pi : \mathbb{L} \to \mathbb{K}^r$. Namely we take a basis of $\mathbb{L}$ over $\mathbb{K}$, $\{\theta_1, \cdots, \theta_r\}$, and define π by $\pi(a_1\theta_1+\cdots+a_r\theta_r) = (a_1, \cdots, a_r)$ for any $a_1, \cdots, a_r \in \mathbb{K}$. It is natural to extend π to two $\mathbb{K}$-linear isomorphisms $\pi_1 : \mathbb{L}^{12} \to \mathbb{K}^{12r}$ and $\pi_2 : \mathbb{L}^{15} \to \mathbb{K}^{15r}$.

In MFE, its encryption map $F : \mathbb{K}^{12r} \to \mathbb{K}^{15r}$ is composed of three maps ϕ_1, ϕ_2, ϕ_3, that is $(y_1, \cdots, y_{15r}) = F(x_1, \cdots, x_{12r}) = \phi_3 \circ \phi_2 \circ \phi_1(x_1, \cdots, x_{12r})$, where $y_1, \cdots, y_{15r}$ are ciphertext variables and $x_1, \cdots, x_{12r}$ are plaintext variables. ϕ_1 and ϕ_3 are invertible affine maps and ϕ_2 is called central map, which is equal to $\pi_2 \circ \bar{\phi}_2 \circ \pi_1^{-1}$.

ϕ_1 and ϕ_3 are taken as the private keys, while the expression of the map $(y_1, \cdots, y_{15r}) = F(x_1, \cdots, x_{12r})$ is the public key. The map $\bar{\phi}_2 : \mathbb{L}^{12} \to \mathbb{L}^{15}$ is defined as follows.

$$\begin{cases} Y_1 = X_1 + X_5X_8 + X_6X_7 + Q_1; \\ Y_2 = X_2 + X_9X_{12} + X_{10}X_{11} + Q_2; \\ Y_3 = X_3 + X_1X_4 + X_2X_3 + Q_3; \\ Y_4 = X_1X_5 + X_2X_7; \\ Y_5 = X_1X_6 + X_2X_8; \\ Y_6 = X_3X_5 + X_4X_7; \\ Y_7 = X_3X_6 + X_4X_8; \\ Y_8 = X_1X_9 + X_2X_{11}; \\ Y_9 = X_1X_{10} + X_2X_{12}; \\ Y_{10} = X_3X_9 + X_4X_{11}; \\ Y_{11} = X_3X_{10} + X_4X_{12}; \\ Y_{12} = X_5X_9 + X_7X_{11}; \\ Y_{13} = X_5X_{10} + X_7X_{12}; \\ Y_{14} = X_6X_9 + X_8X_{11}; \\ Y_{15} = X_6X_{10} + X_8X_{12}. \end{cases} \quad (1)$$

where Q_1, Q_2, and Q_3 form a triple tuple (Q_1, Q_2, Q_3) which is a triangular map from $\mathbb{K}^{3r}$ to itself, more detail please see [21].

The map $\bar{\phi}_2$ can be written by matrix form as follows.

Let $X_1, \cdots, X_{12}$ be the entries of three 2×2 matrices M_1, M_2, M_3, namely,

$$M_1 = \begin{pmatrix} X_1 & X_2 \\ X_3 & X_4 \end{pmatrix}, M_2 = \begin{pmatrix} X_5 & X_6 \\ X_7 & X_8 \end{pmatrix}, \\ M_3 = \begin{pmatrix} X_9 & X_{10} \\ X_{11} & X_{12} \end{pmatrix}. \quad (2)$$

Then $Y_4, \cdots, Y_{15}$ will be the entries in three 2×2 matrices Z_1, Z_2, Z_3, namely,

$$Z_1 = M_1M_2 = \begin{pmatrix} Y_4 & Y_5 \\ Y_6 & Y_7 \end{pmatrix}, \\ Z_2 = M_1M_3 = \begin{pmatrix} Y_8 & Y_9 \\ Y_{10} & Y_{11} \end{pmatrix}, \\ Z_3 = M_2^TM_3 = \begin{pmatrix} Y_{12} & Y_{13} \\ Y_{14} & Y_{15} \end{pmatrix}. \quad (3)$$

Then

$$\begin{cases} \det(M_1) \cdot \det(M_2) = \det(Z_1), \\ \det(M_1) \cdot \det(M_3) = \det(Z_2), \\ \det(M_2) \cdot \det(M_3) = \det(Z_3). \end{cases}$$

Using the determinants of Z_1, Z_2, Z_3, the determinants of M_1, M_2, M_3 can be found. And then one can get the inverse of the central map ϕ_2. More details of decryption are presented in [21].

2.2 High Order Linearization Equation Attack on MFE

The equation of following form is called High Order Linearization Equation (HOLE).

$$\sum_{i=1,j=1}^{n,t} a_{ij}x_if_j(y_1, y_2, \cdots, y_m) \\ + \sum_{j=1}^{l} c_jg_j(y_1, y_2, \cdots, y_m) + d = 0. \quad (4)$$

where f_j, $1 \le j \le t$, g_j, $1 \le j \le l$, are some polynomial functions on ciphertext variables $y_1, y_2, \cdots, y_m$. The highest degree of ciphertext variables y_j is called the order of the Lineraization Equation.

Note that, given a valid ciphertext $y' = (y'_1, y'_2, \cdots, y'_m)$ and substituted it into equation (4), it will become a linear equation on plaintext variables $x_1, \cdots, x_n$.

Once some HOLEs are satisfied by an MPKC, these equations can be used to break the MPKC. The MI scheme was broken by the First Order Linearization Equation (FOLE) method [15]. And the original MEF was broken by the Second Order Linearization Equation (SOLE) method [7].

In the original MFE schemes, the inventors have taken into account the LE attack. They used M_2^T instead of M_2 to avoid the FOLEs.

But many SOLEs were found in the MFE scheme. Denote by M^* the adjoint matrix of a second order matrix M. From

$$Z_1 = M_1M_2, Z_2 = M_1M_3,$$

we have

$$M_3M_3^*M_1^*M_1M_2 = M_3Z_2^*Z_1 = \det(Z_2)M_2. \quad (5)$$

Expanding (4), we get four equations of the following form

$$\sum a'_{ijk}X_iY_jY_k = 0. \quad (6)$$

In [7], 24 equations of this form can be found.

Substituting $(X_1, \cdots, X_{12}) = \pi_1^{-1} \circ \phi_1(x_1, \cdots, x_{12r})$ and $(Y_1, \cdots, Y_{15}) = \pi_2^{-1} \circ \phi_3^{-1}(y_1, \cdots, y_{15r})$ into (8), we get $24r$ equations of the following form

$$\sum_i x_i \left(\sum_{j \le k} a_{ijk} y_j y_k + \sum_j b_{ij} y_j + c_i \right) + \sum_{j \le k} d_{jk} y_j y_k + \sum_j e_j y_j + f = 0. \tag{7}$$

These equations are SOLEs.

Given a public key and a valid ciphertext, after finding all the SOLEs, one can recover the corresponding plaintext efficiently.

3 Cubic MFE

In this section, we will present our Cubic MFE encryption scheme. We use the similar notations as in Section 2.1. The difference is that two $\mathbb{K}$-linear isomorphisms are $\pi_1 : \mathbb{L}^{16} \to \mathbb{K}^{16r}$ and $\pi_2 : \mathbb{L}^{22} \to \mathbb{K}^{22r}$.

3.1 Construction of Central Map

The key point of an MPKC is its central map. Let $X_1, \cdots, X_{16} = \pi_1^{-1} \circ \phi_1(x_1, \cdots, x_{16r})$, and $Y_1, \cdots, Y_{22} = \pi_2^{-1} \circ \phi_3^{-1}(y_1, \cdots, y_{22r})$, the new cencral map $\bar{\phi_2} : \mathbb{L}^{16} \to \mathbb{L}^{22}$ is defined as follows.

$$\begin{cases}
Y_1 = X_1 + X_5X_8 + X_6X_7 + Q_1; \\
Y_2 = X_2 + X_9X_{12} + X_{10}X_{11} + Q_2; \\
Y_3 = X_3 + X_1X_4 + X_2X_3 + Q_3; \\
Y_4 = X_5 + f_5(X_1, X_2, X_3, X_4); \\
Y_5 = X_6 + f_6(X_1, X_2, X_3, X_4, X_5); \\
Y_6 = X_7 + f_7(X_1, X_2, X_3, X_4, X_5, X_6); \\
Y_7 = X_1X_5X_{13} + X_2X_7X_{13} + X_1X_6X_{15} + X_2X_8X_{15}; \\
Y_8 = X_1X_5X_{14} + X_2X_7X_{14} + X_1X_6X_{16} + X_2X_8X_{16}; \\
Y_9 = X_3X_5X_{13} + X_4X_7X_{13} + X_3X_6X_{15} + X_4X_8X_{15}; \\
Y_{10} = X_3X_5X_{14} + X_4X_7X_{14} + X_3X_6X_{16} + X_4X_8X_{16}; \\
Y_{11} = X_1X_9X_{13} + X_2X_{11}X_{13} + X_1X_{10}X_{15} + X_2X_{12}X_{15}; \\
Y_{12} = X_1X_9X_{14} + X_2X_{11}X_{14} + X_1X_{10}X_{16} + X_2X_{12}X_{16}; \\
Y_{13} = X_3X_9X_{13} + X_4X_{11}X_{13} + X_3X_{10}X_{15} + X_4X_{12}X_{15}; \\
Y_{14} = X_3X_9X_{14} + X_4X_{11}X_{14} + X_3X_{10}X_{16} + X_4X_{12}X_{16}; \\
Y_{15} = X_5X_9X_{13} + X_7X_{11}X_{13} + X_5X_{10}X_{14} + X_7X_{12}X_{14}; \\
Y_{16} = X_5X_9X_{15} + X_7X_{11}X_{15} + X_5X_{10}X_{16} + X_7X_{12}X_{16}; \\
Y_{17} = X_6X_9X_{13} + X_8X_{11}X_{13} + X_6X_{10}X_{14} + X_8X_{12}X_{14}; \\
Y_{18} = X_6X_9X_{15} + X_8X_{11}X_{15} + X_6X_{10}X_{16} + X_8X_{12}X_{16}; \\
Y_{19} = X_1X_{13} + X_2X_{15}; \\
Y_{20} = X_1X_{14} + X_2X_{16}; \\
Y_{21} = X_3X_{13} + X_4X_{15}; \\
Y_{22} = X_3X_{14} + X_4X_{16}.
\end{cases} \tag{8}$$

where Q_1, Q_2, and Q_3 form a triple tuple (Q_1, Q_2, Q_3) which is a triangular map from $\mathbb{K}^{3r}$ to itself, and f_5, f_6, f_7 are randomly chosen quadratic polynomials.

The main idea of our improvement is that we use products of three second order matrices in MFE to avoid the HOLEs attack. To do this, it is necessary to introduce a new plaintext variables matrix, M_4 in the matrix form of the central map $\bar{\phi_2}$.

Then the matrix form of the central map $\bar{\phi_2}$ is changed into

$$Z_1 = M_1M_2M_4, Z_2 = M_1M_3M_4, Z_3 = M_2^T M_3 M_4^T.$$

In order to decrypt successfully, we need introduce $Z_4 = M_1M_4$ and Y_4, Y_5, Y_6 in the central map. Let

$$M_4 = \begin{pmatrix} X_{13} & X_{14} \\ X_{15} & X_{16} \end{pmatrix}. \tag{9}$$

Then the matrix form is changed into

$$\begin{aligned}
Z_1 = M_1M_2M_4 &= \begin{pmatrix} Y_7 & Y_8 \\ Y_9 & Y_{10} \end{pmatrix}, \\
Z_2 = M_1M_3M_4 &= \begin{pmatrix} Y_{11} & Y_{12} \\ Y_{13} & Y_{14} \end{pmatrix}, \\
Z_3 = M_2^T M_3 M_4^T &= \begin{pmatrix} Y_{15} & Y_{16} \\ Y_{17} & Y_{18} \end{pmatrix}, \\
Z_4 = M_1M_4 &= \begin{pmatrix} Y_{19} & Y_{20} \\ Y_{21} & Y_{22} \end{pmatrix}.
\end{aligned} \tag{10}$$

Given the values of $Y_1, \cdots, Y_{22}$, the map $\bar{\phi_2}$ can be inverted as follows.

- Firstly, we calculate $\det(Z_1)$, $\det(Z_2)$, $\det(Z_3)$, $\det(Z_4)$. And then we calculate $\det(M_2)$ and $\det(M_3)$ from

  $$\det(Z_1) = \det(M_1)\det(M_2)\det(M_4) = \det(M_2)\det(Z_4),$$

  and

  $$\det(Z_2) = \det(M_1)\det(M_3)\det(M_4) = \det(M_3)\det(Z_4),$$

  respectively.

- Substitute $\det(M_2)$ and $\det(M_3)$ into

  $$\begin{aligned}\det(Z_3) &= \det(M_2^T)\det(M_3)\det(M_4^T) \\ &= \det(M_2)\det(M_3)\det(M_4),\end{aligned}$$

  we can get $\det(M_4)$. Substitute $\det(M_4)$ into $\det(Z_4) = \det(M_1)\det(M_4)$, we can derive $\det(M_1)$.

- Substitute $\det(M_1)$, $\det(M_2)$ and $\det(M_3)$ into

  $$\begin{cases} Y_1 = X_1 + \det(M_2) + Q_1; \\ Y_2 = X_2 + \det(M_3) + Q_2; \\ Y_3 = X_3 + \det(M_1) + Q_3; \end{cases} \tag{11}$$

  We can calculate X_1, X_2, X_3 in turn. And substitute them into $\det(M_1) = X_1X_4 + X_2X_3$, we can get the value of X_4.

- According to the expression of the map $\bar{\phi_2}$, we can calculate X_5, X_6, X_7 in turn. And substitute them into $\det(M_2) = X_5X_8 + X_6X_7$, we can get the value of X_8.

- At last, we can calculate $X_{13}, X_{14}, X_{15}, X_{16}$ and $X_9, X_{10}, X_{11}, X_{12}$ in turn by the expression of the map $\bar{\phi_2}$.

The security analysis can be seen in Section 4.

3.2 Encryption Scheme

Key Generation. Randomly generating two affine maps ϕ_1 and ϕ_3 on $\mathbb{K}^{16r}$ and $\mathbb{K}^{22r}$, respectively. Then calculate the expression of $F : \mathbb{K}^{16r} \to \mathbb{K}^{22r}$, namely,

$$\begin{aligned}(y_1, \cdots, y_{22r}) &= F(x_1, \cdots, x_{16r}) \\ &= \phi_3 \circ \phi_2 \circ \phi_1(x_1, \cdots, x_{16r}).\end{aligned}$$

The private keys are ϕ_1 and ϕ_3.

The public key is the expression of $F : \mathbb{K}^{16r} \to \mathbb{K}^{22r}$, a set of cubic polynomials. The expression of the central map can be public.

Encryption. Given a plaintext $(x'_1, \cdots, x'_{16r})$, the ciphertext $(y'_1, \cdots, y'_{22r})$ is calculated by public key, namely,

$$(y'_1, \cdots, y'_{22r}) = F(x'_1, \cdots, x'_{16r}).$$

Decryption. Given a valid ciphertext $(y'_1, \cdots, y'_{22r})$, the decryption of Cubic MFE is to calculate the inverses of ϕ_3, ϕ_2 and ϕ_1 in turn, namely,

$$(x'_1, \cdots, x'_{16r}) = \phi_1^{-1} \circ \phi_2^{-1} \circ \phi_3^{-1}(y'_1, \cdots, y'_{22r}).$$

4 Security Analysis

In this section, we consider Cubic MFE against several existing attacks, such as linearization equations method and algebraic attacks etc.

Given a public key of an MPKC and a valid ciphertext $y = (y'_1, \cdots, y'_m)$, to break it is equivalent to solve the following system

$$\begin{cases} F_1(x_1, \cdots, x_n) &= y'_1; \\ \cdots & \\ F_m(x_1, \cdots, x_n) &= y'_m. \end{cases} \tag{12}$$

4.1 Linearization Equations Attack

Through theoretical analysis, we did not find any linearization equation satisfied by our Cubic MFE. For example, similar to SOLE attack on MFE, from $Z_1 = M_1M_2M_4$, $Z_4 = M_1M_4$, we can get

$$M_4M_4^*M_1^*M_1M_2M_4 = M_4Z_4^*Z_1 = \det(Z_4)M_2M_4.$$

Expanding it, we get four equations of the form

$$\sum a_{ijkl}X_iX_jY_kY_l + \sum b_{ijk}X_iY_jY_k = 0. \tag{13}$$

Substituting $(X_1, \cdots, X_{16}) = \pi_1^{-1} \circ \phi_1(x_1, \cdots, x_{16r})$ and $(Y_1, \cdots, Y_{22}) = \pi_2^{-1} \circ \phi_3^{-1}(y_1, \cdots, y_{15r})$ into Equation (13), we get $24r$ equations of the form

$$\begin{aligned}&\sum_{i\le j} x_ix_j \left(\sum_{k\le l} a_{ijkl}y_ky_l + \sum_k b_{ijk}y_k + c_{ij} \right) \\ &+ \sum_{k\le l} d_{kl}y_ky_l + \sum_k e_ky_k + f = 0.\end{aligned} \tag{14}$$

From these equations, we can not derive any Linearization Equation.

Furthermore, we did many experiments to verify there is no FOLE and SOLE satisfied by Cubic MFE. This is done as follows. We selected sufficient plaintext/ciphertext pairs and plugged them into the SOLE or FOLE to get a linear system on coefficients of HOLE or FOLE, and then solve it. The experimental results showed that the solutions are all zero, hence no HOLE or FOLE exists.

4.2 Algebraic Attacks

In a direct attack, the attacker wants to recover the plaintext by solving the system (12). The most efficient algorithm for direct attack is Gröbner Basis method such as F_4 and F_5.

According to [5], If $\mathbb{K}$ is big, the complexity of Gröbner Basis method has been proved to be $\mathcal{O}(2^{3n})$ and $\mathcal{O}(2^{2.7n})$ in practice.

In Cubic MFE, if $\mathbb{K} = GF(2^8)$ or $GF(2^{16})$, $r = 3$, $n = 48$, the complexity of Gröbner Basis method is about 2^{129}.

An improvement of Gröbner Basis method, F_5 can be seen in [1]. The complexity of algorithm F_5 relies on the degree of regularity d_{reg} in the algorithm.

Proposition 1. *([1], Proposition 2.2) The complexity of computing a Gröbner basis of a zero-dimensional system of m equations in n variables with F_5 is:*

$$\mathcal{O}\left(m \cdot \binom{n + d_{reg} - 1}{d_{reg}}^{\omega}\right),$$

where d_{reg} is the degree of regularity of the system and $2 < \omega < 3$ is the linear algebra constant.

Unfortunately, we can not determine the degree of regularity in our experiments by Magma. When degree increase to 5, the programs would be out of memory. We estimate the degree of regularity is equal to 6. Hence, the complexity of F_5 on our scheme would be about 2^{83} when $r = 3$.

In summary, the Cubic MFE can resist the direct attack with parameters, $\mathbb{K} = GF(2^8)$ or $GF(2^{16})$, $r = 3$, $n = 48$.

5 Parameter Proposals

Based on the security analysis of Cubic MFE in last section, we recommend $\mathbb{K} = GF(2^8)$ and $GF(2^{16})$, $r = 3$, then $n = 48$ and $m = 66$ for our Cubic MFE.

In Table 1, we present the keys sizes of our Cubic MFE with the paraments recommended and compare them with Cubic Simple Matrix Encryption (CSME) scheme.

From Table 1, we find that the key sizes of our CMFEs are smaller than CSMEs.

The performance of CMFEs ($r = 3$) can be seen in Table 2. We did our experiments with Magma on a normal

Table 1: Parameters and key sizes of CMFEs and comparison with CSMEs

scheme	parameters (k, n, m)	input size(bit)	output size(bit)	public key size(KB)	private key size(KB)
CMFE	$(GF(2^8), 48, 66)$	384	528	1342	6.62
CSME	$(GF(2^8), 49, 98)$	392	784	2115	72.7
CMFE	$(GF(2^{16}), 48, 66)$	768	1056	2684	13.23
CSME	$(GF(2^{16}), 49, 98)$	784	1568	4230	145.4

Table 2: The performance of CMFEs

Field	Encryption Time (ms)	Decryption Time (ms)
$GF(2^8)$	316.72	3.28
$GF(2^{16})$	344.38	3.59

PC with Intel Core i5 CPU@2.53GHz, 3 GB of memory. For each finite field, we randomly chose 100 plaintexts and performed encryptions on them and corresponding decryptions. We calculated the average time in milliseconds of encryptions and decryptions.

6 Conclusion

In this paper, we proposed the Cubic MFE encryption scheme. In our construction, we use multiplications of three second order matrices to get a set of cubic polynomials in the central map. The Cubic MFE is secure against the HOLE attacks and the direct attacks with proper parameters.

The cubic multivariate public key cryptosystems have bigger key sizes than the quadratic multivariate public key cryptosystems. But they can avoided some attacks occurred on the quadratic ones, such as HOLEs attack etc. The security of cubic schemes should be further studied in the future.

Acknowledgments

This work was supported by the National Key Basic Research Program of China under grant 2013CB834203, Major International (Regional) Joint Research Project of China National Science Foundation under grant No.61520106007, The science and technology foundation of Sichuan Province under grant No.2016GZ0065 and the Fundamental Research Funds for the Central Universities under grant No.ZYGX2015J072.

References

[1] L. Bettale, J. C. Faugère, and L. Perret, "Hybrid approach for solving multivariate systems over finite fields," *Journal of Mathematical Cryptology*, vol. 3, no. 3, pp. 177–197, 2009.

[2] O. Billet and M.Matsui, "Cryptanalysis of the square cryptosystems," in *Advances in Cryptology (ASIACRYPT'09)*, pp. 451–468, 2009.

[3] W. W. Cao, X. Y. Nie, L. Hu, X. L. Tang, and J. T. Ding, "Cryptanalysis of two quartic encryption schemes and one improved mfe scheme," in *Proceedings of The Third International Workshop (PQCrypto'10)*, pp. 41–60, May 2010.

[4] C. Clough, J. Baena, J. T. Ding, B. Y. Yang, and M. S. Chen, "Square, a new multivariate encryption scheme," in *Topics in Cryptology (CT-RSA'09)*, pp. 252–264, 2009.

[5] N. Courtois, A. Klimov, J. Patarin, and A. Shamir, "Efficient algorithms for solving overdefined systems of multivariate polynomial equations," in *Advances in Cryptology (EUROCRYPT'00)*, pp. 392–407, 2000.

[6] J. T. Ding and T. J. Hodges, "Inverting HFE systems is quasi-polynomial for all fields," in *Advancees in Cryptology (CRYPTO'11)*, pp. 724–742, 2011.

[7] J. T. Ding, L. Hu, X. Y. Nie, J. Y. Li, and J. Wagner, "High order linearization equation (HOLE) attack on multivariate public key cryptosystems," in *Proceedings of The 10th International Conference on Practice and Theory in Public-Key Cryptography (PKC'07)*, pp. 233–248, Apr. 2007.

[8] J. T. Ding, A. Petzoldt, and L. C. Wang, "The cubic simple matrix encryption scheme," in *6th International Workshop on Post-Quantum Cryptography (PQCrypto'14)*, pp. 76–87, 2014.

[9] J. S. Huang, B. D. Wei, and H. Y. Ou, "An improved MFE scheme resistant against sole attacks," in *Proceedings of Asia Pacific Conference on Postgraduate Research in Microelectronics and Electronics (PrimeAsia'09)*, pp. 157–160, Jan. 2009.

[10] T. Matsumoto and H. Imai, "Quadratic polynomial-tuples for effcient signature verification and message-encryption," in *Advances in Cryptology (Eurocrypt'88)*, pp. 419–453, 1988.

[11] T. Moh, "A public key system with signature and master key functions," *Communication in Algebra*, vol. 18, no. 1, pp. 2207–2222, 1999.

[12] D. Moody, R. Perlner, and D. Smith-Tone, "An asymptotically optimal structural attack on the ABC multivariate encryption scheme," in *6th International Workshop on Post-Quantum Cryptography (PQCrypto'14)*, pp. 180–196, 2014.

[13] X. Y. Nie, C. Y. Hou, Z. H. Xu, and G. Lu, "Analysis of second order matrices construction in MFE public key cryptosystem," *International Journal of Network Security*, vol. 18, no. 1, pp. 158–164, 2016.

[14] X. Y. Nie, Z. H. Xu, L. Lu, and Y. J. Liao, "Security analysis of an improved MFE public key cryptosystem," in *Proceedings of The 10th International Conference on Cryptology and Network Security (CANS'11)*, pp. 118–125, Dec. 2011.

[15] J. Patarin, "Cryptoanalysis of the matsumoto and imai public key scheme of eurocrypt'88," in *Advancees in Cryptology (CRYPTO'95)*, pp. 248–261, 1995.

[16] J. Patarin, "Hidden fields equations (HFE) and isomorphisms of polynomials (IP): two new families of asymmetric algorithms," in *Advances in Cryptology (EUROCRYPT'96)*, pp. 33–48, 1996.

[17] S. Qiao, W. Han, Y. Li, and L. Jiao, "Construction of extended multivariate public key cryptosystems," *International Journal of Network Security*, vol. 18, no. 1, pp. 60-67, 2016.

[18] P. W. Shor, "Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer," *SIAM Review*, vol. 41, no. 2, pp. 303–332, 1999.

[19] C. D. Tao, H. Xiang, A. Petzoldt, and J.T. Ding, "Simple matrix c a multivariate public key cryptosystem (MPKC) for encryption," *Finite Fields and Their Applications*, vol. 35, no. C, pp. 352–368, 2015.

[20] H. W. Tao and Y. X. Chen, "An improved medium-field multivariate public-key encryption scheme," in *Proceedings of The International Conference on Computational Intelligence and Software Engineering*, pp. 1–4, Dec. 2009.

[21] L. C. Wang, B. Y. Yang, Y. H. Hu, and F.P. Lai, "Medium-field multivariate public key encryption scheme," in *Proceedings of The Cryptographers' Track at the RSA Conference*, pp. 132–149, Feb. 2006.

[22] X. Wang, F. Feng, X. M. Wang, and Q. Wang, "A more secure MFE multivariate public key encryption scheme," *International Journal of Computer Science and Applications*, vol. 6, no. 3, pp. 1–9, 2009.

[23] Z. H. Xu, X. Y. Nie, H. Wang, and Y. J. Liao, "Cryptanalysis of an improved MFE public key cryptosystem," *International Journal of Security and Networks*, vol. 7, no. 3, pp. 174–180, 2012.

Biography

Gang Lu is a PH.D candidate in University of Electronic Science and Technology of China now. His research interests include cryptography and security of big data.

Linyuan Xue is pursuing his Master degree from the School of Information and Software Engineering, University of Electronic Science and Technology of China. His research interests include multivariate public key cryptosystems and network security.

Xuyun Nie received the Ph.D. degree in Information security from Graduate university of Chinese Academy of Sciences, Beijing, China, in 2007. He is presently an Associate Professor at University of Electronic Science and Technology of China (UESTC). His research interests include cryptography and information security.

Zhiguang Qin is a professor in the School of Information and Software Engineering, UESTC. He received his PH.D. degree from UESTC in 1996. His research interests include: information security and computer network.

Bo Liu received his Master degree from the School of Information and Software Engineering, University of Electronic Science and Technology of China in 2016. His research interests include multivariate public key cryptosystems and network security.

An Improved Data Hiding Method Based on Lempel-Ziv-Welch Compression Codes

Chin-Chen Chang[1], Ngoc-Tu Huynh[2], Yu-Kai Wang[3], and Yanjun Liu[1]
(Corresponding author: Yanjun Liu)

Department of Information Engineering and Computer Science, Feng Chia University[1]
Taichung 40724, Taiwan
(Email: yjliu104@gmail.com)
College of Information Technology, The University of Danang, Vietnam[2]
Department of Computer Science and Information Engineering, National Chung Cheng University[3]
(Received Dec. 27, 2016; revised and accepted Apr. 11, 2017)

Abstract

Data hiding techniques have been widely used for the last decades to protect secret data by hiding the secret data into a cover file. These techniques are categories based on working domain such as spatial domain, compressed domain and transform domain. In this paper, we propose an improved data hiding method, which employs Lempel-Ziv-Welch (LZW) compression codes to embed secret information since it requires very low computational cost. Our method not only reduces the size of files stored on the disk but also prevents them from being attacked. Experimental results show that the proposed method outperforms some previous schemes in terms of compression rate as well as embedding capacity.

Keywords: Compression Rate; Data Hiding; Embedding Capacity; Lempel-Ziv-Welch (LZW) Compression

1 Introduction

With the development of information technology, there are more and more data kept and transferred. Thus, there are two issues that users always concern about: (1) space of storage and (2) security and confidentiality of their information. For the first issue, in order to save storages, users would prefer some compression methods with faster rates of compression and decompression to compact their data before storing them. Unfortunately, most of effective compression algorithms are computationally expensive. For the second issue there are many traditional block ciphers, such as Data Encryption Standard (DES) and Advanced Encryption Standard (AES) [6,7,14], were proposed to encrypt information but they are not suitable for image encryption because of the redundancy and special storage format of an image. Among various protection methods, steganography technique is one of the most efficient and common methods for image information protection. Steganography techniques, also called data hiding techniques, embed secret information into another host container to avoid suspicion from attackers during transferring information through the Internet or a public channel.

In order to tackle the high complexity of compression methods, a universal lossless compression algorithm called Lempel-Ziv-Welch (LZW) algorithm, was proposed by Abraham Lempel and Jacob Ziv in 1977 and widely used soon due to its achievement of a good compromise between compression performance and speed of execution. Then, LZW was improved by Terry Welch in 1984 [25,28,29]. It is also known as an adaptive compression algorithm that does not assume any prior knowledge of the symbol probabilities.

Data hiding techniques are classified into two categories: reversible [1, 2, 4, 5, 9–13, 15–17, 19, 21–24, 26, 27] and irreversible data hiding [3, 8, 18, 20]. Reversible data hiding schemes are gained more attention since they are suitable for protecting sensitive information such as private information or medical images. In 2013, Wang *et al.* [23] proposed a data hiding method based on LZW code which modifies values of LZW codes to hide secret data. However, in their scheme, the compression rate is still high and data embedding capacity is quite low. In this paper, we propose a new improvement on Wang *et al.*'s method which employed the LZW compression algorithm to solve above two issues at the same time. The proposed scheme achieves a higher protection of secret information while enhancing the compression rate.

The rest of this paper is organized as follows. In Section 2, we briefly review LZW compression method. Our proposed scheme is described in Section 3. In Section 4, we demonstrate the performance of our proposed scheme. Our conclusions are given in Section 5.

2 Related Work

2.1 The LZW Algorithm

The LZ family, including LZ1 [28], LZ2 [29], LZW [25] and their variants are the most popular dictionary-based compression algorithms because they achieve low complexity and good compression rate. LZW algorithm achieves the compression by replacing a repeated sequence of characters with a reference back to its previous occurrence. Performance of the algorithm depends on how such references are represented and on how to select the sequences that are replaced. We now briefly review the LZW algorithm.

The algorithm compresses different lengths of sub-strings to a same length of compression code. That is to say, the user can acquire a same length of index. The characteristic of this algorithm is simple and effective to be implemented. Creating a dictionary including all of the characters in the input file is the first step of this algorithm. Secondly, we find the sub-string which can be expressed in the dictionary. Then, the sub-string will be combined with the next character. After that, we update it in the dictionary and output the index of the sub-string. The file is read repeatedly and the dictionary is updated until all of the characters in the file are read. Finally, we output is the index numbers. The details of the LZW algorithm are described as follows.

Algorithm 1 Compression procedure of the original LZW algorithm

Input. Source file which is needed to be compressed

Output. LZW codes

Step 1. Create a dictionary including all of the characters.

Step 2. Scan a sub-string w from the input file which can be found in the dictionary.

Step 3. Combine the sub-string w with the next character c, and give new index $I(w||c)$ and add it to the dictionary.

Step 4. Output the index of the sub-string $I(w)$ and remove the sub-string w in the input file.

Step 5. Repeat Step 2 through Step 4 until the source file is compressed.

After compressing, we have the LZW compression code with shorter size. Whenever the user wants to decompress to recover the original data, he/she will perform the decompression procedure which is described below.

Algorithm 2 Decompression procedure of the original LZW algorithm

Input. Compression file

Output. Source file which was decompressed

Step 1. Create a dictionary including all of the character.

Step 2. Scan a LZW code $I(w)$ from the compression file and extract the corresponding symbol w.

Step 3. If the next index is equal to the dictionary plus one, combine the symbol w with the first character c of itself; otherwise, combine the symbol w with the first character c of the next LZW code, then give new index $I(w||c)$ and add it to the dictionary.

Step 4. Output the index of the symbol w and remove the LZW code $I(w)$ in the compression file.

Step 5. If all of LZW codes from the compression file are read, then the process is ended and output the decompressed data; otherwise, repeat Step 2 through Step 4.

The following example gives a clearer understanding of LZW algorithm. Assume that we have a dictionary which is represented by 8 bits in ASCII and the length of the compression code is 9 bits. Table 1 shows the LZW algorithm for the input string "aabaababbaabbabaaabb".

Table 1: An example to illustrate LZW algorithm

Compression			Decompression		
Input	Updated dictionary	Output	Input	Output	Updated dictionary
aa	aa=256	97	97	a	
ab	ab=257	97	97	a	aa=256
ba	ba=258	98	98	b	ab=257
aab	aab=259	256	256	aa	ba=258
bab	bab=260	258	258	ba	aab=259
bb	bb=261	98	98	b	bab=260
baa	baa=262	258	258	ba	bb=261
abb	abb=263	257	257	ab	baa=262
baba	baba=264	260	260	bab	abb=263
aaa	aaa=265	256	256	aa	baba=264
abb		263	263	abb	aaa=265

2.2 The HPDH-LZW Scheme

In 2013, Wang *et al.* [23] proposed a high performance reversible data hiding scheme based on LZW algorithm (HPDH-LZW). The main idea of the paper is that the compression code is divided into two ranges: in the dictionary or not. In the hiding phase, if the secret bit is 0, output the original compression index; if the secret bit is 1, calculate the original compression index, and then add the number of dictionary size and output it. In the extraction phase, if the value of current processing code is larger than current size of dictionary, we extract secret bit "1" and recover the original index by calculating the difference between this code and the dictionary size. Otherwise, we extract bit "0" and the original LZW code is equal to the current code. The details of the HPDH-LZW algorithm are described as follows.

The details of the HPDH-LZW decompression and secret extraction procedure are described below.

In the following example, we assume that a dictionary is represented by 8 bits in ASCII, the length of the com-

Algorithm 3 The HPDH-LZW algorithm

Input. Source file which is needed to be compressed and secret data

Output. LZW codes

Step 1. Create a dictionary which includes all characters.

Step 2. Scan a sub-string w from the input file, which can be found in the dictionary.

Step 3. Combine the sub-string w with the next character c from the source file, and add new index $I(w||c)$ to the dictionary.

Step 4.1. If secret bit is "0", then output the index of the sub-string $I(w)$. If the secret bit is "1", then output: $I(w) + ds$.

Step 4.2. Remove the sub-string w in the input file.

Step 5. Repeat Step 2 through Step 4 until the source file is compressed.

Algorithm 4 The HPDH-LZW decompression algorithm

Input. Compressed file

Output. The decompressed source file and secret file

Step 1. Create a dictionary size ds which includes all of the characters.

Step 2. Scan a LZW code $I(w)$ from the compression file. If $I(w)$ is more than the dictionary size ds, we extract the secret bit as "1" and recover the original LZW code $I(w) - ds$; otherwise, we extract the secret bit as "0" and the original LZW code is equal to $I(w)$. The corresponding symbol w is extracted according to the original LZW code.

Step 3. If the next LZW code is equal to $ds + 1$, we combine the symbol w with its first character c; otherwise, combine the symbol w with the first character c of the next LZW code, then give new index $I(w||c)$ and add it to the dictionary.

Step 4. Output the index of the symbol w and the secret bit, then remove the LZW code $I(w)$ in the compressed file.

Step 5. If all LZW codes from the compression file are read, output the decompression data and secret file; otherwise, repeat Step 2 to Step 4.

pression code is 10 bits and the largest size of the dictionary is 9 bits. It means that, the original dictionary indices are from 0 to 255, the new dictionary indices are from 256 to 511 and the highest index is 1023. Table 2 illustrates the HPDH-LZW algorithm for the input string "aabaababbaabbabaaabb" and the string of secret bits "1101101001".

It is noted that in the HPDH-LZW scheme, the size of the compression code and the capacity of secret bits are in inverse proportion, but the size of the compression code and the size of the compression file are in direct proportion.

3 Proposed Scheme

Our proposed scheme is called IDH-LZW, i.e., improved data hiding method based on Lempel-Ziv-Welch compression. In our proposed scheme, the values of LZW codes are modified to embed secret bits of different lengths while avoiding changing the content of the dictionary. That is to say, every new symbol inserted into the dictionary is used to embed secret bits of different lengths.

3.1 IDH-LZW Embedding Algorithm

The proposed embedding algorithm is described in Algorithm 5 as follows.

Algorithm 5 IDH-LZW embedding algorithm

Input. Source file and secret file

Output. Compression codes with secret

Firstly, the user must define how many bits to represent one compression code C_size and how many indices in the dictionary Ds.

Step 1.

- Read the character c_i from the source file.
- Set $s = sp||c_i$, where s is a string variable, sp is previous symbol and $||$ means the concatenation operation. If $i = 0$ then $s = c_0$.

Step 2.

If s exists in the dictionary

- Set the previous symbol $sp = s$.

Else

- Compute $kp = Ds/ds$, where ds is the current size of the dictionary, kp is the hidden fragment and the value of kp is between 2^k and 2^{k+1}.
  - Get k secret bits b_k from the secret file and sb_k is b_k in decimal.
  - Get the code C_i, where C_i is the dictionary index of sp.
  - Set $C_i^{'} = C_i + ds * sb_k$, where ds is the current size of the dictionary.
  - Output $C_i^{'}$, where $C_i^{'}$ is the compression code with secret.

Step 3.

If $ds < Ds$

- Add s into the dictionary.
- Set $s = c_i$, where c_i is the last character of s.

If the source file is remained

- Repeat Step 1 to Step 3.

Step 4.

In case sp still has data without output after Step 3

- Set $kp = Ds/ds$.
- Take a secret digit sb_k from the secret file.
- Get code $C_i^{'}$ and set $C_i^{'} = C_i + ds * sb_k$.
- Output $C_i^{'}$.

Step 5.

- Transform $C_i^{'}$ to binary code Cb_i where binary code size is C_size.
- The final compression code is C, where C is the concatenation of Cb_i.

Table 2: An example to illustrate HPDH-LZW algorithm

Compression				Decompression			
Input	Updated dictionary	Hidden bit	Output	Input	Extracted bit	Output	Updated dictionary
aa	aa=256	1	353	353	1	a	
ab	ab=257	1	354	354	1	a	aa=256
ba	ba=258	0	98	98	0	b	ab=257
aab	aab=259	1	515	515	1	aa	ba=258
bab	bab=260	1	518	518	1	ba	aab=259
bb	bb=261	0	98	98	0	b	bab=260
baa	baa=262	1	520	520	1	ba	bb=261
abb	abb=263	0	257	257	0	ab	baa=262
baba	baba=264	0	260	260	0	bab	abb=263
aaa	aaa=265	0	256	256	0	aa	baba=264
abb		1	529	529	1	abb	aaa=265

3.2 IDH-LZW Decompression and Secret Extraction Algorithm

In the decompression phase, it is noted that the number of bits used to represent one compression code C_size and the number of indices in the dictionary Ds are shared between the dealer and the receiver. This information is used to extract the secret bits and decompress the compressed file. Algorithm 6 is used to demonstrate the decompression and extraction procedure of our method.

Algorithm 6 IDH-LZW decompression and secret extraction algorithm

Input. Compression codes

Output. Reconstructed source file and recovered secret file

Step 1.

- Cut C_size bits to present one LZW compression code $C_i^{'}$.

Step 2.

- Read a LZW code $C_0^{'}$.
- Compute $kp = Ds/ds$, where kp is the embedding range, Ds is the number of the dictionary and ds is the current size of the dictionary. The value of kp is from 2^k to 2^{k+1}.
- Compute $C_0 = C_0^{'}/ds$ and get the remainder $sb_k = C_0^{'} \bmod d$, where C_0 is the original compression code, sb_k are extracted secret bits in decimal and b_k are k secret bits of sb_k in binary.
- Get w where w is the symbol of C_0 in the dictionary.

Step 3.

- Read the next LZW code $C_i^{'}$.

If ds is equal to Ds

$kp = Ds/ds$.

Else

$kp = Ds/(ds + 1)$.

- Compute $C_i = C_i^{'}/ds$, get the remainder $sb_k = C_i \bmod d$ and extract secret bit b_k from sb_k.
- Output b_k.

Step 4.

- Set c_i to be the source file value.

If C_i exists in the dictionary

- Get c_i, where c_i is the symbol of C_i in the dictionary.

Else if C_i is equal to ds

- Set $c_i = w||$(the first symbol of w).

If $(ds < Ds)$

- Put $c_i = w||$(the first symbol of w) into the dictionary.
- Set $w = c_i$.
- Output c_i.

Step 5.

- Concatenate c_i to reconstruct the source file and b_k to recover the secret file.

We illustrate all the steps of our IDH-LZW algorithm in Table 3 as follows. Here, we give a dictionary which is represented by 8 bits in ASCII, the length of compression code is 10 bits and the maximum size of the dictionary is 9 bits as an example. It means that, original dictionary indices are from 0 to 255, new dictionary indices are from 256 to 511 and the maximum index is 1023. Table 3 shows the proposed IDH-LZW algorithm for the input string "aabaababbaabbabaaabb" and the string of secret bits "110110100011".

4 Experimental Results

To show that our proposed scheme is suitable for most data formats, we conduct experiments on different kinds of files such as text files, binary images, grayscale images and color images with different file sizes. Figure 1 shows the set of test images and Figure 2 shows the text files used to implement our method.

4.1 Compression Performance

Our IDH-LZW scheme can embed a large amount of secret bits while compressing the source file. As a result, we can save storage as well as protect our data. We conduct our experiments on different kinds of file formats such as text files, binary images, grayscale images sized

Table 3: Example to illustrate IDH-LZW algorithm

Compression				Decompression			
Input	Updated dictionary	Hidden bit	Output	Input	Extracted bit	Output	Updated dictionary
aa	aa=256	11	865	865	11	a	
ab	ab=257	0	97	97	0	a	aa=256
ba	ba=258	1	356	356	1	b	ab=257
aab	aab=259	1	515	515	1	aa	ba=258
bab	bab=260	0	258	258	0	ba	aab=259
bb	bb=261	1	359	359	1	b	bab=260
baa	baa=262	0	258	258	0	ba	bb=261
abb	abb=263	0	257	257	0	ab	baa=262
baba	baba=264	0	260	260	0	bab	abb=263
aaa	aaa=265	1	521	521	1	aa	baba=264
abb		1	529	529	1	abb	aaa=265

(a) Baboon (grayscale) (b) Lena (grayscale) (c) Boat (grayscale) (d) Peppers (grayscale)

(a) Baboon (color) (b) Lena (color) (c) Boat (color) (d) Peppers (color)

Figure 1: Test images

Four score and seven years ago our fathers brought forth, on this continent, a new nation, conceived in Liberty, and dedicated to the proposition that all men are created equal. Now we are engaged in a great civil war, testing whether that nation, or any nation so conceived and so dedicated, can long endure. We are met on a great

(a) The Gettysburg address

I am happy to join with you today in what will go down in history as the greatest demonstration for freedom in the history of our nation.Five score years ago, a great American, in whose symbolic shadow we stand today, signed the Emancipation Proclamation. This momentous decree came as a great beacon light of hope to

(b) I have a dream

My fellow citizens: I stand here today humbled by the task before us, grateful for the trust you've bestowed, mindful of the sacrifices borne by our ancestors. I thank President Bush for his service to our nation as well as the generosity and cooperation he has shown throughout this transition. Forty-four Americans have now taken the

(c) Obama

THE LITTLE PRINCE Antoine De Saint-Exupery Antoine de Saint-Exupery, who was a French author, journalist and pilot wrote The Little Prince in 1943, one year before his death. The Little Prince appears to be a simple children's tale, some would say that it is actually a profound and deeply moving tale, written in riddles and laced

(d) The little prince

BRAVE NEW WORLD by Aldous Huxley (1894-1963) Chapter One A SQUAT grey building of only thirty-four stories. Over the main entrance the words, CENTRAL LONDON HATCHERY AND CONDITIONING CENTRE, and, in a shield, the World State's motto, COMMUNITY, IDENTITY, STABILITY. The enormous room

(e) Brave new world

Figure 2: Text files

256×256, grayscale images sized 512×512 and color images, as shown in Tables 4-8, respectively. In these tables, the "LZW size" parameter is the size of LZW codes, measured by the number of indices. The "capacity" values are taken under different dictionary sizes (ds) on various images.

4.2 Embedding Performance

The proposed IDH-LZW scheme aims to embed secret information into compressed files. Therefore, the embedding rate is concerned to evaluate the performance of an algorithm. To further demonstrate that our scheme achieves high embedding rate, Tables 9-11 show the number of hidden bits per LZW code measured by byte on different file formats. It can be seen from the tables that, a larger size of dictionary is increased and a larger amount of secret bits is embedded. However, the compression rate is higher. Moreover, among these tables, we can see that the embedding rate in Table 10 is the highest while the compression rate is low. The reason is that in grayscale images, many pixels are the same as their neighbors.

4.3 Comparisons

In this subsection, we implement Wang *et al.*'s scheme (HPDH-LZW) and compare the results with our proposed scheme in terms of low compression rate and high embedding capacity. The "decreased size" parameter is used to evaluate the size of compression code that is reduced while comparing to Wang *et al.*'s scheme. The "increased bits" parameter is conducted to show that our scheme can embed more data into host compression codes. Tables 12-16 obviously show that our proposed scheme is better than HPDH-LZW scheme in terms of compression rate and embedding capacity. Furthermore, graphs in Figure 3 visibly show our comparisons.

5 Conclusions

In this paper, we proposed a novel compression-based data hiding scheme called IDH-LZW which not only solves

Table 4: The capacity performance of IDH-LZW scheme for text files

C_size	10		11		12		13	
Ds	9		10		11		12	
File name	LZW size (index)	Capacity	LZW size (index)	Capacity	LZW size (index)	Capacity	LZW size (index)	Capacity
The Gettysburg address	803	804	717	975	717	1692	717	2409
I have a dream	5090	5091	4040	4298	3538	4565	3325	6145
Obama	7566	7567	5819	6077	5058	6085	4680	7500
Brave new world	237230	237231	179303	179561	146171	147198	126063	128883
The little prince	54821	54822	40392	40650	33474	34501	28055	30875

Table 5: The capacity performance of IDH-LZW scheme for binary images

C_size	10		11		12		13		14		15	
Ds	9		10		11		12		13		14	
File name	LZW size (index)	Capacity	LZW size (index)	Capacity	LZW size (index)	Capacity	LZW size (index)	Capacity	LZW size (index)	Capacity	LZW size (index)	Capacity
Airplane	7610	8112	6717	7730	5885	7921	5501	9584	5467	13645	5467	19112
Boat	9200	9702	8028	9041	7156	9192	6671	10754	6543	14721	6543	21264
Gold	12408	12910	10757	11770	9633	11669	8952	13035	8665	16843	8640	25009
Lena	8025	8527	6807	7820	6220	8256	5911	9994	5811	13989	5811	19800
Peppers	7190	7692	6147	7160	5775	7811	5533	9616	5491	13669	5491	19160

Table 6: The capacity performance of IDH-LZW scheme for grayscale images (sized 256×256)

C_size	16		17		18		19		20	
Ds	15		16		17		18		19	
File name	LZW size (index)	Capacity	LZW size (index)	Capacity	LZW size (index)	Capacity	LZW size (index)	Capacity	LZW size (index)	Capacity
Baboon	40104	70831	39977	103217	39977	143194	39977	183171	39977	223148
Barbara	40497	71224	40257	103497	40257	143754	40257	184011	40257	224268
Boat	34849	65576	34818	98058	34818	132876	34818	167694	34818	202512
Family	38009	68736	37895	101135	37895	139030	37895	176925	37895	214820
Girl	33797	64524	33789	97029	33789	130818	33789	164607	33789	198396
Lena	34861	65588	34843	98083	34843	132926	34843	167769	34843	202612
Peppers	351997	65926	35149	98389	35149	133538	35149	168687	35149	203836
Toys	32193	62920	32193	95113	32193	127306	32193	159499	32193	191692

Table 7: The capacity performance of IDH-LZW scheme for grayscale images (sized 512×512)

C_size	16		17		18		19		20	
Ds	15		16		17		18		19	
File name	LZW size (index)	Capacity	LZW size (index)	Capacity	LZW size (index)	Capacity	LZW size (index)	Capacity	LZW size (index)	Capacity
Baboon	146755	177482	136511	199751	133149	261670	132994	392332	132994	525326
Boat	146376	177103	124444	187684	109856	238377	109856	348233	109856	458089
F16	114876	145603	96009	159249	94785	223306	94785	318091	94785	412876
Lena	127139	157866	117295	180535	114434	242955	114434	357389	114434	471823
Peppers	129000	159727	118875	182115	116070	244591	116070	360661	116070	476731
Gold	140641	171368	122040	185280	118513	247034	118513	365547	118513	484060

Table 8: The capacity performance of IDH-LZW scheme for color images (sized 512×512)

C_size	16		17		18		19		20	
Ds	15		16		17		18		19	
File name	LZW size (index)	Capacity	LZW size (index)	Capacity	LZW size (index)	Capacity	LZW size (index)	Capacity	LZW size (index)	Capacity
Baboon	509088	539815	470992	534232	428694	557215	407172	666510	402207	923434
Boat	559908	590635	510732	573972	426127	554648	344134	603472	337656	858883
F16	430883	461610	399157	462397	320466	448987	288861	548199	288400	809627
Lena	190343	221070	153265	216505	135890	264411	135842	395180	135842	531022
Peppers	516584	547311	490857	524097	415514	544035	387969	647307	382513	903740
Gold	588303	619030	469542	532782	384609	513130	361901	621239	356391	877618

Table 9: The embedding rate of the proposed scheme for text files

C_size	10		11		12		13	
Ds	9		10		11		12	
File name	LZW size (byte)	Hidden bit per byte	LZW size (byte)	Hidden bit per byte	LZW size (byte)	Hidden bit per byte	LZW size (byte)	Hidden bit per byte
The Gettysburg address	1003.75	0.80	985.88	0.99	1075.50	1.57	1165.13	2.07
I have a dream	6362.50	0.80	5555.00	0.77	5307.00	0.86	5403.13	1.14
Obama	9457.50	0.80	8001.13	0.76	7587.00	0.80	7605.00	0.99
Brave new world	296537.50	0.80	246541.63	0.73	219256.50	0.67	204852.38	0.63
The little prince	68526.25	0.80	55539.00	0.73	50211.00	0.69	45589.38	0.68

Table 10: The embedding rate of the proposed scheme for grayscale images (sized 256×256)

C_size	16		17		18		19		20	
Ds	15		16		17		18		19	
File name	LZW size (byte)	Hidden bit per byte	LZW size (byte)	Hidden bit per byte	LZW size (byte)	Hidden bit per byte	LZW size (byte)	Hidden bit per byte	LZW size (byte)	Hidden bit per byte
Baboon	80208.00	0.88	84951.13	1.22	89948.25	1.59	94945.38	1.93	99942.50	2.23
Barbara	80994.00	0.88	85546.13	1.21	90578.25	1.59	95610.38	1.92	100642.50	2.23
Boat	69698.00	0.94	73988.25	1.33	78340.50	1.70	82692.75	2.03	87045.00	2.33
Family	76018.00	0.90	80526.88	1.26	85263.75	1.63	90000.63	1.97	94737.50	2.27
Girl	67594.00	0.95	71801.63	1.35	76025.25	1.72	80248.88	2.05	84472.50	2.35
Lena1	69722.00	0.94	74041.38	1.32	78396.75	1.70	82752.13	2.03	87107.50	2.33
Peppers	70398.00	0.94	74691.63	1.32	79085.25	1.69	83478.88	2.02	87872.50	2.32
Toys	64386.00	0.98	68410.13	1.39	72434.25	1.76	76458.38	2.09	80482.50	2.38

Table 11: The embedding rate of the proposed scheme for color images (sized 512×512)

C_size	16		17		18		19		20	
Ds	15		16		17		18		19	
File name	LZW size (byte)	Hidden bit per byte	LZW size (byte)	Hidden bit per byte	LZW size (byte)	Hidden bit per byte	LZW size (byte)	Hidden bit per byte	LZW size (byte)	Hidden bit per byte
F16	861766	0.54	848209	0.55	721049	0.62	686045	0.80	721000	1.12
Baboon	1018176	0.53	1000858	0.53	964562	0.58	967034	0.69	1005518	0.92
Boat	1119816	0.53	1085306	0.53	958786	0.58	817318	0.74	844140	1.02
Gold	1176606	0.53	997777	0.53	865370	0.59	859515	0.72	890978	0.99
Lena	380686	0.58	325688	0.66	305753	0.86	322625	1.22	339605	1.56
Peppers	1033168	0.53	1043071	0.50	934907	0.58	921426	0.70	956283	0.95

Table 12: Comparisons between IDH-LZW scheme and Wang *et al.*'s scheme conducted on binary images (sized 512×512)

File name	Original file size	HPDH-LZW		Proposed scheme		Decreased size	Increased bits
		LZW size (C_size = 10 and $Ds = 9$)	Capacity	LZW size (C_size = 13 and $Ds = 12$)	Capacity		
F16	32768	9513	7610	8939	9584	6.03%	25.94%
Boat	32768	11500	9200	10840	10754	5.74%	16.89%
Goldhill	32768	15510	12408	14547	13035	6.21%	5.05%
Lena	32768	10031	8025	9605	9994	4.25%	24.54%
Peppers	32768	8988	7190	8991	9616	-0.04%	33.74%

Table 13: Comparisons between IDH-LZW scheme and Wang *et al.*'s scheme conducted on grayscale images (sized 256×256)

File name	Original file size	HPDH-LZW		Proposed scheme		Decreased size	Increased bits
		LZW size (C_size = 10 and $Ds = 9$)	Capacity	LZW size (C_size = 13 and $Ds = 12$)	Capacity		
Baboon	65536	80886	64709	80208	70831	0.84%	9.46%
Barbara	65536	79999	63999	80994	71224	-1.24%	11.29%
Boat	65536	73238	58590	69698	65576	4.83%	11.92%
Family	65536	79809	63847	76018	68736	4.75%	7.66%
Girl	65536	77253	61802	67594	64524	12.50%	4.40%
Lena	65536	76428	61142	69722	65588	8.77%	7.27%
Peppers	65536	77546	62037	70398	65926	9.22%	6.27%
Toys	65536	80878	64702	64386	62920	20.39%	-2.75%

Table 14: Comparisons between IDH-LZW scheme and Wang *et al.*'s scheme conducted on grayscale images (sized 512×512)

File name	Original file size	HPDH-LZW		Proposed scheme		Decreased size	Increased bits
		LZW size (C_size = 10 and $Ds = 9$)	Capacity	LZW size (C_size = 13 and $Ds = 12$)	Capacity		
Airplane	262144	315065	252052	225114	318091	28.55%	26.20%
Baboon	262144	321820	257456	302155	382967	6.11%	748.75%
Barbara	262144	320539	256431	300846	381865	6.14%	48.92%
Boat	262144	279783	223826	260908	348233	6.75%	55.58%
Elaine	262144	313738	250990	284174	367825	9.42%	46.55%
Family	262144	316699	253359	271211	356909	14.36%	40.87%
Girl	262144	299851	239881	257669	345505	14.07%	44.03%
Gold	262144	301103	240882	265326	351953	11.88%	46.11%
Lena	262144	300391	240313	254852	343133	15.16%	42.79%
Peppers	262144	305998.8	244799	252750	341363	17.40%	39.45%
Sailboat	262144	316698.8	253359	274785	359919	13.23%	42.06%
Bridge	262144	277127.5	221702	200623	297467	27.61%	34.17%
Tiffany	262144	286201.3	228961	232441	324261	18.78%	41.62%
Toys	262144	306772.5	245418	243457	333537	20.64%	35.91%
Zelda	262144	301776.3	241421	248214	337543	17.75%	39.82%

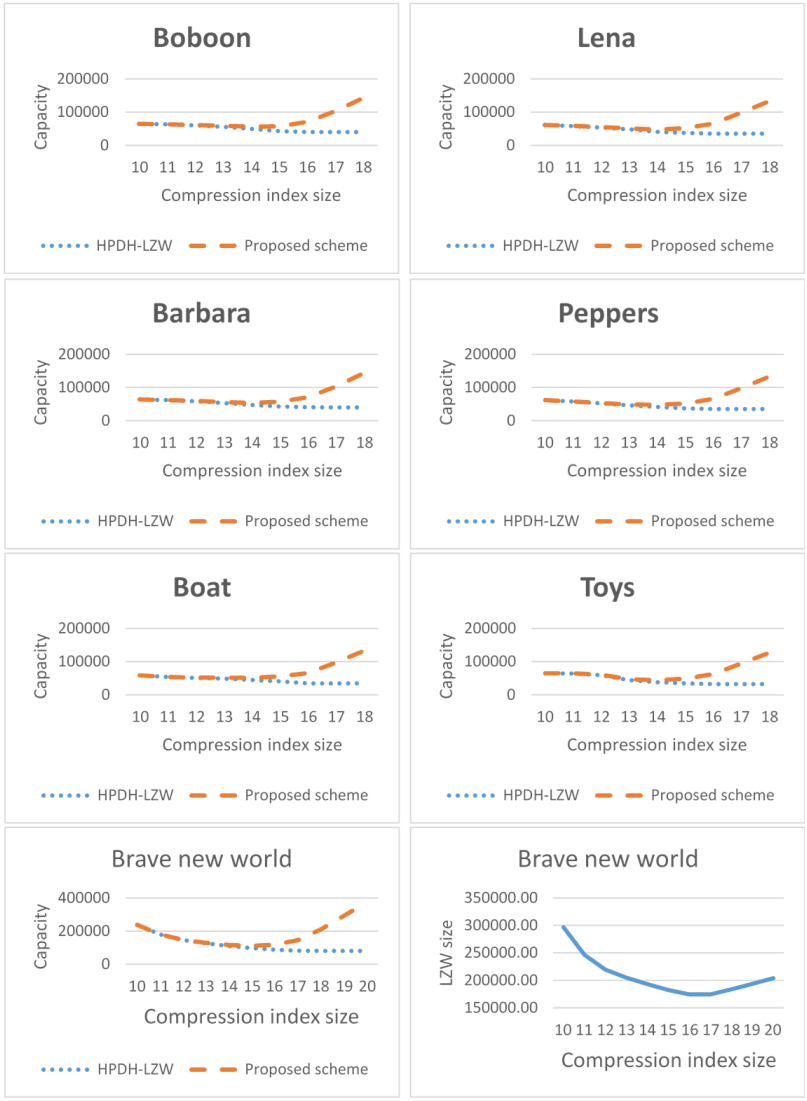

Figure 3: Comparisons between HPDH-LZW and our method

Table 15: Comparisons between IDH-LZW scheme and Wang *et al.*'s scheme conducted on color images (sized 512×512)

File name	Original file size	HPDH-LZW		Proposed scheme		Decreased size	Increased bits
		LZW size (R+G+B) ($C\_size = 10$ and $Ds = 9$)	Capacity	LZW size (R+G+B) ($C\_size = 13$ and $Ds = 12$)	Capacity		
Airplane	786432	894215	715372	711305	984557	20.45%	37.63%
Baboon	786432	971015	776812	970988	1186851	0.00%	52.78%
Boat	786432	912838	730270	775112	1038289	15.09%	42.18%
Gold	786432	948944	759155	835991	1089555	11.90%	43.52%
Lena	786432	693664	554931	464123	771629	33.09%	39.05%
Peppers	786432	934743	747794	829136	1083783	11.30%	44.93%

Table 16: Comparisons between IDH-LZW scheme and Wang *et al.*'s scheme conducted on text files

File name	Original file size	HPDH-LZW		Proposed scheme				Decreased size	Increased bits
		LZW size ($C\_size = 10$ and $Ds = 9$)	Capacity	C_size	Ds	LZW size	Capacity		
The Gettysburg Address	1461	1004	803	11	10	986	975	1.78%	21.42%
I have a dream	9167	6363	5090	14	13	5819	9470	8.55%	86.05%
Obama	13507	9458	7566	14	13	8160	11324	13.72%	49.67%
Brave new world	375467	296538	237230	19	18	193791	291713	34.65%	22.97%
The little prince	91141	68526	54821	17	16	46963	74927	31.47%	36.68%

the issue of data security but also reduces the size of storage. Experimental results show that our proposed scheme achieves good compression rate and high embedding capacity. Moreover, our proposed scheme has very low computation cost but guarantees the efficiency, that is suitable for real time applications.

References

[1] K. Bharanitharan, C. C. Chang, H. R. Yang, and Z. H. Wang, "Efficient pixel prediction algorithm for reversible data hiding," *International Journal of Network Security*, vol. 18, no. 4, pp. 750–757, 2016.

[2] D. Cavagnino, M. Lucenteforte, and M. Grangetto, "High capacity reversible data hiding and content protection for radiographic images," *Signal Processing*, vol. 117, pp. 258–269, 2015.

[3] H. Chen, X. Du, Z. Liu, and C. Yang, "Optical color image hiding scheme by using gerchberg-saxton algorithm in fractional fourier domain," *Optics and Lasers in Engineering*, vol. 66, pp. 144–151, 2015.

[4] D. Coltuc, "Low distortion transform for reversible watermarking," *IEEE Transactions on Image Processing*, vol. 21, no. 1, pp. 412–417, 2012.

[5] G. Galambos and J. Bekesi, "Data compression: theory and techniques. Department of informatics, teacher's training college," *Database and Data Communication Network Systems*, vol. 1, 2002.

[6] T. Gulom, "The encryption algorithms GOST28147-89-IDEA8-4 and GOST28147-89-RFWKIDEA8-4," *International Journal of Electronics and Information Engineering*, vol. 6, no. 2, pp. 59-71, 2017.

[7] T. Gulom, "The encryption algorithm GOST28147-89-PES16-2 and GOST28147-89-RFWKPES16-2," *International Journal of Electronics and Information Engineering*, vol. 6, no. 1, pp. 1–11, 2017.

[8] R. Jafari, D. Ziou, and M. M. Rashidi, "Increasing image compression rate using steganography," *Expert Systems with Applications*, vol. 40, no. 17, pp. 6918–6927, 2013.

[9] B. Jana, D. Giri and S. K. Mondal, "Dual-image based reversible data hiding scheme using pixel value difference expansion," *International Journal of Network Security*, vol. 18, no. 4, pp. 633–643, 2016.

[10] F. Li, Q. Mao, and C. C. Chang, "A reversible data hiding scheme based on IWT and the sudoku method," *International Journal of Network Security*, vol. 18, no. 3, pp. 410–419, 2016.

[11] J. J. Li, Y. H. Wu, C. F. Lee, C. C. Chang, "Generalized PVO-K embedding technique for reversible data hiding," *International Journal of Network Security*, vol. 20, no. 1, pp. 65-77, 2018.

[12] X. Li, B. Li, B. Yang, and T. Zeng, "General framework to histogram-shifting-based reversible data hiding," *IEEE Transactions on Image Processing*, vol. 22, no. 6, pp. 2181–2191, 2013.

[13] K. Ma, W. Zhang, X. Zhao, N. Yu, and F. Li, "Reversible data hiding in encrypted images by reserving room before encryption," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 3, pp. 553–562, 2013.

[14] S. D. Putra, S. Sutikno, Y. Kurniawan, A. S. Ahmad, "Design of an AES device as device under test in a DPA attack," *International Journal of Network Security*, vol. 20, no. 2, pp. 256-265, 2018.

[15] C. Qin, C. C. Chang, and T. J. Hsu, "Reversible data hiding scheme based on exploiting modification direction with two steganographic images," *Multimedia Tools and Applications*, vol. 74, no. 15, pp. 5861–5872, 2015.

[16] C. Qin and Y. C. Hu, "Reversible data hiding in vq index table with lossless coding and adaptive switching mechanism," *Signal Processing*, vol. 129, pp. 48–55, 2016.

[17] V. Sachnev, H. J. Kim, J. Nam, S. Suresh, and Y. Q. Shi, "Reversible watermarking algorithm using sorting and prediction," *IEEE Transactions Circuit Systems for Video Technology*, vol. 19, no. 7, pp. 989–999, 2009.

[18] E. Satir and H. Isik, "A compression-based text steganography method," *Journal of Systems and Software*, vol. 85, no. 10, pp. 2385–2394, 2012.

[19] J. Wang, J. Ni, and Y. Hu, "An efficient reversible data hiding scheme using prediction and optimal side information selection," *Journal of Visual Communication and Image Representation*, vol. 25, no. 6, pp. 1425–1431, 2014.

[20] S. Wang, J. Sang, X. Song, and X. Niu, "Least significant qubit (lsqb) information hiding algorithm for

quantum image," *Measurement*, vol. 73, pp. 352–359, 2015.
[21] Y. L. Wang, J. J. Shen, M. S. Hwang, "A survey of reversible data hiding for VQ-compressed images," *International Journal of Network Security*, vol. 20, no. 1, pp. 1-8, 2018.
[22] Y. L. Wang, J. J. Shen, M. S. Hwang, "An improved dual image-based reversible hiding technique using LSB matching", *International Journal of Network Security*, vol. 19, no. 5, pp. 858–862, 2017.
[23] Z. H. Wang, H. R. Yang, T. F. Cheng, and C. C. Chang, "A high-performance reversible data-hiding scheme for lzw codes," *The Journal of Systems and Software*, vol. 86, no. 11, pp. 2771–2778, 2013.
[24] Z. H. Wang, X. Zhuang, C. C. Chang, C. Qin, Y. Zhu, "Reversible data hiding based on geometric structure of pixel groups", *International Journal of Network Security*, vol. 18, no. 1, pp. 52–59, 2016.
[25] T. A. Welch, "A technique for high-performance data compression," *IEEE Computer*, vol. 17, no. 6, pp. 8–19, 1984.
[26] S. Zhang, T. Gao, L. Yang, "A reversible data hiding scheme based on histogram modification in integer DWT domain for BTC compressed images," *International Journal of Network Security*, vol. 18, no. 4, pp. 718–727, 2016.
[27] X. Zhang, "Reversible data hiding with optimal value transfer," *IEEE Transactions on Multimedia*, vol. 15, no. 2, pp. 316–325, 2013.
[28] J. Ziv and A. Lempel, "A universal algorithm for sequential data compression," *IEEE Transactions on Information Theory*, vol. IT-23, no. 3, pp. 337–343, 1977.
[29] J. Ziv and A. Lempel, "Compression of individual sequences via variable-rate coding," *IEEE Transactions on Information Theory*, vol. IT-24, no. 5, pp. 530–536, 1978.

Biography

Chin-Chen Chang received his Ph.D. degree in computer engineering from National Chiao Tung University. His current title is Chair Professor in Department of Information Engineering and Computer Science, Feng Chia University, from February 2005. He is currently a Fellow of IEEE and a Fellow of IEE, UK. And, since his early years of career development, he consecutively won Outstanding Talent in Information Sciences of the R. O. C., AceR Dragon Award of the Ten Most Outstanding Talents, Outstanding Scholar Award of the R. O. C., Outstanding Engineering Professor Award of the R. O. C., Distinguished Research Awards of National Science Council of the R. O. C., Top Fifteen Scholars in Systems and Software Engineering of the Journal of Systems and Software, and so on. On numerous occasions, he was invited to serve as Visiting Professor, Chair Professor, Honorary Professor, Honorary Director, Honorary Chairman, Distinguished Alumnus, Distinguished Researcher, Research Fellow by universities and research institutes. His current research interests include database design, computer cryptography, image compression, and data structures.

Ngoc-Tu Huynh received the BS degree in mathematics – informatics in 2006 from Danang University, Vietnam, and the MS degree in information engineering and computer science in 2010 from Feng Chia University. Since 2006, she has been a lecturer of Department of Computer Science, College of Information Technology, Danang University, Vietnam. She is currently pursuing her Ph.D in information engineering and computer science, Feng Chia University, Taichung, Taiwan. Her research interests include visual cryptography, watermarking, steganography and image processing.

Yu-Kai Wang received his MS degree from Department of Computer Science and Information Engineering, National Chung Cheng University, Chiayi, Taiwan. His current research interests include information hiding and image processing.

Yanjun Liu received her Ph.D. degree in 2010, in School of Computer Science and Technology from University of Science and Technology of China (USTC), Hefei, China. She has been an assistant professor serving in Anhui University in China since 2010. She currently serves as a senior research fellow in Feng Chia University in Taiwan. Her specialties include E-Business security and electronic imaging techniques.

IoT-based Efficient Tamper Detection Mechanism for Healthcare Application

Ahmed A. Elngar
(Corresponding author: Ahmed A. Elngar)
Faculty of Computer and Information, Beni-Suef University, Egypt
62511, Beni Suef, Salah Salem Str., Egypt
(Email: elngar_7@yahoo.co.uk)
(Received Apr. 28, 2017; revised and accepted Aug. 26 & Sept. 2, 2017)

Abstract

Security of networks is the most important challenge of the Internet of Things (IoT) that need smarter security mechanisms. Therefore, a tamper detection (TD) is an efficient security mechanism based on networks of IoT for healthcare applications, which used to deal with security violations. In this paper, a new TD mechanism based IoT for real data of healthcare application called (IOT-TD) model has been proposed. This paper effectively proposed (ANN-GA) tamper detection mechanism. Where, Genetic Algorithm (GA) is used to optimize weight and bias values of artificial neural networks (ANN) which lead to maximize the ANN detection accuracy, minimize the timing detection and efficiency energy saving. The experimental results showed that the tamper detection performance of (ANN-GA) is 98.51%. In addition, the proposed model showed that the (ANN-GA) enhances the timing to 0.03 sec which is important for real time of (IOT-TD) model healthcare application and the efficiency energy saving transmission is 1980 times better than full transmission. Also, the proposed Model relies on the certificate-based DTLS handshake protocol as it is the main security for (IoT-TD) model.

Keywords: Artificial Neural Network; Genetic Algorithm; Healthcare Applications; Internet of Things; Tamper Detection

1 Introduction

Nowadays, IoT is becoming one of the hottest research topics. IoT describes the future, where every day physical objects will connect to the Internet and be able to identify themselves to each others [19]. Hence, the IOT realizing smart environments such as: smart living, smart home, smart manufacturing, and smart healthcare applications. Due to the spread of chronic diseases and rising the cost of traditional healthcare application around the world; so it urgently demand transform the healthcare from hospital centered systems to remote personal healthcare systems [12]. Sensors, equipments and detectors around us have a significant impact on our everyday activities. Which It is becoming more pervasive for attempting to fulfill end users' need and provide easy of usability, specially in healthcare applications [14]. Therefore, one of the most important challenges of IoT based healthcare is a data security [9]. Where security is a major issue concerned of the most devices and their communications in nature [6]. These devices have a capability to send / receive data between each others using different communication protocols. The communication protocols must allowed low energy consumption and sufficient data security. Therefore, communication protocols are very important to secure the networks of IoT [17]. Hence, different types of communication protocols such as $CoAP$, $IEEE$ 802.15.4, $ZigBee$, $6LOWPAN$ and $Ethernet$ are used [13, 15]. The following Figure 1 shows some of security and management protocols for IoT.

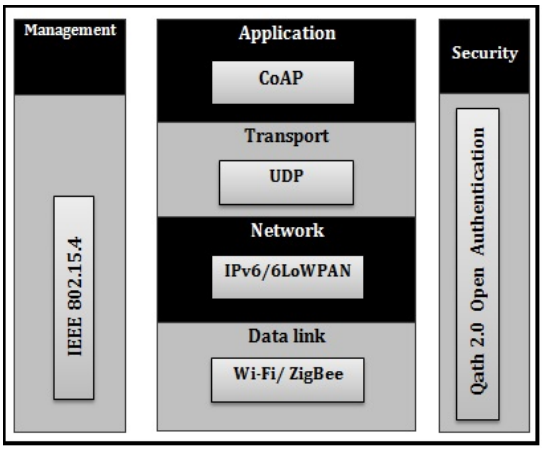

Figure 1: The framework protocols for IoT based healthcare

The main objective of internet of things based healthcare application is enhancing the interaction of device-to-

device, as well as the interaction of device-to-human via Internet. [14]. Although the collected data from harmless wearable sensors, such data is vulnerable to top privacy concerns. Where as the networks of IoT is secured based encryption and authentication mechanisms, but it so vulnerable against cyber-attacks [18]. So, this paper aims to secure data for digital communication in healthcare application [9, 21].

This paper, present a secure and efficient model for IoT-based healthcare using TD mechanism that monitor the malicious traffic in IOT-networks. Which it can defend the IOT-networks from intruders [1]. Intrusions are malicious activities that harmful to sensor nodes. Therefore, tTD can be used to inspect and investigate devices, user actions, and identifies the malicious activities for IOT-networks [11]. The TD works as IOT-networks observer, which avoids the damage of data by generating an alert before the attackers begin to attack. Also, TD can detect both external and internal attacks. An external attacks are launched by third party who is initiated by outside IOT-networks, whereas internal attacks are launched by nodes that belong to the IoT-networks.

There are mainly three components of TD: monitoring, detection and alarm [10]. The monitoring component monitors the network traffics, patterns and resources, detection is a core component of TD which detects the intrusions according to specified algorithm and alarm component raised an alarm if intrusion is detected [10].

In this paper, $IOT-TD$ model employs the ANN which have been used to solve classification problems. The performance of a ANN depends directly on the design of the hidden layers, and in the calculation of the weights that connect the different nodes [20]. In order to obtain a feasible results, the weights of ANN are calculated using a GA [4]. The GA is a meta-heuristic algorithm based on the concept of evolution processes. So, from all the search spaces of possible weights, the GA will generate new points of possible solutions to ANN.

Also, this paper employs $DTLS$ handshake protocol as it is a main security solution for the $IoT-TD$ model. To the best of our knowledge, $IoT-TD$ model is the first effort for proposing a secure and efficient model for IoT-based healthcare application using $(ANN-GA)\,TD$ mechanism. The elaboration of proposed model from the viewpoint of security as well as performance analysis is conducted. Also, the results reveal that the proposed $IOT-TD$ model based healthcare application increases the detection accuracy, speed up the detection time and the efficiency of energy saving compared to other well-known approaches.

Table 1 is the nomenclature of the paper.

The rest of this paper is organized as follows: Section 2 gives a literature survey. Section 3 presents a concept of the tamper detection mechanism. Section 4 gives the problem formulation. Section 5 introduces the proposed $IOT-TD$ model for healthcare application. Section 6 gives the implementation results and analysis. Finally, Section 7 contains the conclusion remarks.

Table 1: Nomenclature

IOT	Internet of Things
TD	Tamper Detection
ANN	Artificial Neural Network
GA	Genetic Algorithm
$DTLS$	Datagram Transport Layer Security
IDS	Intrusion Detection System
UDP	User Datagram Protocol
TCP	Transmission Control Protocol
$COAP$	Constrained Application Protocol

2 Literature Survey

The IOT performs the complicated functions in a simple way, which lead to structure more intelligent environments to make it a safe places for live in. Many researchers have been working on IoT-based healthcare applications and wireless sensor areas to provide the best mechanisms for data security. This section describes a variant contributions which are proposed in recent years.

Jun in [7] proposes event processing based IDS. Which solves the problem of real time of IDS in IoT-networks. Authors claimed a design of IDS approach based on the basis of Event Processing Model EPM. It is rule-based IDS in which rules are stored in Rule Pattern Repository and takes SQL and EPL of Epser as a reference. According to the results, this proposed consumed more CPU resources, less memory and took less processing time than traditional IDS for IoT-networks.

Alsadhan in [2] proposed an optimized IDS for IoT-networks using soft computing mechanisms. The objective of this proposed is increasing the performance of the system and identify each activities in a robust way. Where authors implemented the soft computing mechanisms like PCA, PSO and Greedy Search in IDS. In this proposed, the number of features are reduced with increasing of detection rate.

In [5], is proposed an IoT-based health monitoring system architecture which uses star-based $6LoWPAN$ motes that are integrated with an AFE device. The system uses a gateway which collects the data from the motes and transmits them to server, so they can provide many services for the connectivity conservation and the reinforcement of the system.

In [8], Kasinathan proposed IoT-networks based DoS detection IDS architecture within the EU FP7 project ebbits network framework. In this approach, IDS can listen or monitor $6LoWPAN$ traffic by using IDS probe. They used hybrid approach for placement of IDS. DoS protection manager is core component of proposed system which raised an alert by using information available on network manager component.

3 Tamper Detection Mechanism (TD)

TD mechanism is an ability of a device to sense with an active attempt which compromised the device or the data associated with that device. Hence, it enables this device to start appropriate defensive actions against any attacks [22]. The methods used for TD are typically designed as a suite of sensors each specialized for a single threat type. Also, TD mechanism enables the device to be aware of tampering and typically fall into one of three groups:

- **Switches**: to detect the opening of a device.
- **Sensors**: to detect environmental changes, voltage and power sensors to detect glitch attacks.
- **Circuitry**: to detect drilling or penetrating the device boundary.

The idea behind of a TD mechanism is to be a sensitive enough to detect the presence of a tangible threats. Also, it be able to distinguish from "false alarms" situations. There are several methods for applying TD mechanism; such as ANN and $C4.5$ methods [16].

4 Problem Formulation

Let z_t be set of patient sensors values acquired at time t.

$$z_t(a) = D_t[b_t](a),\ \forall\ a\ \in A \tag{1}$$

Where, D_t denoted as an operator transforming the original record b_t, and $a \in z^2$ indicates the sensors values that belonging to the regular record $A \subset z^2$. As far as there are no tampering attacks/events.

$$D_t[b_t](a) = b_t(a) + \eta_t(a),\ \forall\ a\ \in A \tag{2}$$

Where, η_t is a random variable accounting for record noise values, and b_t are acquired from the same sensor even though typically $b_t \neq b_t - 1$; because values of patient record are changed.

When, at time τ^* an external disturbance introduces a tampering, the record b_t is degraded by an unknown tamper attack and z_t becomes:

$$D_t[b_t](a) = \int b(e)h_t(a,j)d_j + \eta_t(a),\ \forall\ a\ \in A, t \succeq \tau^* \tag{3}$$

Where, $h_t(a,j)$ is the value-spread function at value $a \in A$.

The proposed $ANN - GA\ TD$ mechanism analyzes a sequence of $\{z_t, t = 1, ...,\ number\ of\ sensors\}$ to detect the time instant τ^* when tampering like 3 occur. We assume that T_0 tampering-free values are provided for training.

5 The Proposed $IOT - TD$ Model for Healthcare Application

This section contains the description for the proposed model $IOT - TD$. So, the main aim of the proposed model is to detect the tampering and ensure authentication of the biomedical information in IoT based healthcare application. Our proposed model designed to detect a huge types of attacks, which compromise the security of the biomedical information. These attacks such as: imitating and alteration. In which intruder can interfere and send an altered data that causes the tampering, bugging, and interruption of the biomedical information. The proposed model have a combination of IoT technologies, and communication protocols to design an efficient healthcare application.

5.1 Model Architecture

The architecture of the proposed $IOT - TD$ model shown in Figure 2 consists of three main modules: the digital environment module, local data processing environment module and remote doctor workstation module.

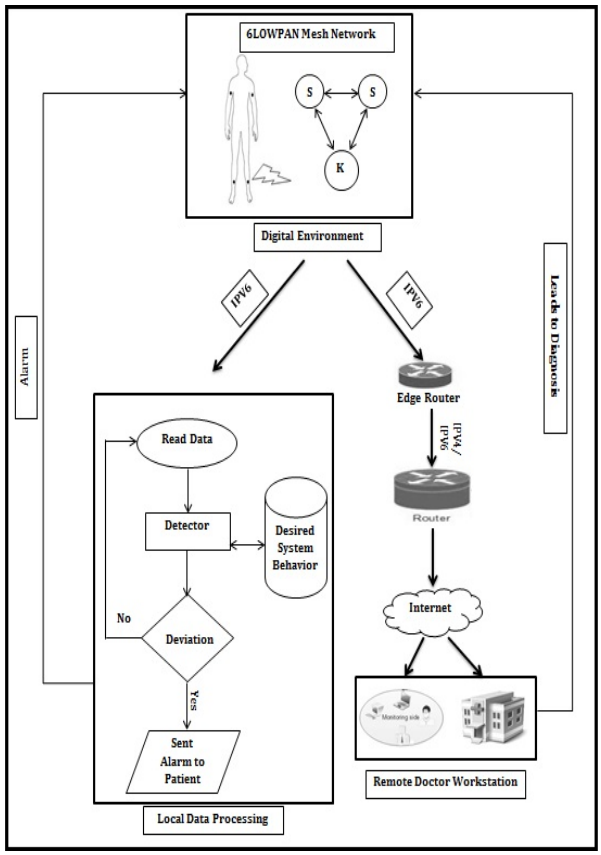

Figure 2: Structure $IOT - TD$ for healthcare model

5.1.1 Digital Environment Module

The digital environment module represented by the Arduino UNO Board and some medical sensors such as: Body temperature sensor, and pulse sensor, etc., which measure some variables, such as: blood pressure, temperature, and heart rate, etc.. Then the values will gathered to create a database as shown in Figure 3.

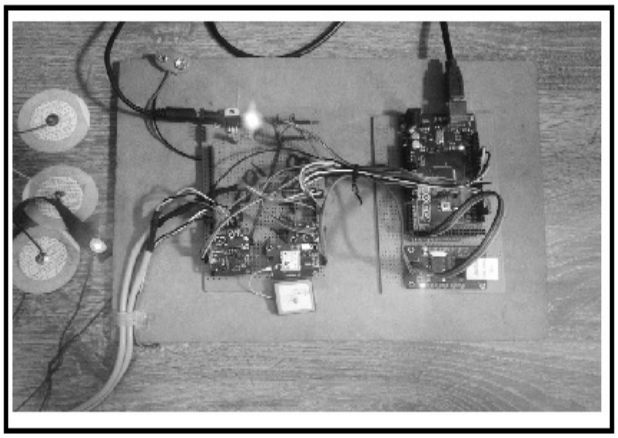

Figure 3: $IOT - TD$ hardware architecture

5.1.2 Local Data Processing Module

The local data processing module which consists of the TD component. Which it will receive the data (i.e., temperature, etc.) transmitted from user monitoring sensors in real-time. Then will be analyzing these values using $ANN - GA$ TD mechanism and compared it with the normal values of same patient. Where, ANN composed of digital nodes explained in subsection 5.1.1 module (equivalent to neurons of a human brain) which are interconnected by weighted links (equivalent to synapses between neurons) as shown in Figure 4.

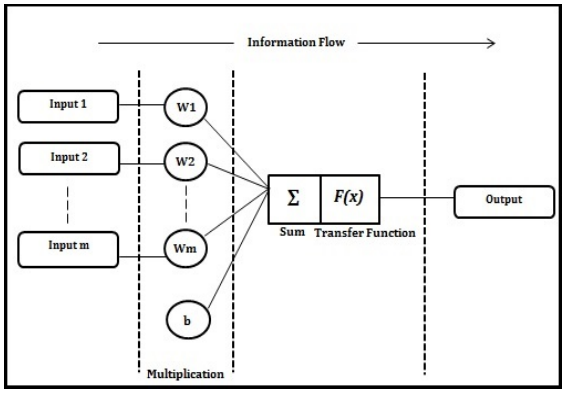

Figure 4: Design of ANN-GA

Hence, the outcome of the ANN is altered by changes of the weights of links. So, the weights of the ANN are calculated using GA approach. Such that from all the search hypothesis space of all the possible weights, the GA will generate new points of the possible solutions. Therefore, The mathematical description of $ANN - GA$ as follows:

$$Y_i = F(\Sigma_{i=0}^{m} W_i \times X_i + b) \quad (4)$$

Where,

- X_i: is input variable associated with each node.
- W_i: is connections' weights between inputs x_i estimated by GA.
- b: is the bias of the node.
- F: is the transfer function.
- Y_i: is the desired output of ANN.

If the output of $ANN - GA$ within the normal range; which indicates that the patient values observed is normal, then the data processing module will continue read the data from patient' medical sensors. While, if these values are not within the normal range; it indicating that the user's values monitor is abnormal, then the TD will send a reminder of warning to the patient.

5.1.3 Remote Doctor Workstation Module

Also, the remote doctor workstation module consists of the same TD mechanism component which observe the data for patient medical sensors in the real-time , then set up a personal database for each patient. Hence, TD will send an alert to patients when sense with abnormal values at any time then gives disposal proposals.

5.2 Model Topology

In this section, we focuses on the topology of the connections between the sensors and the actuators; which all nodes are connected to each other by links. In which it can used to communicate for the data and the signals transferring. The proposed model' network is consisting of $IEEE$802.15.4 which used in the physical layer. So, it can provide a wireless communications, where the bits of data after they have been converted into signals can be transmitted and received. Moreover, in the data link layer, 6$LoWPAN$ is used as (adaptation layer) where; the adjustment from $IPv6$ to $IEEE$ 802.15.4 is done. In the addressing and routing of data the Internet Protocol IP is used. So, we assign to every node a unique $IPv6$ address. The next layer is a transport layer, where UDP is used for the carriage of the data. The UDP is supplying lower latency and it is faster than TCP. The application layer is last layer, where it uses the $CoAP$. The proposed model' network also connects to other networks via $Wi - Fi$.

The notations used throughout this work which describes the proposed model shown in Table 2.

Suppose that the proposed network involves a set of sensors $\{S_1, S_2, ..., S_n\}$ and set of actuators

Table 2: Notations

Notation	Description
S_n	Total set of sensors
K_m	Total set of actuators
L_{node}	total set of links
S_{data}	Total sensor data
K_{sig}	Total actuator signals
P_{loss}	packet/signal loss
D_{rec}	Total data received
D_{sent}	Total data sent
N_{nf}	Number of nodes fails to transmit
A_{nodes}	Number of nodes succeeded

$\{K_1, K_2, ..., K_m\}$ which connected with each other. This situation is described by the two equations below:

$$(S_n + K_m) - N_{nf} = A_{nodes} \quad (5)$$

$$(S_{data} + K_{sig}) - P_{loss} = D_{sent} \quad (6)$$

The proposed model used the duplex mesh communications, which means if nodes failed to transmit data/signal; it doesn't affect the transmission from other nodes. This is shown below:

$$[(S_n + K_m) \times \frac{(S_n + K_m - 1)}{2}] = L_{nodes} \quad (7)$$

6 Implementation Results and Security Analysis

The proposed $IOT - TD$ model is evaluated for the 36 normal and 36 abnormal patients. All experiments have been performed using Intel Core i3 2.13 GHz processor with 2 GB of RAM. The experiments have been implemented using Java language environment with C, Eclipse, Cloud Interface Linux Operating System 64-bit, and Windows Operating System 64-bit.

6.1 Performance Measurements

The detection effectiveness of the proposed $IOT - TD$ model is measured in term of TP Rate, FP Rate and $F - measure$; which are calculated based on the confusion matrix (CM). The CM is square matrix where columns correspond to the predicted class, whereas, rows correspond to the real classes. Table 3 presents the CM, which shows the four possible prediction outcomes. Here,

True negatives (TN): indicates the number of normal events are successfully labeled as normal.

False positives (FP): refer to the number of normal events being predicted as abnormal.

False negatives (FN): The number of abnormal events are incorrectly predicted as normal.

True positives (TP): The number of abnormal events are correctly predicted as abnormal.

$$TPRate = \frac{TP}{TP + FN}$$

$$FPRate = \frac{FP}{FP + TN}$$

$$F - measure = \frac{2 * TP}{(2 * TP) + FP + FN}.$$

Table 3: Confusion matrix

	Predicted Class	
Real Class	Normal	Abnormal
Normal	TN	FP
Abnormal	FN	TP

6.2 Experiment Results

The detection performance measurements by $ANN - GA$ TD are shown in Tables 4 and 5. Table 4 shows the accuracy measurements achieved for $C4.5$ method. While, Table 5 gives the accuracy measurements of $ANN - GA$ TD for the proposed $IOT - TD$ model.

Table 4: C4.5 tamper detection

Class name	TP Rate	FP Rate	F-Measure
Normal	0.793	0.267	0.791
Abnormal	0.733	0.207	0.736

Table 5: $ANN - GA$ tamper detection

Class name	TP Rate	FP Rate	F-Measure
Normal	1	0.033	0.987
Abnormal	0.967	0.0	0.983

From Tables 4 and 5, it is clear that the detection accuracy achieved using $ANN - GA$ as TD method is better than using $C4.5$.

Table 6 compares the TD accuracy and timing speed of $C4.5$ and proposed $ANN - GA$. Table 6 illustrate that the propose gives better detection performance (98.51%) than the $C4.5$.

Also, the proposed enhances the timing speed to 0.03 sec which is important for real time $IOT - TD$ model in healthcare application.

The performance comparison of the proposed model over two other approaches based on several features are listed in Table 7.

6.3 Energy-Saving Transmission Efficiency Analysis

The nodes in the $IoT - TD$ base healthcare application are usually battery energy hence; energy is a scarce re-

Table 6: Testing accuracy and timing comparison

System	Test accuracy	Model building Time
IOT-C4.5	76.66%	0.06 sec.
Proposed IOT-ANN-GA	98.51%	0.03 sec.

Table 7: Comparative analysis between the proposed model and other approaches

Feature	[3]	[5]	Proposed
IoT-based	✓	✓	✓
TD-based	×	×	✓
Coap	×	×	✓
Topology	Star	Mesh	Full Duplex Mesh
Security	Basic	AES Block	AES-128, DTLS, WPA2
Energy Efficiency	×	✓	✓
802.15.4	×	✓	✓
Scalability	High	Low	High
Adaptability	High	High	High

source. Here, this paper compared the transmission efficiency every 5-minutes according to different abnormal patient ratios with "full transmission" and "energy-saving transmission". In the energy-saving transmission mode, data is transmitted to the remote doctor workstation module in a 32- byte package at a 5-minute interval. In the full transmission mode, data is transmitted continuously at a 1-second interval. In total, the size of continuous data transmitted over 5-minutes is 57,652 bytes ($12\ bits * 128/sec * 300s + 20\ bytes\ of\ package\ field + 32\ bytes\ of\ ANN - GA\ result$). A twenty-four hour of data-set is emulated using 288 every five-minute data sets. The normal ratio is defined as the percentage of normal $ANN-GA$ results analyzed by the multi-pattern abnormal disease matching in the remote doctor workstation module. Once the $ANN - GA$ parameters are transmitted to the remote doctor workstation module, they are analyzed to decide whether to transmit the raw data. For instance, if the normal ratio is 80%, the twenty-four hour energy-saving transmission transmits $288 * 32\ bytes + 57652\ bytes * 288 * (1 - 80\%) = 3.176\ Mb$, and the efficacy is $15.84\ Mb/3.176Mb = 4.99$. Suppose that 100% normal patterns can be detected from the patients; the transmission efficiency is then 1980 times. It is useful when analysis a huge amount of data such in healthcare application as shown in Table 8.

6.4 Security Analysis

The security of the proposed model "$IoT - TD$" is an important issue for healthcare application. Where as the healthcare information is very sensitive and the internet will never be safe. So, this section is going to discuss the security architecture of the proposed model in each layer.

This paper propose $(ANN - GA)\ TD$ efficient mechanism at both local data processing and remote doctor workstation. Where, GA with ANN will produce a hybrid neural network. So, the weights of ANN are calculated using GA algorithm. From all possible weights of search space, the GA will generate new points of possible solutions. Which implies that, it possible to optimize the ANN by modifying the structure of weights calculation. Hence, $(ANN - GA)\ TD$ mechanism leads to maximize the TD accuracy, minimize the detection timing and efficiency energy saving.

Also, $6LoWPAN$ in data link layer security which is responsible for the encryption and authentication of the links. $6LOWPAN$ provides secure data packets delivery. besides, in the transport layer the proposed model use UDP over $DTLS$ mechanisms that could also be used for $CoAP$ security, in order to save the communications between the objects. Furthermore, the $IEEE$ 802.15.4 standard has many security protocols, such as the $Wi-Fi$ Protected Access $WPA2$ which provides data integrity, confidentiality and authentication.

7 Conclusions

This paper proposed a new $IOT - TD$ model which employs $(ANN - GA)\ TD$ mechanism for secure the sensitive information in healthcare applications. Therefore, $ANN - GA$ can be used to satisfy the security requirements of IoT-networks environment. According to the primary and earlier experiments, the proposed $ANN - GA$ mechanism achieved 98.51% TD rate, which can be considered as the best tamper detection rate compared with the $C4.5$ algorithm which achieved 76.66%. Also, the proposed $ANN-GA$ mechanism enhances the timing detection to 0.03 sec compared with the $C4.5$ algorithm which achieved 0.06 sec and efficiency energy saving which is important for the real-time $IOT - TD$ model of healthcare applications.

Acknowledgments

The author gratefully acknowledge the editor and the anonymous reviewers for their valuable comments.

References

[1] R. Aarthi, A. R. Renold, "Coap based acute parking lot monitoring system using sensor networks," *ICTACT Journal On Communication Technology: Special Issue on Advances In Wireless Sensor Networks*, vol. 5, no. 2, pp. 923–928, 2014.

[2] A. Alsadhan, N. Khan, "A proposed optimized and efficient intrusion detection system for wireless sensor network," *International Journal of Electrical, Computer, Energetic, Electronic and Communication Engineering*, vol. 7, no. 12, 2013.

Table 8: Energy-saving transmission efficiency results

Normal ration	0%	20%	40%	60%	80%	100%
Full transmission (Mb/24-hour)	15.84	15.84	15.84	15.84	15.84	15.84
Energy-saving transmission (Mb/24-hour)	15.84	12.67	9.54	6.33	3.176	0.008.
Efficiency (times)	1	1.25	1.66	2.50	4.99	1980

[3] S. R. Anurag, A. M. Rahmani, T. Westerlund, G. Yang, P. Liljeberg, H. Tenhunen, "Pervasive health monitoring based on internet of things: Two case studies," in *EAI 4th International Conference omn Wireless Mobile Communication and Healthcare (Mobihealth'14)*, pp. 275-278, 2014.

[4] A. A. Elngar, D. A. El A. Mohamed, F. M. Ghaleb, "A fast accurate network intrusion detection system," *International Journal of Computer Science and Information Security*, vol. 10, no. 9, Sept. 2012.

[5] T. N. Gia, A. M. Rahmani, T. Westerlund, P. Liljeberg, H. Tenhunen, "Fault tolerant and scalable IoT-based architecture for health monitoring," in *IEEE Sensors Applications Symposium (SAS'15)*, pp. 1-6, 2015.

[6] J. He, C. Hu,, and X. Wang, "A smart device enabled system for autonomous fall detection and alert," *International Journal of Distributed Sensor Networks*, vol. 1, 2016.

[7] C. Jun, C. Chi, "Design of complex event-processing IDS in internet of things," in *IEEE Sixth International Conference on Measuring Technology and Mechatronics Automation*, 2014.

[8] P. Kasinathan, *et al.*, "Denial-of-service detection in 6LoWPAN based internet of things," in *IEEE 9th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob'13)*, 2013.

[9] H. Kayarkar, "Classification of various security techniques in databases and their comparative analysis," *ACTA Technica Corviniensis*, vol. 5, pp. 135-138, 2012.

[10] M. Kovatsch, "CoAP for the web of things: From tiny resource-constrained devices to the web browser," in *Proceedings of ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication*, pp. 1495-1504, 2013.

[11] Y. Ma, "NFC communications-based mutual authentication scheme for the internet of things," *International Journal of Network Security*, vol. 19, no. 4, pp. 631-638, 2017.

[12] S. R. Maynard, H. Thapliyal, and A. Caban-Holt, "Smart home system for patients with mild cognitive impairment," in *Proceedings of the 2015 International Conference on Computational Science and Computational Intelligence*, pp. 738-742, 2015.

[13] A. Mayzaud, R. Badonnel, and I. Chrisment, "A taxonomy of attacks in RPL-based internet of things," *International Journal of Network Security*, vol. 18, no. 3, pp. 459-473, May 2016.

[14] S. P. Mohanty, U. Choppali, and E. Kougianos, "Everything you wanted to know about smart cities," in *IEEE Consumer Electronics Magazine*, vol. 5, no. 3, pp. 60-70, 2016.

[15] P. Pongle, G. Chavan, "A survey: Attacks on RPL and 6LoWPAN in IoT," in *IEEE International Conference on Pervasive Computing (ICPC'15)*, 2015.

[16] J. R. Quinlan, *C4.5 Programs for Machine Learning*, Morgan Kaufmann San Mateo Ca, 1993.

[17] E. Raptopoulou, *CoAP Enabled Sensors for the Internet of Things*, Department of Applied Informatics and Multimedia, Technological Educational Institute of Crete, 2014.

[18] W. L. Tai, Y. F. Chang, "Comments on a secure authentication scheme for IoT and cloud servers," *International Journal of Network Security*, vol. 19, no. 4, pp. 648-651, 2017.

[19] L.M.R. Tarouco, L.M. Bertholdo, L.Z. Granville, and L.M.R. Arbiza, "Internet of things in healthcare: Interoperatibility and security issues," in *IEEE International Conference on Communications*, pp. 621-6125, 2012.

[20] G. Wang, J. Hao, J. Ma and L. Huang , "A new approach to intrusion detection using artificial neural networks and fuzzy clustering," *Expert Systems with Applications*, vol. 37, pp. 6225-6232, 2010.

[21] D. Xu, Z. Wu, Z. Wu, Q. Zhang, L. Qin, J. Zhou, "Internet of things: Hotspot-based discovery service architecture with security mechanism," *International Journal of Network Security*, vol. 17, no. 2, pp. 208-216, 2015.

[22] Y. Ye, Y. He, Y. Wang, "SHVC, the scalable extensions of HEVC and its applications," *ZTE Communications*, vol. 14, no. 1, 2016.

Biography

Ahmed A. Elngar graduated with a B.Sc. in computer Science from computer science Department, Al-Azhar University, Master of computer science in Intrusion Detection System (IDS) from Ain Shanm university. He obtains his P.hD at computer science Department, Al-Azhar University. Also he is a member in Egyptian Mathematical Society (EMS) and International Rough Set Society(IRSS). Now he is a member of Scientific Research Group in Egypt (SRGE). He is Asst. Prof. of computer scienceat Faculty of Computer and Information, Beni-Suef University.

Energy Aware and Trust Based Cluster Head Selection for Ad-hoc Sensor Networks

Zhe Wei[1] and Shuyan Yu[2]
(Corresponding author: Zhe Wei)
School of Computer Science, Civil Aviation Flight University of China[1]
46 Nanchang Rd 4th Section, Guanghan 618300, China
(Email: findwei@foxmail.com)
College of Management and Information, Zhejiang Post and Telecommunication College[2]
(Received May 1, 2017; revised and accepted July 5 & Aug. 20, 2017)

Abstract

Clustering provides an efficient management method and energy balancing scheme for ad-hoc sensor networks. Cluster head is the most important role in a cluster and it acts as a local data coordinator and maintains cluster information. Once malicious nodes or lower energy nodes are selected as cluster heads, the system would be greatly affected. Thus selection of trusted cluster heads with proper residual energy becomes critical for the overall network performance. In this research, we propose an energy aware and trust based cluster head selection method for ad-hoc sensor networks. The proposed method relies on an effective distributed trust model for cluster head selection and it also considers the residual energy in the selection process. Simulations show that more trusted nodes with proper residual energy are selected as cluster heads, which in turn provides a higher packet delivery ratio from the cluster member nodes to the base station and a better balanced energy consumption of the network.

Keywords: Ad-hoc Sensor Network; Cluster Head Selection; Energy Aware; Trust

1 Introduction

Ad-hoc sensor networks normally consist of spatially distributed devices using wireless sensor nodes to collaboratively collect, process, and transmit physical or environmental parameters [21]. In practice, individual sensor nodes collet data of interest, process them locally for certain purposes, and send the processed data to the base station directly or indirectly with the help of intermediate nodes [15]. Autonomy is one of the most important characteristics of ad-hoc sensor networks where each node is self-configured without the centralized administration. Further, ad-hoc sensor networks are instant in that no pre-established infrastructure is needed for the network deployment, so they have been used for a variety of applications such as security surveillance, intrusion detection, disaster management, animal tracking and so on.

To ensure the full functioning of the various applications, security is an important issue to be addressed for the autonomous and unattended ad-hoc sensor networks [14]. This is because sensor nodes are vulnerable to attacks such as selective forwarding attack, Sybil attack and wormhole attack. Most security solutions like cryptography are software based and they are designed to mainly deal with the outside attacks for traditional networks, but such soft security is hard to be implemented in sensor nodes to counter the attacks especially from inside malicious nodes. To solve this problem, trusted computing [4] has been adopted to tackle the malicious nodes within the network. Trust is essentially a stimulating mechanism for nodes' cooperation and its computing is based on a node's action or behavior such as delivering or dropping data packets upon request. Under trust mechanism, higher trust nodes will receive more services from its peers and less trust nodes get fewer or no services from the others. Sensor nodes are also featured with limited power supply and they are usually disposed when their batteries are exhausted. Clustering techniques [22] provide an efficient energy balancing method for the sensor network. In a clustering scheme, all the nodes in the network are virtually partitioned into sub networks called clusters. In each cluster, member nodes have one or more elected Cluster Heads (CHs). CH is the most important element in a cluster and it acts as a local coordinator for data transmission within the cluster and maintains the cluster members and topology information [20].

However, once the malicious nodes are selected as CHs, the system performance would be greatly affected since all the member nodes depend on CHs for packet transmission to their respective destinations. In addition, some CHs with high trust value will be repetitively selected, which drains their energy faster. In this context, selection of trusted CHs with proper residual energy becomes critical for the overall network performance. In this research, we

propose an Energy Aware and Trust (EAT) based CH selection method for ad-hoc sensor networks. The proposed method relies on an effective distributed trust model and it also considers the residual energy in the process of CH selection.

The rest of this article is organized as follows. Section 2 discusses the related work about the classical node clustering algorithm and the trust computing, Section 3 describes the proposed CH election method, simulation tests are carried out in Section 4, and the conclusions and the future research are discussed in Section 5.

2 Related Work

LEACH (Low Energy Adaptive Clustering Hierarchy protocol) [9] is a classical hierarchical clustering algorithm for wireless sensor networks (WSNs) and many clustering algorithms such as C-LEACH [18], P-LEACH [12], A-LEACH [2], H-LEACH [3], N-LEACH [16], K-LEACH [24], E-LEACH [26], T-LEACH [11], W-LEACH [23], V-LEACH [1], LEACH-FL [8], and LEACH-ERE [13] have derived from LEACH by either modifying the threshold criteria or optimizing the algorithm parameters.

In LEACH, clustering is based on the signal strength and CHs are randomly selected. The operation process for LEACH is split into rounds and each round consists of the setup phase and the steady phase. In the setup phase, each sensor node that has not been selected as CH chooses a random number between 0 and 1 to decide whether it will become a CH or not for the current round. The decision of a node to be a CH is independent of other nodes. If the number of a sensor node is less than the predefined threshold value $T(n)$, this sensor node will convert from an ordinary node into a CH for the current round. The threshold $T(n)$ is defined by

$$T(n) = \begin{cases} \frac{p}{1-p*(r \bmod \frac{1}{p})} & if\ n \in G \\ 0 & otherwise \end{cases}$$

Where r denotes the current round, G represents the set of nodes that have not been selected as CHs in the last $\frac{1}{p}$ rounds, p is a pre-determined percentage of CHs in the round, and n is the number of nodes in the network. After a node is elected, it informs the member nodes about its election as CH through advertisement packet. Upon receiving the advertisement packet, the member node sends its ID in the join packet to the CH. In the steady phase, member nodes collect and transmit data to their CHs which aggregate the received data and forward these data to the BS. After a given period of time, the algorithm returns to the setup phase and enters into a new round of CH selection. LEACH balances the energy consumption of cluster members by rotating the CH, but the drawback of LEACH is that the CH selection is random without considering node's residual energy.

Trust and reputation mechanisms [7, 19, 25] have been gradually studied by researchers. In the trust computing, trust is defined as the degree of beliefs about the behaviors of others and it can help to identify the malicious nodes, predict the future behavior of a node, and select trustworthy nodes under certain trust strategies. The basis of trust mechanism is that its calculation is either directly based on the historical behaviors of participating nodes or indirectly based on the references from other nodes. Among these models, Bayesian theory that attempts to discover the behavior patterns through historical actions fundamentally complies with the procedure of trust evaluation. Bayesian based trust computing first calculates the prior probability of an event, then applies the prior probability into the binomial distribution, and finally modifies or updates such probability by using a posterior inference according to the relevant evidences [25].

RFSN (Reputation based Framework for high integrity Sensor Networks) [7] is a representative application of Bayesian theory for the trust computing. In RFSN, each sensor holds trust metrics representing past behaviors of other nodes in order to predict these nodes' future behaviors. According to the trust metrics built for other nodes by the behavior monitoring, a sensor node can rate them as *positive* or *negative* and evaluate the trustworthiness of these nodes. RFSN uses a completely decentralized management manner and can run on each sensor node, the latter of which in RFSN only interacts with nodes within its wireless communication range and thus only maintains the reputation of nodes in its vicinity. In RFSN, a transaction is defined as two nodes making an exchange of information or participating in a collaborative process. After each transaction, one partner will rate the other as *cooperative* or not. Let Θ represent the probability that a certain node will cooperate when asked to exchange information in RFSN, and a prior distribution that reflects the probability that a node would cooperate with another one is defined by

$$P(\Theta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\Theta^{\alpha-1}(1-\Theta)^{\beta-1}$$

where $0 \le \Theta \le 1, \alpha \ge 0$, and $\beta \ge 0$. Θ can be used as the success probability in Bernoulli observations. For example, let $T \in [0,1]$ be the node i's rating for node j in one transaction, then

$$P(T|\Theta) = \Theta^T(1-\Theta)^{1-T}$$

After the transaction the posterior of Θ is:

$$P(\Theta|T) = \frac{P(T|\Theta)P(\Theta)}{\int P(T|\Theta)P(\Theta)d\Theta} \sim Beta(\alpha+T, \beta+1-T)$$

The mathematical expectation of Θ is:

$$E(\Theta) = \frac{\alpha+T}{\alpha+T+\beta+1-T} \quad (1)$$

In Equation (1), $E(\Theta)$ can be regarded as the trust value of a node, and the shape parameter (α, β) can be interpreted as the observed number of positive outcomes

(cooperation) and the observed number of negative outcomes (non-cooperation) in one transaction respectively. According to Equation (1), the limitation of the trust calculation in RFSN is that trust has to be computed after each transaction so as to update the trust values.

3 EAT Based CH Selection

3.1 Energy Model

In the proposed method, the information of remaining energy about each sensor node is exchanged periodically among one-hop neighbors and based on [10], when k-bit data packet is transmitted within distance d in ad-hoc sensor networks, the transmitter energy consumption $E_t(k,d)$ is defined by

$$E_t(k,d) = \begin{cases} kE_{elec} + k\varepsilon_{FS}d^2 & d < d_0 \\ kE_{elec} + k\varepsilon_{MP}d^4 & d \geq d_0 \end{cases}$$

where E_{elec} is the electronics energy such as signal coding and spreading, $\varepsilon_{FS}d^2$ and $\varepsilon_{MP}d^4$ are the amplifier energy in the free space fading channel with d^2 power loss and multi path fading channel with d^4 power loss respectively. If the distance d is less than the predefined threshold d_0, the power loss can be modeled as the free space model, or else, if d is greater than or equal to d_0, the power loss is modeled as the multi path model. After receiving this k-bit data packet, the receiver energy consumption $E_r(k)$ is defined by

$$E_r(k) = kE_{elec}$$

Thus the remaining energy of a certain node i is

$$E_{remaining} = E_{initial} - E_t(k,d) - Er(k)$$

In practice, the free space model is used and a threshold $E_{threshold} = 0.00005J$ is set so as to check whether a node has enough remaining energy to work as a CH.

3.2 Trust Calculation

Unlike the binomial distribution based trust method in RFSN, *negative* binomial distribution based method is more flexible and with more applications. In our previous work [25], we proposed a negative binomial distribution based trust that can well be applied in WSNs. The definition is as follows.

$$E(\Theta) = \frac{\alpha + r}{\alpha + \beta + r + s} \tag{2}$$

where r and s are the corresponding increments, $E(\Theta)$ and the shape parameter $(\alpha,\ \beta)$ has the similar meanings to those in Equation (1), but the increment in Equation (2) can be 2 or more and neighboring nodes need not update the trust of the monitored node every time. For example, many newly designed MAC protocols such as SW-MAC [17] and ASS-MAC [5] support sleep mode in sensor networks where sleep-wake scheduling is set to achieve the energy efficiency in communications and the energy consumption can be significantly reduced by putting nodes into sleep state when their services are not needed for certain period of time [17]. It means that nodes in sleep mode cannot respond to the requests from others. Assume node j makes a series of requests within a fixed period of time ΔT from node i and i works alternatively between sleep and wake mode during the requests. If j receives $r(\Delta T)$ positive outcomes and $s(\Delta T)$ negative outcomes from i within ΔT, then the trust value of i maintained by j is defined by

$$E_{i,j}(\Theta) = \frac{\alpha_{i,j} + r_{i,j}(\Delta T)}{\alpha_{i,j} + \beta_{i,j} + r_{i,j}(\Delta T) + s_{i,j}(\Delta T)}$$

Further, trust from the third parties should be added as indirect references. According to the *D-S* belief theory [6], suppose j receive the trust about i from h. Let (α_i^h, β_i^h) denote such indirect trust and j has the past trust values about i and h that are denoted by (α_i, β_i) and (α_h, β_h) respectively. Thus combined with indirect trust from h, the shape parameters are redefined by

$$\alpha_i^{'} = \alpha_i + \frac{2\alpha_h\alpha_i^h}{(\beta_h + 2) + (\alpha_i^h + \beta_i^h + 2) + 2\alpha_h}$$

$$\beta_i^{'} = \beta_i + \frac{2\alpha_h\beta_i^h}{(\beta_h + 2) + (\alpha_i^h + \beta_i^h + 2) + 2\alpha_h}$$

where $\alpha_i^{'}$ and $\beta_i^{'}$ are the direct-indirect integrated trust parameters about node i respectively. The proposed CH selection algorithm is presented in Table 1 and in case no qualified CH is selected in the current round, a random node within the cluster is designated as the acting CH.

Table 1: Algorithm of EAT based CH selection

Algorithm

```
//Use LEACH to form clusters in the sensor network.
//In each round, the CH is selected as follows.
Input: Cluster members
Output: CH of a cluster
Begin
loop1: for(i=1;i<=ClusterNumber;i++)
          if (MaxTrust < Trust[i]){
             MaxTrust=Trust[i];
              j=i;}
        if (E_remaining(j) >= E_threshold)
             Node j is selected as CH;
         else {
           ClusterNumber--;
           remove Node j from the ClusterNumber;
           goto loop1; }
End
```

4 Simulations

The problem that this study deals with is to select CHs with descent trust value and proper residual energy so as to effectively prevent malicious nodes from becoming CHs and efficiently balance the energy consumption of the network. Although LEACH can balance the energy consumption of cluster members by rotating the CH, the drawback is that the CH selection is random without considering node's residual energy and trust values. Thus both malicious nodes and nodes with low residual energy can become CHs in LEACH, which could deteriorate the system performance. In this section, to test the performance of the proposed method, NS-2 is used for the simulation and LEACH is selected for comparison.

4.1 Settings

Assume that 500 sensor nodes are randomly deployed in a $400m$*$400m$ square region. The BS is set at the center of the area and all the CHs can directly communicate with the BS. When requested by the BS, all cluster members send fixed 200-sized data packets containing node ID and meaningful information directly to the CH through which these packets are transmitted to the BS. Suppose that there are 10% evenly deployed malicious nodes, and when working as cluster members they selectively send void data packets to the CH in order to drain the network energy, and when these malicious nodes are selected as CHs, they randomly drop some or all the data packets sent by the cluster members. It is also assumed that new round of CH selection is carried out in every 20 requests from the BS. Other settings are as follows: the initial reputation of each node is 0.5; 802.11 protocol with TDMA and sleep mode is implemented in MAC; $EI = 0.5J$, $E_{elec} = 50nJ/bit$, $d = 1m$, and $\varepsilon_{fs}d^2 = 10pJ/bit/m^2$; the channel bandwidth is set to 1 Mb/s; sensor nodes are capable of bidirectional communication on every link and they work in the promiscuous mode so that nodes can over hear the ongoing packets from its neighbors.

4.2 Test 1

Under ideal conditions, CHs are the trusted entities for packet transmission and data packets from the cluster member nodes should be completely transmitted by the corresponding CHs to the BS. But due to the existence of malicious nodes or malicious CHs, not all the CHs are trusted and some packets may not be delivered by malicious CHs and eventually cannot reach the BS. In this part, CH average trust value and the packet delivery ratio are tested and results are presented in Figure 1 and Figure 2 respectively.

In LEACH, malicious nodes can be selected as CHs without any prevention, thus the CH average trust in LEACH fluctuates around 0.52 during the queries as can be noticed in Figure 1. It indicates that some selected CHs have lower trust values than 0.5 and these CHs could present malicious behaviors such as dropping the packets sent from the member nodes in the current scenario. While in EAT, the CH average trust increases steadily and reach about 0.61 on the 100th query meaning that more and more trusted nodes are selected as CHs.

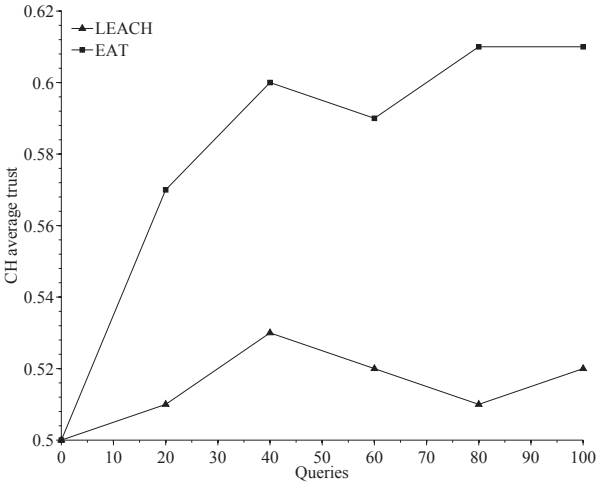

Figure 1: CH average trust

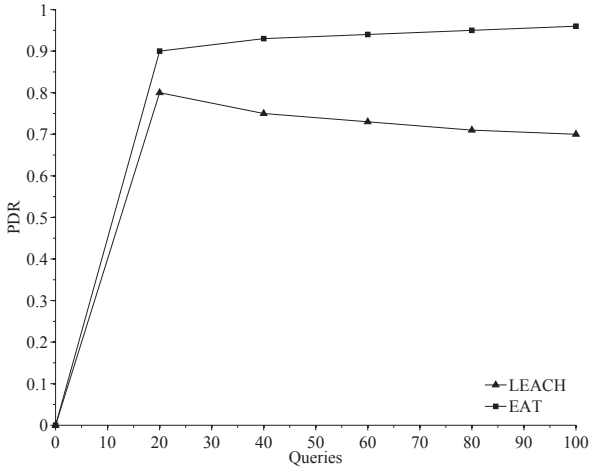

Figure 2: Packet delivery ratio

Definition 1. *Packet deliver ratio, or PDR is the number of packets received by the BS to the number of packets sent by the member sensor nodes.*

In Figure 2, as the query number increases, the PDR in LEACH reaches its maximum value around 0.8 on the 20th query and then drops constantly and reaches around 0.7 on the 100th query. This is because in LEACH malicious nodes can be selected as CHs without any precautions. Once malicious nodes become CHs, they can do considerable damage to the network such as dropping some or all the received packets from the cluster member nodes in this case. On contrast, EAT can maintain

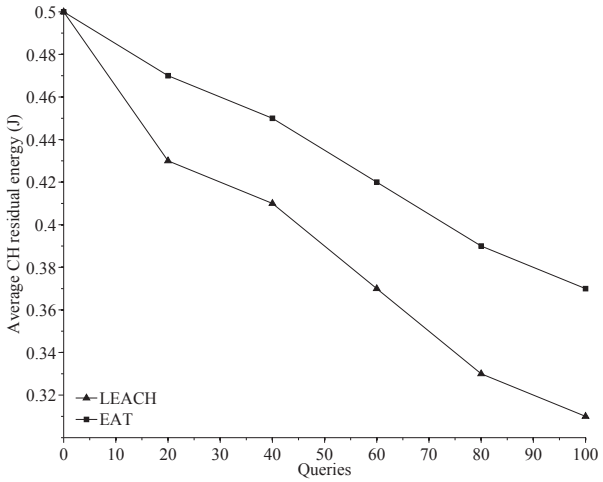

Figure 3: Average CH residual energy

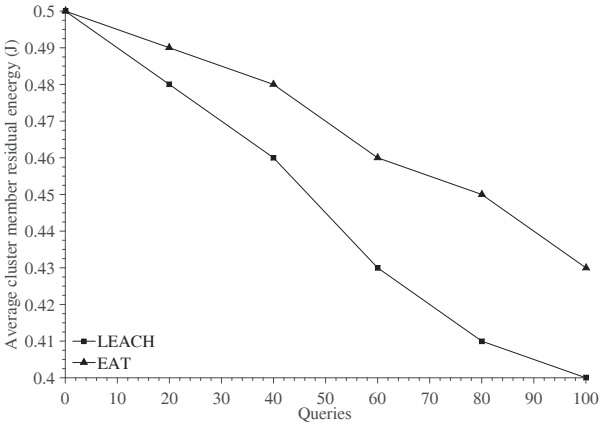

Figure 4: Average cluster member residual energy

a higher PDR during the queries by incorporating trust mechanism and avoiding malicious nodes becoming CHs and hence can obtain PDR about 0.95 on the 100th query. Compared to LEACH, EAT has an average of 19.8% improvement in PDR.

4.3 Test 2

In this part, the average CH residual energy and the average cluster member residual energy are tested and results are shown in Figure 3 and Figure 4 respectively.

In Figure 3, the average CH residual energy in both methods declines as the query goes on, but EAT always maintain a higher average residual energy than LEACH, e.g., about 0.46J and 0.41J respectively on the 40th query. It indicates that by considering the node residual energy during the CH selection, EAT can choose the potential candidate CH with more residual energy. By contrast, LEACH rotates the CH and randomly selects the CH without taking the remaining energy into consideration resulting in lower average CH residual energy than EAT. It can also be found in Figure 1 and Figure 3 that EAT can not only maintain a higher CH trust value but also keep a higher average CH residual energy than LEACH.

The similar result can be found in Figure 4 that in EAT, the average cluster member residual energy is always higher than LEACH, e.g., about 0.46J and 0.43J respectively on the 60th query. It indicates that EAT can better balance the energy consumption within the clusters. Compared to LEACH, EAT has an average of 4.2% improvement regarding the average CH residual energy and 2.2% improvement on the average cluster member residual energy.

5 Conclusions

Improper CH selection could severely degrade the performance of the clustered ad-hoc sensor networks especially when malicious nodes or low-energy nodes are selected and become CHs. In this study, with the help of trust mechanism and by considering nodes' residual energy, an energy aware and trusted CH selection method is proposed aiming to select CHs with descent trust value and proper residual energy. Simulation tests have confirmed that the proposed method can effectively prevent malicious nodes from becoming CHs and efficiently balance the energy consumption of the network. But due to the random behavior of malicious nodes, some malicious nodes can still be elected as CHs in the proposed method as can be seen the PDR test. Thus how to further enhance the PDR and better spot the malicious nodes with more different random misbehaving patterns will be our future work.

Acknowledgment

This work is partially supported by the National Science Foundation of China under grant No. 61073177, the Scientific Research Program of CAFUC under grant Nos. J2018-03, F2017KF02, J2013-41, and Q2014-053. The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which has improved the presentation.

References

[1] A. Ahlawat and V. Malik, "An extended vice-cluster selection approach to improve LEACH protocol in WSN," *International Conference on Advanced Computing & Communication Technologies*, pp. 236-240, 2013.

[2] M. Ali, T. Dey and R. Biswas, "ALEACH: Advanced LEACH routing protocol for wireless microsensor networks," in *Proceedings of the International Conference on Electrical and Computer Engineering*, pp. 909-914, 2008.

[3] A. Azim and M. Islam, "Hybrid LEACH: A relay node based low energy adaptive clustering hierarchy for wireless sensor networks," in *Proceedings of IEEE Malaysia International Conference on Communication*, pp. 911-916, 2009.

[4] D. Challener, K. Yoder, and R. Catherman, "A practical guide to trusted computing," *Pearson Education*, 2007.

[5] I. Dbibih and *et al.*, "ASS-MAC: adaptive sleeping sensor MAC protocol designed for wireless sensor networks," in *International Conference on Information Technology for Organizations Development*, pp. 1-5, 2016.

[6] A. Dempster, "Upper and lower probabilities induced by multivalued mapping," *The Annals of Mathematical Statistics*, vol. 38, no. 2, pp. 325-339, 1967.

[7] S. Ganeriwal, *et al.*, "Reputation based framework for high integrity sensor networks," *ACM Transactions on Sensor Networks*, vol. 4, no. 3, pp. 15-37, 2008.

[8] R. Ge, H. Zhang and S. Gong, "Improving on LEACH protocol of wireless sensor networks using fuzzy logic," *Journal of Information & Computational Science*, vol. 7, no. 3, pp. 767-775, 2010.

[9] W. B. Heizelman, A. P. Chandrakasan and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *IEEE Transactions on Wireless Sensor Communications*, vol. 1, no. 4, pp. 660-670, 2002.

[10] W. B. Heinzelman, *et al.*, "An application-specific protocol architecture for wireless microsensor networks," *IEEE Transactions on Wireless Communications*, vol. 1, no. 4, pp. 660-670, 2002.

[11] R. Hou, W. Ren and Y. Zhang, "A wireless sensor network clustering algorithm based on energy and distance," in *Proceedings of the International Workshop on Computer Science and Engineering*, pp. 439-442, 2009.

[12] K. Jin, Y. Zhang and D. Tian, "Based on the improvement of leach protocol for wireless sensor network routing algorithm," in *Proceedings of the International Conference on Intelligent System Design and Engineering Application*, pp. 1525-1528, 2012.

[13] J. Lee and W. Cheng, "Fuzzy logic based clustering approach for wireless sensor networks using energy predication," *IEEE Sensors Journal*, vol. 12, no. 9, pp. 2891-2897, 2012.

[14] C. T. Li and M. S. Hwang, "A lightweight anonymous routing protocol without public key en/decryptions for wireless ad hoc networks," *Information Sciences*, vol. 181, no. 23, pp. 5333-5347, 2011.

[15] C. T. Li, M. S. Hwang, C. Y. Liu, "An electronic voting protocol with deniable authentication for mobile Ad hoc networks," *Computer Communications*, vol. 31, no. 10, pp. 2534–2540, June 2008.

[16] Y. Li, L. Ding and F. Liu, "The improvement of LEACH protocol in WSN," in *Proceedings of the International Conference on Computer Science and Network Technology*, pp. 1345-1348, 2011.

[17] L. Liang and *et al.*, "SW-MAC: a low-latency mac protocol with adaptive sleeping for wireless sensor networks," *Wireless Personal Communications*, vol. 77, no. 2, pp. 1191-1211, 2014.

[18] R. Mehta, A. Pandey and P. Kapadia, "Reforming clusters using C-LEACH in wireless sensor networks," in *Proceedings of the International Conference on Computer Communication and Information*, pp. 1-4, 2012.

[19] P. Mukherjee and S. Sen, "Comparing reputation schemes for detecting malicious nodes in sensor networks," *The Computer Journal*, vol. 3, no. 54, pp. 482-498, 2011.

[20] R. Mylsamy and S. Sankaranarayanan, "A preference-based protocol for trust and head selection for cluster-based MANET," *Wireless Personal Communications*, vol. 86, no. 3, pp. 1611-1627, 2016

[21] Y. B. Sailed and A. Olivereau, "A lightweight threat detection system for industrial wireless sensor networks," *International Journal of Network Security*, vol. 18, no. 5, pp. 842-854, 2016.

[22] T. Sanu and M. Thomaskutty, "Lossless address data compression using quadtree clustering of the sensors in a grid based WSN," *Ad Hoc Networks*, vol. 56, pp. 84-95, 2017.

[23] C. So-In, *et al*, "Performance evaluation of LEACH on cluster head selection techniques in wireless sensor networks," in *Proceedings of the International Conference on Computer and Information Technology*, pp. 51-61, 2013.

[24] M. Thein and T. Thein, "An energy efficient cluster-head selection for wireless sensor networks," in *Proceedings of the International Conference on Intelligent System, Modeling and Simulation*, pp. 287-291, 2010.

[25] F. Wang, *et al.*, "SONR: A reliable reputation system of self organized network," *Journal of Network and Computer Applications*, vol. 35, pp. 914-926, 2012.

[26] J. Xu and *et al.*, "Improvement of LEACH protocol for WSN," in *Proceedings of the International Conference on Fuzzy System and Knowledge Discovery*, pp. 2174-2177, 2012.

Biography

Zhe Wei received his Ph.D. degree in computer science and technology in 2015. Now he is a lecturer in Civil Aviation Flight University of China. His main research includes WSN security and applications.

Shuyan Yu received her master degree in software engineering in 2006. Now she is an associate professor in Zhejiang Post and Telecommunication College. Her main research includes computer security and applications.

A Data Sorting and Searching Scheme Based on Distributed Asymmetric Searchable Encryption

Lina Zou, Xueying Wang and Shoulin Yin
(Corresponding author: Shoulin Yin)

Department of Computer and Mathematics, Shenyang Normal University
253 Huanghe N. St, Huanggu Qu, Shenyang 110034, China
(Email: zln0781@sina.com)
(Received Dec. 25, 2016; revised and accepted Mar. 12 & May 2, 2017)

Abstract

Searchable encryption algorithm is a hot issue nowadays. It can sort the results of searching and return the optimal matching files. The essence of Asymmetric searchable encryption is that users exchange the data of encryption, one party sends a ciphertext with key encryption, the other party with another key receives the ciphertext. Encryption key is not the same as the decryption key, and cannot deduce another key from any one of the key, thus it greatly enhances the information protection, and can prevent leakage the user's search pattern. In order to get higher efficiency and security in information retrieval, in this paper we introduce the concept of distributed Searchable asymmetric encryption, which is useful for security and can enable search operations on encrypted data. Moreover, we give the proof of security. Finally, experiments results show that our method has better retrieval efficiency.

Keywords: Asymmetric Searchable Encryption; Data Sorting; Distributed; Searchable Encryption

1 Introduction

With the rapid development of communication technology, cloud service has entered the large number of people's live and work. Exposing the user data security of the third party service providers leads to security issues [12, 28]. To protect user's data privacy has become more and more important and urgent, which requires encryption. However, the cloud service that its characteristics of convenient and flexible way to charge, more and more users choose the local data migration to the cloud server. Many netters delegate to a third party provider or service provider right to search for their data. There are many scholars having done some research about the Public Key Encryption with Keyword Search. For example, sensitive data (i.e. private data of user or commercial data) is put into cloud, cloud service providers can directly read and use these data, which may result in bad effect of violating the privacy of users and damaging the security of data [7, 19, 20, 22, 23]. Therefore, the cloud service is not absolutely reliable. We need to use new technologies to protect privacy and data. And on the premise of guaranteeing safety, it needs to safeguard the normal operation of user as soon as possible. So protection of user privacy and safety of user data becomes a hot issue.

To solve this problem, Searchable Encryption is introduced [2, 8, 9]. Using SE mechanism encrypts data, and the ciphertext is stored in the cloud server. When users need to search some keywords, they can use the keyword to search documents sent to the cloud server [13]. The cloud will receive the search proof test matching for each file, if the match is successful, it implies that the file contains the keyword. Finally, the cloud will return all files matching success back to the user. After receiving the search results, users only need to return to the encrypted files. The majority of the schemes study single keyword, conjunctive keywords and complex search query of public key cryptography based SE schemes [4, 5, 16].

Here is a motivating example for PEKS. This example is according to the reference [1, 15, 24]. Suppose user Alice wants to read her emails from her laptop or smart phone or PAD after she stores her emails in the servers of some email service provider. Because of previous cloud accidents, Alice does not believe the third-party service provider or fears that powerful agencies may require the service provider to surrender all her data. Any user with Alice's public key can send her encrypted emails from the many transmission mediums that only she can decrypt based on standard public key encryption. PEKS scheme produces some email searchable ciphertexts, Alice prepare to find a unique email then, the sender could also attach to the searchable ciphertexts. Alice could make use of keywords to search for this email. Once delegated, the ciphertexts can be searched. Across Alice's email the service provider searches those search ciphertexts contained that match the issued trapdoor, and returns to her a positive match.

There is a standard application in searchable encryption that it supports the order with keyword matching degree in returned document. Hwang [10] proposed an efficient secure channel free public key encryption with conjunctive field keyword search scheme that could stand against the off-line keyword-guessing attacks, which was more suitable for the weak devices used by users. Tsai [25] proposed a publicly verifiable authenticated encryption scheme based on factoring and discrete logarithms. He pointed out that even if either factoring or discrete logarithms was broken, this scheme still could keep the authentication, integration, and confidentiality of the message. Ling [14] proposed an efficient and secure onetime password authentication scheme for wireless sensor networks according to the Lamport's concept to consider the limitations of computation and lower power in a wireless sensor networks. Cao [3] defined and solved the challenging problem of privacy preserving multi-keyword ranked search over encrypted cloud data (MRSE). A set of strict privacy requirements are established for such a secure cloud data utilization system. Among various multikeyword semantics, it chose the efficient similarity measure of "coordinate matching" to capture the relevance of data documents to the search query. They further used "inner product similarity" to quantitatively evaluate such similarity measure and proposed a basic idea for the MRSE based on secure inner product computation, and then gave two significantly improved MRSE schemes to achieve various stringent privacy requirements in two different threat models. Wang [26] used order preserving encryption to encrypt relevance, which could get accuracy results. Wang [27] presented that ranked search greatly enhanced system usability by enabling search result relevance ranking instead of sending undifferentiated results, and further ensured the file retrieval accuracy. Specifically, he explored the statistical measure approach, i.e. relevance score, from information retrieval to build a secure searchable index, and developed a one-to-many order-preserving mapping technique to properly protect those sensitive score information. Jiang [11] proposed a novel privacy preserving keyword search scheme over encrypted cloud data to address this problem. To enable users to search over encrypted data, he firstly adopted a structure named as Inverted Matrix (IM) to build search index. The IM was consisted of a number of index vectors, each of which was associated with a keyword. Then he mapped a keyword to a value as an address used to locate the corresponding index vector. Finally, he masked index vectors with pseudo-random bits to obtain an Encrypted Enlarged Inverted Matrix (EEIM) to preserve the privacy of users. However, the above methods are used for symmetric searchable encryption or some asymmetric searchable encryption methods have low efficiency.

Therefore, we propose a data sorting and searching based on distributed asymmetric searchable encryption in cloud environment. And we combine order-preserving encryption algorithm based on symmetric encryption to realize sorting and searching in asymmetric searchable encryption algorithm. The following is the structure of this paper. In Section 2, we construct the new scheme. Section 3 and Section 4 give the security proof and performance analysis respectively. There is a conclusion in Section 5.

2 Structure of Distributed Asymmetric Encryption Algorithm

Firstly, we introduce distributed asymmetric encryption algorithm (DAEA). A DAEA system includes four probabilistic polynomial time algorithms as follows:

- $KeyGen$: $(K_C, K_{SP_1}, K_{SP_2}, K_{QP_1}, K_{QP_2})$. This algorithm is executed by the client C, inputs a security parameter λ and outputs a secret key K_C to the client C, public keys K_{SP_1} and K_{QP_1} to SP and QP ($K_{SP_1} = K_{QP_1}$) and private keys K_{SP_2} and K_{QP_2} to SP and QP, respectively.

- $Encrypt$: $(I_1, I_2) = Encrypt(K_C, D)$. This algorithm is executed by the client C, inputs a key KC and a set of documents D, outputs an encrypted index I_1 to SP and I_2 to QP.

- $Trapdoor$: (T_1^s, T_2^s). This algorithm is executed by the client C, inputs the private key K_{SP_2}, K_{QP_2} and a query keyword $s \in W$, and outputs a trapdoor T_1^s to SP and trapdoor T_2^s to QP.

- $Test$: $(a = Test(K_{SP_1}, K_{QP_1}, I_1, I_2, T_1^s, T_2^s))$. SP provides K_{SP_1}, I_1, T_1^s and QP provides K_{QP_1}, I_2, T_2^s as input. According to the matching results of W and W' outputs judgment value a, $a \in 0, 1$.

Given the above definition, the public-key encryption scheme with keyword search does the following processes. Firstly, the receiver uses the Setup algorithm to produce his/her public or private key pair. Then, he/she runs the Trapdoor algorithm to create trapdoors T_W (the third service providers can search) for any keyword W. The given trapdoors are as input for the Test algorithm. The third service provider determines whether one sender gives message encrypted by ksEnc containing one of the keywords W specified by the receiver.

Sorting and searching function indicates that all the matched documents will be ordered by a standard. Finally, it returns the most relative k documents to user. Its SQL form is "**ORDERED BY keyword**". We use the order-preserving encryption algorithm to compute correlation score. Sorting function can record encrypted correlation score and construct an index $< keyword, score >$ Key-value pair. Therefore, sorting function can acquire score and order with computing time complexity $O(1)$.

To store index record, this paper utilizes indirect addressing scheme [18, 21] based on sparse matrix to construct a 2-dimension index table A and record $< keyword, correlationscore >$. All the data are encrypted.

When it executes query, server searches all the correlation score of matched documents and selects the optimal k documents. Before encrypting data, it needs a preprocess to safety use order-preserving encryption algorithm. So we construct a order–preserving encryption table to preprocess all the plaintexts. The followings are the steps to store encrypted correlation score.

1) Given a document set $D = (d_1, d_2, \cdots, d_n)$. It scans o^k keywords W for each document $d_k (1 \le k \le n)$ in document set. In d_k, each keyword $w_i^k \in W$. According to the appear frequency of keyword, it calculates correlation score $s_i^k (1 \le i \le o^k)$ and records a $o^k \times 3$ matrix corresponding to d_k. In this matrix, the record of $i-th$ row is $R_i^k = (w_i^k, s_i^k, p_i^k)$.

2) For all documents, data quantity of order-preserving encryption is $N = o^1 + o^2 + \cdots + o^n$ and numbered $s_1, s_2, \cdots, s_N$. Order-preserving encryption result of each data $s_j (1 \le j \le N)$ is e_j. The previous matrix is transformed into a order-preserving encryption table, $i-th$ row records $R_i^k = (w_i^k, s_i^k, p_i^k)$, where e_i^k is the order-preserving encryption result of s_i^k.

3) For a document, it contains $\frac{|d|}{2} + 1$ keywords at most (there is a separator behind each keyword). Therefore, the index table will be filled with $\frac{|d|}{2} + 1$ data items at last, which guarantees that document content has nothing to do with data item number.

The definitions in asymmetric searchable encryption algorithm will be used in this paper as follows.

Definition 1. $s \leftarrow PEKS(K_{pub}, w)$. *Construction subalgorithm of searchable structure. Input public key K_{pub} and a keyword w. Output searchable structure s. This algorithm is executed by user. s and encryption information of document will be submitted to server.*

Definition 2. $c \leftarrow Enc(K_{pub}, d)$. *Document encryption subalgorithm of asymmetric searchable. Input public key K_{pub} and a message d. Output ciphertext c. This algorithm is executed by data sender. c and searchable encryption structures will be submitted to server.*

Definition 3. *Hash function $f_h : 0,1^* \to 0,1^l$. Where l is mapping length. For example, $f_h = MD5$, then $l = 128bits$.*

Our new data sorting and searching scheme based on distributed asymmetric searchable encryption includes two parts: Build algorithm and Filter algorithm. First, it inputs one document d and encryption key in $Build(d, K_{pub})$ algorithm. Second, it scans d and constructs words list. Third, keywords are compared. Fourth, it outputs functional structure and word mapping. The detailed processes of two algorithms are used for encryption function construction and file query respectively as follows (**Algorithm 1** and **Algorithm 2**).

The main idea of Build algorithm is that it extracts keywords from document to construct index table combining data structure of asymmetric searchable encryption

Algorithm 1 Build algorithm $Build(d, K_{pub})$

Input: Document d, encryption public key $K_e = K_{pub}$.
Output: Function structure $L_d = A$ and word mapping $V_d = (v_1, v_2, \cdots, v_r)$.

1) Compute $c \leftarrow ASE.Enc(K_{pub}, d)$.
2) Scan d and get r words. Construct a word table $W = (w_1, w_2, \cdots, w_r)$.
3) Select a different keywords $W' = (w_1, w_2, \cdots, w_a)$ from W.
4) for each word $w_x (1 \le x \le a)$ in W' do
5) Compute $s_x \leftarrow ASE.PEKS(K_{pub}, w_x)$
6) Compute $h_x \leftarrow f_h(s_x)$
7) end for
8) Let $G = (s_1, s_2, \cdots, s_a)$ map to $H = (h_1, h_2, \cdots, h_a)$ and W'
9) for each word $w_y (1 \le y \le r)$ in W do
10) search $h \in H_i$
11) Let $v_y = h$
12) end for
13) Let $V_d = (v_1, v_2, \cdots, v_r)$
14) Data item of document d in order-preserving encryption table is (w_1, e_1, p_1), (w_2, e_2, p_2), $\cdots$, (w_o, e_o, p_o)
15) for each $i \in [0, 1]$, build an index $A[v_{p_i}] = e_i$
16) the rest $\frac{|d|}{2} + 1$ items are filled with random character string.
17) Output.

Algorithm 2 Filter algorithm $Filter(C, L, V_T)$

1) Input n ciphertexts $C = (c_1, c_2, \cdots, c_n)$, corresponding function structure $L = (L_1, L_2, \cdots, L_n)$, mapping $V_T = (V_1, V_2, \cdots, V_n) = (v_1, v_2, \cdots, v_n)$
2) for n function structures, compute $r_1 = L_1[v_1]$, $r_2 = L_2[v_2]$, $\cdots$, $r_n = L_n[v_n]$.
3) Output order $C' = (c_1, c_2, \cdots, c_k)$.

algorithm based on order-preserving encryption scheme. And Filter algorithm, according to encryption index table of each encryption document, ranks the query result and returns the order result. So it can return the optimal match query document.

3 Security Proof

For SP and QP, we suppose secure channels between the three parties which does not collude. It indicates that admissible $Q-query$ protocol running $\prod_{DASE}^{Q}$ $(Q \in N)$ are executed. Normally, to all participants, the protocol $\prod_{DASE}^{Q}$ has the unique public output access pattern $(id_{w_1}(D), \cdots, id_{w_Q}(D))$. If a DSAE scheme is secure, it leaks no information. The following is that we first define ideal functionality of a DSAE scheme:

Functionality X_{DASE}^{Q}. Consider a DSAE scheme with keyword set W, output $(K_C, K_{SP_1}, K_{SP_2}, K_{QP_1}, K_{QP_2}) = Keygen(\lambda)$, and a document set D. X_{DASE}^{Q} $(Q \in N)$ is the functionality that takes as input:

- K_C and keywords $w_1, \cdots, w_Q$ from client C.
- K_{SP_1}, K_{SP_2} from provider SP.
- K_{QP_1}, K_{QP_2} from query proxy QP.
- $id_{w_1}(D), \cdots, id_{w_Q}(D) \rightarrow id_Q(D)$ to all C, SP, QP.

Then, we consider that a DSAE scheme is secure if all the admissible $Q-query$ run $\prod_{DASE}^{Q}$ to compute functionality X_{DASE}^{Q}.

Safety protection model of sorting query function is that given two same documents set D_1 and D_2. Challenger uses LSE to encrypt data D_b. Adversary can query a keyword and acquire ordered document subset, nevertheless, he dose not know which document subset is chosen by challenger.

According to the proposed non-adaptive indistinguishability chosen keyword attack model [6] and order-persevering concept, we present a non-adaptive indistinguishability chosen order attack model as follows.

Definition 4. ***non-adaptive indistinguishability chosen order attack model****. Let Σ be the order query function component. $k \in N$ is secure parameter. Considering the following simulation experiment. A denotes adversary. S is simulator.*

1) $Adv_{\sum,A}(k)$: challenger executes key generation algorithm $Gen(1^k)$ to generate a key K. Adversary generates a document set $D = (d_1, d_2, \cdots, d_n)$ (size of each document is constant), he receives the encrypted document $C = (c_1, c_2, \cdots, c_n)$ and function structure $L = (L_1, L_2, \cdots, L_n)$. Adversary submits a query w, where $w \in d_1 \cap d_2 \cap \cdots \cap d_n$ and receives mapping v from challenger. At last, A returns $b \in [0, 1]$ as the output of experiment.
2) $Simulate_{\sum,A,S}(k)$. Given documents number n, size of each document $|d|$ and mapping size $|v|$. S generates C^*, L^*, v^* and sends the results to adversary A. A returns $b \in [0, 1]$ as the output of experiment.

Then we call that order query function component is CRKA secure. Only for all polynomial time adversary A, there is a simulator S meeting the following formula:

$$|Pr[Adv_{\sum,A}(k) = 1] - Pr[Simulate_{\sum,A,S}(k) = 1]| \leq negl(k)$$

Its probability depends on key generation algorithm $Gen(1^k)$.

Theorem 1. *If interface of LSE has CPA security, OPE algorithm has the POPF-CCA security, then order query function component has CRKA security.*

Proof. Simulator generates (C^*, L^*, v^{ast}) through the following steps. For C^*, simulator generates n random character string $(c_1^*, c_2^*, \cdots, c_n^*)$. Size of each string is $|d|$. For L^*, let $m = |d|/2 + 1$, simulator generates m random character string $(v_1^*, v_2^*, \cdots, v_m^*)$. Size of each string is $|v|$. Simulator constructs a $m \times n$ matrix $E_{m \times n} = (e_{ij}^*)$. Where e_{ij}^* is random number. For every document, simulator generates index item $A_j^*[v_i^*] = e_{ij}^* (1 \leq i \leq m, 1 \leq j \leq n)$. For v^*, simulator randomly selects $v^* = v_i^* \in V^*$. □

Theorem 2. *There is no polynomial resolving device which can distinguish (C, L, v) and (C^*, L^*, v^*).*

Proof. Key K is encrypted for adversary, hence CPA security of interface directly guarantees the indistinguishability between C and C^*, as well as v and v^*. Meanwhile, when receiving $v = v_i \in V$ or $v^* = v_i^* \in V$, adversary can call the function $Filter(C, L, v)$ or $Filter(C^*, L^*, v^*)$ to acquire $r_1 = L_1[v_i] = (e_1, e_2, \cdots, r_n = L_n[v_i] = e_n)$ or $r_1^* = L_1^*[v_i^*] = (e_1^*, e_2^*, \cdots, r_n^* = L_n^*[v_i^*] = e_n^*)$. It is undistinguish between POPF-CCA security guarantee set $(r_1, r_2, \cdots, r_n)$ and set $(r_1^*, r_2^*, \cdots, r_n^*)$. That is to say, adversary cannot distinguish the encryption result of OPE and output of randomly order-preserving function. Therefore, L and L^* are indistinguishable. □

4 Performance Analysis

This new scheme is coded by MATLAB. Each document is set as 10KB. Its content is the word combination randomly selected from dictionary. Running results of Filter algorithm are as Figure 1.

From Figure 1, we can know that the reason why order asymmetric searchable encryption scheme has high running efficiency is that asymmetric searchable encryption structure is transformed into symmetric searchable encryption structure when encrypting data. They can have the same encryption efficiency. For order query, its main operation is to acquire correlation score from index table. This table is maintained by indirect address of spare

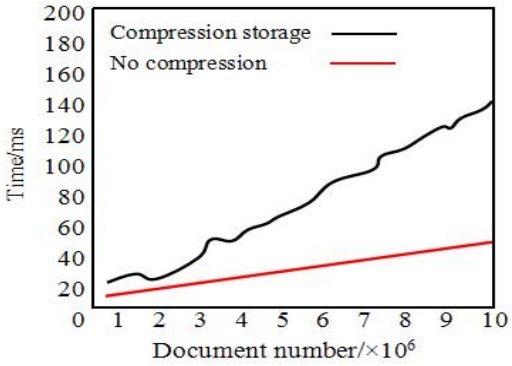

Figure 1: Document retrieval performance

matrix (i.e. compression storage). Let n be number of further required documents. So the total query time complexity is $O(n)$. In addition, we compare the score from the final result and select the optimal k matched documents. Its total computation complexity is $O(n)$ too.

The following is the comparison experiment. The experience data consists of 200 short messages. For the random input keywords set, cloud server will search all the data and find the required keywords set. This experiment is conducted 100 times. And we use three aspects to evaluate the performance of our method including index time cost (Time cost on data owner building retrieval index includes time of extracting and encrypting each keyword in data document.), trapdoor generated time cost (Time cost on data user building trapdoor includes the time of encrypting time and the time of generating trapdoor.) and query time cost (Time cost on cloud server completing a retrieval request includes time of computing document similarity degree and ranking time.).

The document number ranges from 100 to 600, and we select 60 keywords in each document. Table1 is the index time cost with different documents. And a comparison to PRMSN [29] and NMRSS [17] is shown in Table 1. From Table 1, we can know that the time cost of creating index will increase with the adding of documents. In addition, because the time of building sub-index for each document is unchanged, the relation between time cost and document number is quasi-linear.

Table 1: Index time cost with different documents

Document number	The Proposed Method	PRMSN	NMRSS
100	0.25s	0.35s	0.34s
200	0.49s	0.52s	0.51s
300	0.74s	8.79s	0.81s
400	0.93s	1.13s	1.11s
500	1.07s	1.22s	1.22s
600	1.61s	1.71s	1.77s

Then we change the keywords number in each document. The total number of document is 600 under the different keywords. So the index time with different keywords is as shown in Table 2. Table 2 shows the effectiveness of our new method. The more keywords are, the retrieval time is higher. However, our method can take less time.

Table 2: Index time cost with different keywords

Keyword Number	The Proposed Method	PRMSN	NMRSS
10	0.47ms	0.57ms	0.58ms
20	1.12ms	1.33ms	1.30ms
30	1.35ms	1.53ms	1.54ms
40	1.57ms	1.69ms	1.75ms
50	1.74ms	1.92ms	1.89ms
60	1.77ms	1.92ms	1.92ms

Furthermore, we make experience for generating index time with different keyword number as Table 3. In Table 3, when the maximum number of keyword is $d = 30$, the trapdoor generating time is the optimal with our method. From Table 3 we can know that the keyword number cannot affect the time cost and all the time cost is less than 0.002s, which on account of the fact that it adds virtual keywords into the input keyword set. This can ensure that the total number of keyword is d. And Table 4 is the index generating time with different maximum keyword number. Input keyword number is 10 in Table 4 whose time change trend is closely to Table 3. In that keyword number increases, the order of polynomial function increases too, our method costs less time than other two methods.

Table 3: Generating trapdoor time with different keywords number

Keyword Number	The Proposed Method	PRMSN	NMRSS
5	1.14ms	1.40ms	1.42ms
10	1.14ms	1.44ms	1.43ms
15	1.17ms	1.40ms	1.38ms
20	1.17ms	1.40ms	1.40ms
25	1.17ms	1.43ms	1.44ms
30	1.17ms	1.44ms	1.44ms

Table 4: Generating trapdoor time with different maximum keywords number

Maximum Keyword No.	The Proposed Method	PRMSN	NMRSS
10	0.45s	0.79s	0.78s
20	0.94s	1.08s	1.11s
30	1.14s	1.46s	1.47s
40	1.23s	1.57s	1.58s

Finally, we use the inquiry time to demonstrate the

effective of our new method. We first keep that the maximum number of input keyword is 30 and 40 keywords in each document, then we change the document number. Inquiry time with different documents number is as shown in Table 5. Cloud server will cost longer time to search all the data set with the increase of document. Then we set document number as 600 and change the maximum number of input keyword as table6. Data in table6 shows that maximum number of input keyword has none effect on our method. Therefore, our method can execute effectively multi-keyword retrieval in cloud environment.

Table 5: Inquiry time with different documents

Document Number	The Proposed Method	PRMSN	NMRSS
100	0.11s	0.12s	0.13s
200	0.21s	0.21s	0.21s
300	0.28s	0.34s	0.33s
400	0.34s	0.44s	0.43s
500	0.42s	0.54s	0.53s
600	0.62s	0.71s	0.69s

Table 6: Inquiry time with different maximum keywords number

Maximum Keyword No.	The Proposed Method	PRMSN	NMRSS
10	0.57s	0.64s	0.61s
20	0.57s	0.66s	0.62s
30	0.61s	0.68s	0.65s
40	0.78s	0.79s	0.83s

5 Conclusions

In this paper, we propose a data sorting and searching scheme based on asymmetric searchable encryption in cloud environment, which offsets the deficiency of asymmetric searchable encryption. This new scheme combines the advantage of asymmetric searchable encryption and symmetric searchable encryption, it can be extended to other data structure based on symmetric searchable encryption. Therefore, we will study more advanced searchable encryption schemes to improve our method in the future.

6 Acknowledgement

The authors gratefully acknowledge the anonymous reviewers for their valuable comments.

References

[1] A. Arriaga, Q. Tang, P. Ryan, "Trapdoor privacy in asymmetric searchable encryption schemes," *International Conference on Cryptology in Africa*, pp. 31-50, 2014.

[2] M. Bellare, A. Boldyreva, A. O'Neill, "Deterministic and efficiently searchable encryption," in *Annual International Cryptology Conference*, pp. 535-552, 2007.

[3] N. Cao, C. Wang, M. Li, K. Ren and W. Lou, "Privacy-preserving multi-keyword ranked search over encrypted cloud data," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 1, pp. 222-233, 2014.

[4] D. Cash, S. Tessaro, "The locality of searchable symmetric encryption," in *International Conference on the Theory and Applications of Cryptographic Techniques*, pp. 351-368, 2014.

[5] Q. Chai, G. Gong, "Verifiable symmetric searchable encryption for semi-honest-but-curious cloud servers," in *IEEE International Conference on Communications (ICC'12)*, pp. 917-922, 2012.

[6] M. Chase, E. Shen, "Substring-searchable symmetric encryption," *Proceedings on Privacy Enhancing Technologies*, vol. 2, pp. 263-281, 2015.

[7] P. S. Chung, C. W. Liu, and M. S. Hwang, "A study of attribute-based proxy re-encryption scheme in cloud environments," *International Journal of Network Security*, vol. 16, no. 1, pp. 1–13, 2014.

[8] R. Curtmola, J. Garay, S. Kamara and R. Ostrovsky, "Searchable symmetric encryption: improved definitions and efficient constructions," *Journal of Computer Security*, vol. 19, no. 5, pp. 895-934, 2011.

[9] S. T. Hsu, C. C. Yang, and M. S. Hwang, "A study of public key encryption with keyword search", *International Journal of Network Security*, vol. 15, no. 2, pp. 71-79, Mar. 2013.

[10] M. S. Hwang, S. T. Hsu, C. C. Lee, "A new public key encryption with conjunctive field keyword search scheme," *Information Technology and Control*, vol. 43, no. 3, pp. 277-288, Sep. 2014.

[11] X. Jiang, J. Yu, F. Kong, X. Cheng and R. Hao, "A novel privacy preserving keyword search scheme over encrypted cloud data," in *10th IEEE International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC'15)*, pp. 836-839, 2015.

[12] S. Kamara, K. Lauter, "Cryptographic cloud storage," in *International Conference on Financial Cryptography and Data Security*, pp. 136-149, 2010.

[13] C. C. Lee, S. T. Hsu, M. S. Hwang, "A study of conjunctive keyword searchable schemes," *International Journal of Network Security*, vol. 15, no. 5, pp. 321-330, 2013.

[14] C. H. Ling, C. C. Lee, C. C. Yang and M. S. Hwang, "A secure and efficient one-time password authentication scheme for WSN," *International Journal of Network Security*, vol. 19, no. 2, pp. 177-181, Mar. 2017.

[15] B. Martens, F. Teuteberg, "Risk and compliance management for cloud computing services: Designing a reference model," in *Americas Conference on Information Systems (Amcis'11)*, 2011.
[16] T. Moataz, A. Shikfa, "Boolean symmetric searchable encryption," in *Proceedings of the 8th ACM SIGSAC Symposium on Information, Computer and Communications Security*, pp. 265-276, 2013.
[17] R. R. Netinti, S. Madhri, "A novel multi-keyword ranked search system in encrypted and Synonym queries supported cloud," *International Journal of Science Engineering and Advance Technology*, vol. 3, no. 12, pp. 1370-1373, 2016.
[18] C. Nit, L. M. Itu, C. Suciu, "GPU accelerated blood flow computation using the lattice boltzmann method," in *IEEE High Performance Extreme Computing Conference (HPEC'13)*, pp. 1-6, 2013.
[19] D. Patel, "Accountability in cloud computing by means of chain of trust dipen contractor," *International Journal of Network Security*, vol. 19, no. 2, pp. 251-259, 2017.
[20] S. Subashin, V. Kavitha, "A survey on security issues in service delivery models of cloud computing," *Journal of Network and Computer Applications*, vol. 34, no. 1, pp. 1-11, 2011.
[21] C. T. Sungur, P. Spiess, N. Oertel and O. Kopp, "Extending BPMN for wireless sensor networks," in *IEEE 15th Conference on Business Informatics*, pp. 109-116, 2013.
[22] H. Takabi, J. B. D. Joshi, G. J. Ahn, "Security and privacy challenges in cloud computing environments," *IEEE Security & Privacy*, vol. 8, no. 6, pp. 24-31, 2010.
[23] H. Takabi, J. B. D. Joshi, G. J. Ahn, "Securecloud: Towards a comprehensive security framework for cloud computing environments," in *IEEE 34th Annual Conference on Computer Software and Applications Conference Workshops (COMPSACW'10)*, pp. 393-398, 2010.
[24] Q. Tang, X. Chen, "Towards asymmetric searchable encryption with message recovery and flexible search authorization," in *Proceedings of the 8th ACM SIGSAC Symposium on Information, Computer and Communications Security*, pp. 253-264, 2013.
[25] C. Y. Tsai, C. Y. Liu, S. C. Tsaur and M. S. Hwang, "A publicly verifiable authenticated encryption scheme based on factoring and discrete logarithms," *International Journal of Network Security*, vol. 19, no. 3, pp. 443-448, May 2017.
[26] C. Wang, N. Cao, J. Li and K. Ren, "Secure ranked keyword search over encrypted cloud data," in *IEEE 30th International Conference on Distributed Computing Systems (ICDCS'10)*, pp. 253-262, 2010.
[27] C. Wang, N. Cao, K. Ren and W. Lou, "Enabling secure and efficient ranked keyword search over outsourced cloud data," *IEEE Transactions on Parallel and Distributed Systems*, vol. 23, no. 8, pp. 1467-1479, 2012.
[28] C. Wang, S. Chow, Q. Wang, K. Ren and W. Lou, "Privacy-preserving public auditing for secure cloud storage," *IEEE Transactions on Computers*, vol. 62, no. 2, pp. 362-375, 2013.
[29] W. Zhang, S. Xiao, Y. Lin, T. Zhou and S. Zhou, "Privacy preserving ranked multi-keyword search for multiple data owners in cloud computing," *IEEE Transactions on Computers*, vol. 65, no. 5, pp. 1566-1577, 2016.

Biography

Lina Zou was born in 1980. She graduated and received her M.S. degrees from Northeastern University in 2007. Lina Zou is a lecturer in the department of computer and mathematics, Shenyang Normal University, Shenyang, China. She has published many international papers indexed by EI and SCI. Her research interests include the theory of computer, Mathematics, Information Security and Intelligence Algorithm.

Xueying Wang was born in 1966. She is a professor in department of computer and mathematics and Software College, Shenyang Normal University. She received her B.S. and M.S. degrees from Wuhan University. Her research interests include Enterprise informatization and Innovation and entrepreneurship education.

Shoulin Yin received the B.Eng. And M.Eng. degree from Shenyang Normal University, Shenyang, Liaoning province, China in 2013 and 2015 respectively. His research interests include Multimedia Security, Network Security, Filter Algorithm and Data Mining. He received School Class Scholarship in 2015. Email: 352720214@qq.com.

Cryptanalysis of Novel Extended Multivariate Public Key Cryptosystem with Invertible Cycle

Gang Lu[1], Linyuan Xue[1], Xuyun Nie[1,2,3] and Zhiguang Qin[1,3]
(Corresponding author: Xuyun Nie)
School of Information and Software Engineering & University of Electronic Science and Technology of China[1]
Section 2, North Jianshe Road, Chengdu 610054, China
(Email: xynie@uestc.edu.cn)
State Key Laboratory of Information Security & Institute of Information Engineering[2]
Network and Data Security Key Laboratory of Sichuan Province[3]
(Received Dec. 1, 2016; revised and accepted Mar. 31, 2017)

Abstract

In 2016, Qiao *et al.* proposed a novel extended multivariate public key cryptosystem (EMC) to enhance the security of multivariate public key cryptosystem. They applied it on Matsumoto-Imai (MI) encryption scheme and claimed that the enhanced MI scheme can be secure against Linearization Equation (LE) attack. Through analysis, we found that the enhanced MI scheme satisfied Quadratization Equations (QE). After finding all the quadratization equations, we can recover the plaintext corresponding to a valid ciphertext of the enhanced MI scheme.

Keywords: Extended Multivariate Public Key Cryptosystem; Invertible Cycle; Multivariate Cryptography; Quadratization Equation

1 Introduction

In recent years, more and more researches have been made on the quantum computer. Once large-scale quantum computers are successfully built, the traditional public key cryptosystems such as RSA and ElGamal were no longer secure [19, 25]. The study of Post-quantum cryptography is urgent. Multivariate public key cryptosystem (MPKC) is one of promising alternative public key cryptosystem. The security of the MPKC is depended on the difficulty of solving systems of randomly chosen multivariate nonlinear polynomial equations over finite fields. Up to now, quantum computers do not appear to have advantage over the traditional computers to handle with this problem.

From 1988, many cryptosystems have been proposed in MPKCs, such as Matsumoto-Imai (MI) cryptosystem[15], Oil-Vinegar signature scheme [14, 22], Hidden Field Equation cryptosystem (HFE) [21], Tamed Transformation Method (TTM) [16], Medium Field Equation (MFE) cryptosystem [26] etc. But most of them are not secure. Hence, many security enhancement methods have been put forward, for example, Plus/Minus [23], Internal perturbation [4, 6], Piece in Hand method [13] etc. All of these methods are subjected to different levels of attacks [7, 8, 11, 12, 17, 18].

In 2016, Qiao *et al.* [24] proposed an idea named novel Extended Multivariate public key Cryptosystems(EMCs), which introduce nonlinear invertible transformations to enhance the security of defective MPKCs. They used three different nonlinear invertible transformations, invertible cycle, tame transformation and special oil and vinegar, and applied them on MI scheme. The original MI scheme was broken by Patarin [20] using Linearization Equation(LE) attack. Three enhanced MI schemes can resist LE attack.

In this paper, we focus on the enhanced MI scheme with invertible cycle. MI scheme satisfied LEs of form

$$\sum a_{ij}x_iy_j + \sum b_ix_i + \sum c_jy_j + d = 0, \qquad (1)$$

where x_i are the plaintext variables and y_j are the ciphertext variables. In the enhanced MI scheme with invertible cycle, they only applied a quadratic map on plaintext variables before performing MI encryption function. So, this scheme would satisfied a type of equation named Quadratization Equation(QE) of form

$$\sum a_{ijk}x_ix_jy_k + \sum b_{ij}x_iy_j + \sum c_iy_i + \sum d_{ij}x_ix_j + \sum e_ix_i + f = 0.$$

After finding all QEs for a given public key, substitute a valid ciphertext into these QEs, we can derive a set of quadratic equations on plaintext variables. Combining these quadratic equations with the public key and the valid ciphertext, we can recover the corresponding plaintext easily by Gröbner bases method.

This paper is organized as follows. In Section 2, we give some necessary fundamental notion and a brief description of the EMC with invertible cycle. Then, we present theoretical analysis and experimental results of our QE attack in Section 3. Finally, a conclusion was made in Section 4.

2 Preliminaries

2.1 General Structure of MPKC

Let m, n are two positive integers and $k = GF(q)$ is a finite filed. $\bar{F} : k^n \to k^m$ is built as a composition of three maps:

$$\bar{F} = L_1 \circ F \circ L_2$$

where $F : k^n \to k^m$, named central map, is an invertible map. $L_1 : k^m \to k^m$ and $L_2 : k^n \to k^n$ are two invertible affine maps used to hide the structure of F.

The public key of MPKC consists of a set of multivariate quadratic polynomials over a finite field, which is the expression of map $\bar{F}$, that is

$$\begin{aligned}(y_1, \cdots, y_m) &= \bar{F}(x_1, \cdots, x_n) \\ &= L_1 \circ F \circ L_2(x_1, \cdots, x_n) \\ &= (\bar{f}_1, \cdots, \bar{f}_m),\end{aligned}$$

where $\bar{f}_1, \cdots, \bar{f}_n \in k[x_1, \cdots, x_n]$ are a set of nonlinear polynomials. The private keys are L_1 and L_2.

2.2 Direct attack

The direct attack to recover plaintext is to find a solution by solving the following system

$$\begin{cases} y'_1 = \bar{f}_1(x_1, \ldots, x_n) \\ \quad\vdots \\ y'_m = \bar{f}_m(x_1, \ldots, x_n) \end{cases} \tag{2}$$

where $\bar{f}_i$ $(1 \le i \le m)$ are the components of a given public key $\bar{F}$ and $\boldsymbol{y}' = (y'_1, \ldots, y'_m)$ is a ciphertext under this public key. A straightforward way to solve this system is Gröbner basis [1] method and its variants $\mathbf{F}_4$ [9] and $\mathbf{F}_5$ [10]. According to [2], the complexity of $\mathbf{F}_5$ is bounded by

$$O\left(\binom{n + d_{reg}}{n}^{\omega}\right)$$

where n is the number of the plaintext variables, d_{reg} is the degree of regularity in Gröbner basis method and $2 \le \omega \le 3$.

2.3 Matsumoto-Imai Scheme

MI [15] scheme was proposed by Matsumoto and Imai in 1988. Let $k = GF(q)$ is a finite filed with characteristic 2, and K is a degree n extension of k. Let $\phi : K \to k^n$ is a standard k-linear isomorphism between K and k^n as follow:

$$\phi(a_0 + a_1 x + \cdots + a_{n-1} x^{n-1}) = (a_0, a_1, \cdots, a_{n-1}).$$

Choose θ $(0 < \theta < n)$ such that $gcd(q^\theta + 1, q^n - 1) = 1$ and define the map $\widetilde{F}$ over K by $\widetilde{F}(X) = X^{q^\theta + 1}$.

The condition of θ ensure that $\widetilde{F}$ is an invertible map. Indeed, if t is an integer such that $t(1 + q^\theta) = 1 \bmod (q^n - 1)$, then $\widetilde{F}^{-1}$ is simply $\widetilde{F}^{-1} = X^t$.

The MI scheme uses $F(x_1, \cdots, x_n) = \phi \circ \widetilde{F} \circ \phi^{-1}(x_1, \cdots, x_n) : k^n \to k^n$ as its central map. Let L_1 and L_2 be two invertible affine transformations over k^n. The MI encryption map was defined as follows

$$\bar{F}(x_1, \cdots, x_n) = L_1 \circ F \circ L_2(x_1, \cdots, x_n) = (\bar{f}_1, \cdots, \bar{f}_n).$$

where $\bar{f}_1, \cdots, \bar{f}_n \in k[x_1, \cdots, x_n]$.

The public keys of MI are n polynomials, $(\bar{f}_1, \cdots, \bar{f}_n)$, and private keys are (L_1, L_2, θ).

2.4 Linearization Equation

The linearization equation(LE) is put forward by Patarin in 1995 [20] to break MI scheme.

In general, the form of a linearization equation given by

$$\sum_{i=1}^{n} a_i x_i A_i(y_1, \cdots, y_m) + B(y_1, \cdots, y_m) + c = 0,$$

where $x_i, (1 \le i \le n)$ are plaintext variables, $y_i, (1 \le i \le n)$ are ciphertext variables, $A_i, (1 \le i \le n)$ and B are polynomial functions with respect to the ciphertext variables.

It is obvious that LE is linear on plaintext variables. In other words, given a valid ciphertext $(y_1, \cdots, y_m)$ and substituted it into LE, LE will become a linear polynomial equation on plaintext variables.

We usually refer to the maximum degree of ciphertext variables as the order of the LE.

For example, the first order linearization equation (FOLE) is given by

$$\sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij} x_i y_j + \sum_{i=1}^{m} b_i y_i + \sum_{i=1}^{n} c_i x_i + d = 0.$$

And the second order linearization equation (SOLE) is of form

$$\sum_{i=1}^{n}\sum_{j=i}^{m}\sum_{k=j}^{m} a_{ijk} x_i y_j y_k + \sum_{i=1}^{m}\sum_{j=i}^{m} b_{ij} y_i y_j + \sum_{i=1}^{n}\sum_{j=i}^{m} c_{ij} x_i y_j + \sum_{i=1}^{m} d_i y_i + \sum_{i=1}^{n} e_i x_i + f = 0.$$

Linearization Equation can help us to do elimination on the system (2). For more information about LE attack, please refer to [5] and [17].

2.5 Quadratization Equation

The quadratization equation attack is proposed by Cao *et al.* [3] in 2010. The general form of a quadratization equation is

$$\sum_{i=1}^{n}\sum_{j=i}^{n} a_{ij}x_ix_jA_{ij}(y_1,\cdots,y_m)+\sum_{i=1}^{n} b_ix_iB_i(y_1,\cdots,y_m) +C(y_1,\cdots,y_m)+c=0$$

where $x_i,(1 \le i \le n)$ are plaintext variables, $y_i,(1 \le i \le n)$ are ciphertext variables, $A_{ij}(y_1,\cdots,y_m), B_i(y_1,\cdots,y_m)$ and $C(y_1,\cdots,y_m)$ are polynomial functions in the ciphertext variables.

We can find that substituting a valid ciphertext $(y_1,\cdots,y_m)$ into a QE, the QE will become a quadratic equation on plaintext variables. If we can derive a set of QEs, we will derive a set of quadratic equations on plaintext variables for a valid ciphertext. Combine these equations with the system (2), the degree of regularity might be lower down in solving the system (2) by Gröbner basis method. Hence, the complexity of solving the system (2) will be smaller. Similar to the LE, we can also define the order of the QE as the maximum degree of ciphertext variables.

The first order quadratization equation (FOQE), an example of QE, is given by

$$\sum_{i=1}^{n}\sum_{j=i}^{n}\sum_{k=j}^{m} a_{ijk}x_ix_jy_k+\sum_{i=1}^{n}\sum_{j=i}^{m} b_{ij}x_iy_j+ \sum_{i=1}^{m} c_iy_i+\sum_{i=1}^{n}\sum_{j=i}^{n} d_{ij}x_ix_j+\sum_{i=1}^{n} e_ix_i+f=0.$$

2.6 The Novel EMC

The Novel EMC, designed by Qiao *et al.* [24], may serve as an security enhancement method both on encryption system and signature system. The main idea of the novel EMC is that they introduced a nonlinear invertible transformation L_3 and applied it on the plaintext variables before the original encryption map work, namely, as in Equation (3):

$$\begin{aligned}\widetilde{F}(x_1,\cdots,x_n) &= \bar{F}\circ L_3(x_1,\cdots,x_n)\\ &= L_1\circ F\circ L_2\circ L_3(x_1,\cdots,x_n).\end{aligned} \tag{3}$$

where $(x_1,\cdots,x_n)\in k^n$, $k=GF(q)$.

The public key of the novel EMC is the expression of map $\widetilde{F}$ and the private keys are L_1, L_2 and L_3.

In [24], they chose three types of nonlinear invertible transformation L_3, invertible cycle, tame transformation and special oil and vinegar. In the following parts of this paper, we only give cryptanalysis of the novel EMC with invertible cycle.

The L_3 as invertible cycle is described as follows.

Let μ is an invertible map on positive integer, given by

$$\mu:\{1,\cdots,n\}\to\{1,\cdots,n\}:\mu(i)=\begin{cases}1 & \text{for } i=n\\ i+1 & \text{otherwise}\end{cases}$$

For $n\ge 2$, let $L_3:(x_1,\cdots,x_n)\to(t_1,\cdots,t_n)$ be a nonlinear transformation over k^n, defined as Equation (4):

$$\begin{cases} t_1=\begin{cases}c_1x_1x_2 & \text{for } n \text{ odd}\\ c_1x_1^qx_2 & \text{for } n \text{ even}\end{cases},\\ t_i=c_ix_ix_{\mu(i)} \text{ for } 2\le i\le n\end{cases} \tag{4}$$

Remark. *Due to* $(x_1,\cdots,x_n)\in k^n$ *and* $k=GF(q)$, $x_i^q=x_i$. *When* n *is an even,* L_3 *is not invertible because that from (4), we can derive* $x_1=\frac{c_2c_4\cdots c_nt_1t_3\cdots t_{n-1}}{c_1c_3\cdots c_{n-1}t_2t_4\cdots t_n}\cdot x_1$, *that is, we can not derive* x_1 *from* L_3. *Hence, the map* L_3 *is not invertible when* n *is an even. So we consider only the case* n *is an odd.*

The public keys of the novel EMC with invertible cycle are a set quartic polynomials. More detail about process of encryption and decryption please refer to [24].

Practical Parameters. In [24], the authors chose MI encryption scheme as the original MPKC and they recommended $k=GF(2^{16})$, and $n=27$.

3 Cryptanalysis of Novel EMC

Although the enhanced MI scheme with invertible cycle can resist linearization equations attack, the design of the L_3 based on "Invertible Cycle" will bring new security hazards to the scheme. Since it is vulnerable to quadratization equation attack, it appears that L_3, at some level, is not an appropriate method to raise the security of the original scheme.

3.1 Quadratization Equations

As we know, the original scheme MI satisfies the first order linearization equation. So the ciphertext variables $y_i,(1\le i\le n)$ and the intermedium variables $t_i,(1\le i\le n)$ satisfy the first order linearization equation, namely,

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}t_iy_j+\sum_{i=1}^{n} b_iy_i+\sum_{i=1}^{n} c_it_i+d=0. \tag{5}$$

Substituting Equation (4) into Equation (5), Equation (5) will change into the following equation:

$$\sum_{i=1}^{n}\sum_{j=i}^{n}\sum_{k=1}^{n} a_{ijk}x_ix_jy_k+\sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}x_iy_j+\sum_{i=1}^{n} c_iy_i +\sum_{i=1}^{n}\sum_{j=i}^{n} d_{ij}x_ix_j+\sum_{i=1}^{n} e_ix_i+f=0 \tag{6}$$

Equation (6) is exactly a Quadratization Equation. To continues our attack, we need find out all quadratization equations. This can be done as follows.

To find all quadratization equations equivalent to find a basis of the space V spanned by all QEs.

The number of coefficients in Equation (6) is equal to $\frac{(n+1)^2(n+2)}{2}$. Then we can randomly generate slightly

over $\frac{(n+1)^2(n+2)}{2}$ plaintext/ciphertext pairs from the public key and substitute them into Equation (6). It is clear that we obtain a system of linear equation on unknown coefficients$(a_{ijk}, b_{ij}, c_{ij}, d_i, e_i, f \in k)$. Solving this system, we can get a basis of the solution space of this system, namely denote by Equation (7).

$$\begin{cases} \sum_{i=1}^{n}\sum_{j=i}^{n}\sum_{k=1}^{n} a_{ijk}^{(\rho)} x_i x_j y_k + \sum_{i=1}^{n}\sum_{j=i}^{n} b_{ij}^{(\rho)} x_i x_j \\ + \sum_{i=1}^{n}\sum_{j=i}^{n} c_{ij}^{(\rho)} x_i y_j + \sum_{i=1}^{n} d_i^{(\rho)} x_i \\ + \sum_{i=1}^{n} e_i^{(\rho)} y_i + f^{(\rho)} = 0 \\ 1 \le \rho \le D \end{cases} \tag{7}$$

where D is the dimension of the space V.

The process above relies merely on any given public key and it can be executed once for all cryptanalysis under that public key.

3.2 Ciphertext-only Attack

For a given cipheretext $\boldsymbol{y}' = (y_1', \cdots, y_n')$, substitute them into Equation (7) and do Gaussian elimination on it, we can get D' quadratic equations on variables, namely:

$$\begin{cases} \sum_{i=1}^{n}\sum_{j=i}^{n} \widetilde{a}_{ij}^{(\rho)} x_i x_j + \sum_{i=1}^{n} \widetilde{b}_i^{(\rho)} x_i + \widetilde{c}^{(\rho)} = 0 \\ 1 \le \rho \le D' \end{cases} \tag{8}$$

Combining these quadratic equations (8) with the system (2), we obtain a new system with $D' + n$ equations on plaintext variables. Then we solve the new system by Grönber basis algorithm. Experiments results show the corresponding plaintext can be recovered efficiently.

The algorithm of our attack can be seen in Algorithm 1.

Algorithm 1 Steps of QE Attack

1: **Input:** public key $\bar{F}$ of a MPKC, ciphertext $\boldsymbol{y}' \in k^n$
2: **Output:** corresponding plaintext $\boldsymbol{x}' \in k^n$
3: Determine the number of QE. It is $\frac{(n+1)^2(n+2)}{2}$;
4: Compute $N > \frac{(n+1)^2(n+2)}{2}$ plaintext/ciphertext pairs from the public key;
5: Substitute these plaintext/ciphertext pairs into Equation (6) and solve the resulted linear system;
6: Substitute the ciphertext $\boldsymbol{y}'$ into the quadratization equation found by last step and obtain D' quadratic equations on the plaintext variables.
7: Combine the quadratic equations with the system (2) to get a new system on plaintext variables. Solve the system directly via Gröbner Basis method.

3.3 Complexity and Experiments Results

In our attack, we set $k = GF(2^{16})$, $n = 27$, and the original MPKC is MI encryption scheme with $\theta = 4$. We chose randomly a valid ciphertext $\boldsymbol{y}' = (y_1', \cdots, y_n')$ and we want to find the corresponding plaintext $\boldsymbol{x}' = (x_1', \cdots, x_n')$.

In the first step, the number of coefficients in QE is equal to $\frac{(n+1)^2(n+2)}{2} = \frac{22736}{2} = 11368$. We computed 11370 plaintext/ciphertext pairs and substituted them into Equation (6) and did Gaussian Elimination on the resulted linear system. The complexity of $\frac{(n+1)^2(n+2)}{2}$ plaintext/ciphertext pairs generation is about $O(n^8)$. It is about 2^{38} for $n = 27$. And the complexity of the Gaussian Elimination is less than $(\frac{(n+1)^2(n+2)}{2})^3$, which is less than 2^{41} for $n = 27$. The dimension D of the space spanned by all QEs is equal to 26 in our experiments. This step is the most time consuming step in our attack. But this can be done once for a given public key.

In the second step, we substituted a valid ciphertext $\boldsymbol{y}' = (y_1', \cdots, y_n')$ into Equation (7) and we obtained 26 linear independent quadratic polynomials equation on plaintext variables.

In last step, combining the quadratic equations derived in step 2 with the system (2), we used Gröbner basis solving it and obtained the corresponding plaintext. Extensive experimental evidence has shown that the degree of regularity in solving the system is 3, hence the complexity of this step is $O\left(\binom{n+3}{n}^{\omega}\right)$, which is less than 2^{36} for $n = 27, 2 \le \omega \le 3$.

We performed our attack via Magma on a PC with Intel Core i5-3350P CPU 3.10 GHz and 4 GB of memory. In our experiments, we chose different parameters to illustrated our attack.

In Table 1, we showed the time of three stages under different parameters. T_1 indicates the time of generating $\frac{(n+1)^2(n+2)}{2}$ plaintext-ciphertext pairs from the public key. T_2 indicates the time of obtaining the quadratization equations. T_3 indicates the time of recovering the plaintext.

In Table 2, we compared the efficiency of our attack with the direct attack on the EMC with invertible cycle. The results showed that the degree of regularity in Gröbner basis method is reduced, so as to the execution time. In Table 2, $Time_Q$ and $d_{reg}(Q)$ express the time of recovering the plaintext and the degree of regularity in our attack and $Time_D$ and $d_{reg}(D)$ express the time and the degree of regularity in direct attack. According to the results in our experiments, the complexity of direct attack is $O\left(\binom{n+6}{n}^{\omega}\right)$, which is greater than the complexity of our attack, $O\left(\binom{n+3}{n}^{\omega}\right)$.

4 Conclusions

In this paper, we presented the cryptanalysis of the novel EMC with invertible cycle by Quadratization Equation attack. The same method can also be applied on the novel EMC with tame transformation. The emergence of Quadratization Equation can damage the security of

Table 1: The time comparison of practical attack under different parameters

n	**q**	D	D'	T_1 [**s**]	T_2[**s**]	T_3[**s**]	d_{reg}
21	2^8	20	20	25.265	131.922	0.36	3
21	2^{16}	20	20	27.487	719.523	0.92	3
23	2^8	22	22	41.969	279.891	0.813	3
23	2^{16}	22	22	48.875	1720.364	2.262	3
25	2^8	24	24	67.328	563.907	1.625	3
25	2^{16}	24	24	81.619	333.726	4.914	3
27	2^8	26	26	105.39	1105.969	3.625	3
27	2^{16}	26	26	114.037	6771.333	17.503	3

Table 2: The efficiency comparison of Quadratization attack & Direct attack

n	**q**	$Time_Q$	$d_{reg}(Q)$	$Time_D$	$d_{reg}(D)$
21	2^8	0.36	3	8.219	6
23	2^8	0.813	3	17.922	6
25	2^8	1.625	3	34.828	6
27	2^8	3.625	3	54.515	6

MPKCs. We should avoid it in designing MPKCs.

Acknowledgments

This work was supported by the National Key Basic Research Program of China under grant 2013CB834203, Major International (Regional) Joint Research Project of China National Science Foundation under grant No.61520106007 and The science and technology foundation of Sichuan Province (No.2016GZ0065). The authors gratefully acknowledge the anonymous reviewers for their valuable comments.

References

[1] W. W. Adams and P. Loustaunau, "An introduction to gröbner bases," *Graduate Studies in Mathematics*, vol. 60, no. 1, p. 167–168, 1994.

[2] M. Bardet, J. C. Faugere, and B. Salvy, "On the complexity of gröbner basis computation of semi-regular overdetermined algebraic equations," in *Proceedings of the International Conference on Polynomial System Solving*, pp. 71–74, 2004.

[3] W. W. Cao, X. Y. Nie, L. Hu, X. L. Tang, and J. T. Ding, "Cryptanalysis of two quartic encryption schemes and one improved MFE scheme," in *International Workshop on Post-Quantum Cryptography*, pp. 41–60, 2010.

[4] J. T. Ding, "A new variant of the Matsumoto-Imai cryptosystem through perturbation," in *International Workshop on Public Key Cryptography*, pp. 305–318, 2004.

[5] J. T. Ding, L. Hu, X. Y. Nie, J. Y. Li, and J. Wagner, "High order linearization equation (HOLE) atttack on multivariate public key cryptosystems," in *International Workshop on Public Key Cryptography*, pp. 233–248, 2007.

[6] J. T. Ding and D. Schmidt, "Cryptanalysis of HFEv and internal perturbation of HFE," in *International Workshop on Public Key Cryptography*, pp. 288–301, 2005.

[7] V. Dubois, P. A. Fouque, A. Shamir, and J. Stern, "Practical cryptanalysis of SFLASH," in *Annual International Cryptology Conference*, pp. 1–12. Springer, 2007.

[8] V. Dubois, L. Granboulan, and J. Stern, "Cryptanalysis of HFE with internal perturbation," in *International Workshop on Public Key Cryptography*, pp. 249–265, 2007.

[9] J. C. Faugere, "A new efficient algorithm for computing gröbner bases (F_4)," *Journal of Pure and Applied Algebra*, vol. 139, no. 1, pp. 61–88, 1999.

[10] J. C. Faugere, "A new efficient algorithm for computing gröbner bases without reduction to zero (F_5)," *Issac Proceedings of the International Symposium on Symbolic Algebraic Computation*, vol. 139, no. 1-3, pp. 75–83, 2004.

[11] J. C. Faugere, A. Joux, L. Perret, and J. Treger, "Cryptanalysis of the hidden matrix cryptosystem," in *International Conference on Cryptology and Information Security in Latin America*, pp. 241–254, 2010.

[12] P. A. Fouque, L. Granboulan, and J. Stern, "Differential cryptanalysis for multivariate schemes," in *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pp. 341–353, 2005.

[13] R. Fujita, K. Tadaki, and S. Tsujii, "Nonlinear piece in hand perturbation vector method for enhancing security of multivariate public key cryptosystems," in *International Workshop on Post-Quantum Cryptography*, pp. 148–164, 2008.

[14] A. Kipnis, J. Patarin, and L. Goubin, "Unbalanced oil and vinegar signature schemes," in *International*

Conference on the Theory and Applications of Cryptographic Techniques, pp. 206–222, 1999.

[15] T. Matsumoto and H. Imai, "Public quadratic polynomial-tuples for efficient signature-verification and message-encryption," in *Workshop on the Theory and Application of of Cryptographic Techniques*, pp. 419–453. Springer, 1988.

[16] T. T. Moh, "A public key system with signature and master key functions," *Communications in Algebra*, vol. 27, no. 5, pp. 2207–2222, 1999.

[17] X. Y. Nie, P. Albrecht, B. Johannes, and F. G. Li, "Linearization equation attack on 2-layer nonlinear piece in hand method," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. 97, no. 9, pp. 1952–1961, 2014.

[18] X. Y. Nie, C. Y. Hou, Z. H. Xu, and G. Lu, "Analysis of second order matrices construction in MFE public key cryptosystem," *International Journal of Network Security*, vol. 18, no. 1, pp. 158–164, 2016.

[19] V. Padmavathi, B. Vishnu Vardhan, A. V. N. Krishna, "Provably secure quantum key distribution by applying quantum gate," *International Journal of Network Security*, vol. 20, no. 1, pp. 88-94, 2018.

[20] J. Patarin, "Cryptanalysis of the Matsumoto and Imai public key scheme of Eurocrypt88," in *Annual International Cryptology Conference*, pp. 248–261, 1995.

[21] J. Patarin, "Hidden fields equations (HFE) and isomorphisms of polynomials (IP): two new families of asymmetric algorithms," in *Advances in Cryptology (EUROCRYPT'96)*, pp. 33–48, 1996.

[22] J. Patarin, "The oil and vinegar signature scheme," in *Dagstuhl Workshop on Cryptography*, 1997.

[23] J. Patarin, L. Goubin, and N. Courtois, "C^*_{-+} and HM: Variations around two schemes of T. Matsumoto and H. Imai," in *International Conference on the Theory and Application of Cryptology and Information Security*, pp. 35–50, 1998.

[24] S. T. Qiao, W. B. Han, Y. F. Li, and L. Y. Jiao, "Construction of extended multivariate public key cryptosystems," *International Journal of Network Security*, vol. 18, no. 1, pp. 60–67, 2016.

[25] M. S. Srinath, V. Chandrasekaran, "Isogeny-based quantum-resistant undeniable blind signature scheme," *International Journal of Network Security*, vol. 20, no. 1, pp. 9-18, 2018.

[26] L. C. Wang, B. Y. Yang, Y. H. Hu, and F. P. Lai, "A "Medium-Field" multivariate public-key encryption scheme," in *The Cryptographers' Track at the RSA Conference (CT-RSA'06)*, pp. 132–149, 2006.

Biography

Gang Lu is a PH.D candidate in University of Electronic Science and Technology of China now. His research interests include cryptography and security of big data.

Linyuan Xue is pursuing his Master degree from the Department of Information and Software Engineering, University of Electronic Science and Technology of China. His research interests include multivariate public key cryptosystems and network security.

Xuyun Nie received the Ph.D. degree in Information security from Graduate university of Chinese Academy of Sciences, Beijing, China, in 2007. He is presently an Associate Professor at University of Electronic Science and Technology of China (UESTC). His research interests include cryptography and information security.

Zhiguang Qin is a professor in the School of Information and Software Engineering, UESTC. He received his PH.D. degree from UESTC in 1996. His research interests include: information security and computer network.

Global Stability of Worm Propagation Model with Nonlinear Incidence Rate in Computer Network

Ranjit Kumar Upadhyay and Sangeeta Kumari
(Corresponding author: Ranjit Kumar Upadhyay)
Department of Applied Mathematics, Indian Institute of Technology (Indian School of Mines)
Dhanbad - 826004, India
(Email: ranjit.chaos@gmail.com)
(Received Oct. 27, 2016; revised and accepted Feb. 20, 2017)

Abstract

In this paper, an e-epidemic $SVEIS$ model describing the transmission of worms with nonlinear incidence rate through horizontal transmission is formulated in computer network. The existence of two equilibrium points: worm-free and endemic equilibria have been investigated. The stability analyses are determined by the basic reproduction number. It has been observed that if the basic reproduction number,

$$R_0 = \frac{\sigma\beta(\mu+\theta)\Lambda}{\eta_2(\mu+\gamma)(\Lambda(\mu+\theta)+\eta_1\mu c)} < 1,$$

the system is globally asymptotically stable and the infected nodes get vanish at worm-free equilibrium state; worms fade out from the network. However, if $R_0 > 1$, the infected node exists; worm persists in the network at an endemic equilibrium state and is globally stable with some conditions. Further, the transcritical bifurcation at $R_0 = 1$, has obtained using the center manifold theorem. The effect of vaccination and non-linear transmission rate on the dynamics of the model system has been observed. The dynamical behavior of the susceptible, exposed and infected nodes with real parametric values is examined. We also observe that the critical vaccination rate is required to eradicate the worm. Our results illustrate several administrative and executive insights.

Keywords: Computer Network; E-Epidemic Model; Nonlinear Incidence Rate; Stability Analysis; Transcritical Bifurcation; Worm Propagation

1 Introduction

Over the last several decades, a rigorous global effort is speeding up the developments, in the establishment of a worldwide surveillance network for the propagation of computer malicious objects (Viruses, Worms and Trojans). Researchers from computer science and applied mathematics have collaborated for fast assessment of potentially critical conditions. To achieve this goal, mathematical modeling plays a vital role in efforts; that focuses on predicting, assessing and controlling potential outbreak. Also, the epidemic modeling and its applications have been used to understand the effect of changes in the behavior of solutions of the model system. To better understand such dynamics, the papers of Kermack and McKendrick [13] and Capasso and Serio [2] can be studied which established the deterministic compartmental epidemic modeling. In this way, many research articles have published and discussed the main approaches that are used for the surveillance and modeling of biological diseases as well as computer viruses dynamics. Wang *et al.* [26] proposed a novel worm attack $SVEIR$ model using saturated incidence rate and partial immunization rate. In which they have shown the partial immunization is highly effective for eliminating worms. A propagation model with varying node numbers of removable memory device(RMD) virus have been formulated and obtained three threshold parameters to control the RMD-virus in [11].

Worm exploits security vulnerabilities and does not require any user action to propagate. It is a self-propagating malicious program that focuses mainly on infecting as many nodes as possible to the network. Several network phenomena are well modeled as transmissions (through both horizontally and vertically) of viruses or worms through a network. We consider the vaccination strategies that are used to control the spreading of malicious objects [24, 30]. The regular pattern of periodic occurrences have been observed in the epidemiology of many infectious diseases and computer viruses. To predict and control the spread of computer worms, it is necessary to understand such periodic patterns and identify the specific factors that exhibit such periodic outbreak [16]. Zhang *et al.* [29] have employed an impulsive state feedback model to study the transmission of computer worm and the preventive ef-

fect of operating system patching.

Recent studies have demonstrated that the nonlinear incidence rate is one of the important factor for the modeling of epidemic and e-epidemic systems that induces periodic oscillations in epidemic models [23, 27]. For modeling the transmission process, researchers have employed different forms of incidence rate on which the dynamics of the model system depends extensively. The e-epidemicity of worm is closely related to the stability of the equilibria of the model system. Many researchers have considered bilinear incidence rate (βSI) [7, 14], standard incidence rate $\frac{\beta SI}{N}$ [18], nonlinear incidence rate $\frac{\beta SI}{f(I)}$ [8], modified saturated incidence rate $f(S,I) = \frac{\beta SI}{(1+\alpha S+\gamma I)}$(Beddington-DeAngelis type), where $\alpha, \beta, \gamma > 0$ [5, 12], *etc.* in their studies. The use of classical e-epidemic transmission for studying computer virus propagation has been investigated by Piqueira *et al.* [21]. In the present work, we have taken nonlinear incidence rate as $f(S,I) = \frac{\beta SI}{(S+I+c)}$ [25].

In the e-epidemiology of worm propagation in the computer network, Mishra and Pandey [19, 18] described the effect of anti-virus software and vaccination on the attack of computer worms with global stability. Also, some research articles appear on the computer worm or virus model with different recovery rates and dynamics [22, 28]. Recently, Upadhyay *et al.* [25] proposed a $SVEIR$ model with nonlinear incidence rate for modeling the virus dynamics in computer network. In this paper, $SVEIS$ model with nonlinear incidence rate and vaccination strength are presented. This work is basically for the implementation and practice to predict and minimize the severe attack of worms in the computer network.

Here, we have investigated the global stability of the proposed e-epidemic $SVEIS$ model using modified nonlinear incidence rate and predict its optimal vaccination and eradicate worms from the network. The paper is structured as follows. In Section 2, we formulate an e-epidemic $SVEIS$ model as the system of ordinary differential equations and give the descriptions of all the parameters used in the model system. We find the two possible equilibrium points and its existence criteria and also calculate the basic reproduction number in Section 3. The stability analysis for both the equilibrium points are analyzed and transcritical bifurcation analysis is executed when basic reproduction number $R_0 = 1$ in Section 4. Section 5 presents the numerical simulations to verify the results found analytically by taking computer relevant value of parameters and discusses the stability of the model system using MATLAB and Mathematica. Finally, we conclude this article in Section 6.

2 Formulation of the Mathematical Model

Consider N nodes which have been divided into four subclasses as susceptible (S), vaccinated (V), exposed (E) and infectious (I) nodes and $N = S + V + E + I$. Some assumptions for formulating the model system are as follows:

1) We assume that any new node entering into the network is susceptible. The crashing rate of a node (due to hardware or software problems) μ is constant throughout the network.

2) The nodes are interacting heterogeneously. Worms are transmitted to the node through horizontal transmission.

3) The worms propagate into network when an infected file is transferred from an infectious node to the susceptible node. We have considered the modified nonlinear incidence rate $f(S,I) = \frac{\beta SI}{S+I+c}$. This represents the fact that the number of nodes carrying the worms can interact with other nodes, reaches some finite maximum value due to limitation of time or the network slowdown problems of the particular nodes [25].

4) Software offers temporary immunity to the nodes that is, when the software loss their efficiency or removed from the node of the computer network, the node becomes susceptible to attack again.

5) The worm induces temporary immunity is a fraction of recovered node, remaining recovered nodes again become susceptible [1]. A small fraction of exposed node recovers, rather than being infected and develops worm acquired temporary immunity due to self-prevention and detection techniques of operating system and becomes vaccinated [24]. A fraction of infected node after recovery gains temporary immunity against the worm and joins vaccinated class, remaining node becomes re-susceptible.

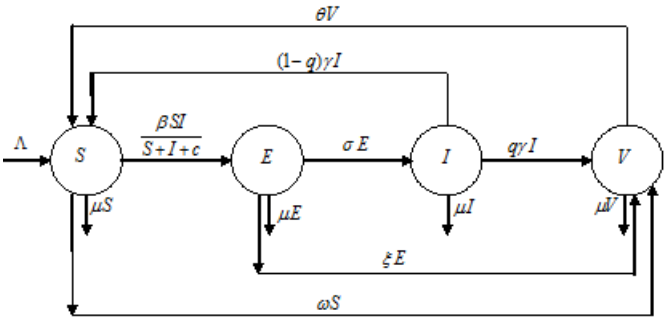

Figure 1: Schematic diagram for model system (1)

The schematic diagram for the model (1) is shown in Figure 1. The worm transmission between the different classes can be expressed by the following model:

$$\begin{cases} \frac{dS}{dt} &= \Lambda - \mu S - \omega S + \theta V - \frac{\beta SI}{S+I+c} + (1-q)\gamma I, \\ \frac{dV}{dt} &= \omega S - \theta V - \mu V + \xi E + q\gamma I, \\ \frac{dE}{dt} &= \frac{\beta SI}{S+I+c} - \mu E - \xi E - \sigma E, \\ \frac{dI}{dt} &= \sigma E - \mu I - \gamma I. \end{cases} \quad (1)$$

Table 1: Definition of parameters

Parameter	Descriptions	Units
S	Susceptible node	In number
V	Vaccinated node	"
E	Exposed node but not yet infectious	"
I	Infectious node	"
Λ	Rate at which new nodes are connected to the network	Day^{-1}
μ	Crashing rate of node due to hardware or software problems	"
ω	Vaccination rate	"
θ	Rate at which vaccinated nodes lose their immunity and join susceptible class	"
β	Contact rate or rate of transfer of worms from an infectious node to susceptible node	"
c	Half saturation constant	In number
q	Fraction of recovered nodes gaining worm-acquired immunity	Day^{-1}
ξ	Recovery rate of exposed class due to self-prevention technique of operating system	"
γ	Duration of infection of infected nodes	"
σ	Rate at which exposed node become infectious	"

with initial conditions: $S(0) = S_0 > 0, V(0) = V_0 > 0, E(0) = E_0 > 0, I(0) = I_0 > 0$, where all the parameters are positive and $0 \leq q \leq 1$. The definitions of all parameters are summarized in Table 1.

3 Existence of Equilibrium Points and Basic Reproduction Number

The existence of worm-free and endemic equilibrium points are established and basic reproduction number has been calculated.

We observe that total number of nodes N satisfies the equation $\frac{dN}{dt} = \Lambda - \mu N$ and hence $N(t) \to \frac{\Lambda}{\mu}$, as $t \to \infty$. The solutions of the model system (1) are non-negative for all $t \geq 0$. Therefore, the feasible region

$$U = \left\{ (S, V, E, I) : 0 \leq S, V, E, I, N \leq \frac{\Lambda}{\mu} \right\},$$

is positively invariant in which the usual existence, uniqueness of solutions and continuation results hold.

The model system (1) always has the worm-free equilibrium $P^0 = \left(S^0, V^0, 0, 0\right)$, where

$$S^0 = \left(\frac{\mu+\theta}{\eta_1}\right) N^0, V^0 = \left(\frac{\omega}{\eta_1}\right) N^0, N^0 = \frac{\Lambda}{\mu},$$

with $\eta_1 = \mu + \theta + \omega$, represent the level of susceptible, vaccinated and total number of nodes respectively, in the absence of infection.

Now, we calculate the basic reproduction number [4]. Let $x = (E, I)$, then from the model system (1), it follows:

$$\frac{dx}{dt} = f - \upsilon,$$

where $f = \begin{bmatrix} \frac{\beta SI}{S+I+c} \\ 0 \end{bmatrix}$ and $\upsilon = \begin{bmatrix} \eta_2 E \\ -\sigma E + (\mu+\gamma) I \end{bmatrix}$.

We obtain $F =$ Jacobian of f at $WFE = \begin{bmatrix} 0 & \frac{\beta S^0}{S^0+c} \\ 0 & 0 \end{bmatrix}$ and $M =$ Jacobian of υ at $WFE = \begin{bmatrix} \eta_2 & 0 \\ -\sigma & \mu+\gamma \end{bmatrix}$.

The next generation matrix approach is used to compute the basic reproduction number, R_0 and is defined as the spectral radius of the next generation operator. The formation of the operator involves determining two compartments, infected and non-infected nodes for the considered model system.

The next generation matrix for the system is

$$K = FM^{-1} = \begin{bmatrix} \frac{\sigma\beta S^0}{\eta_2(\mu+\gamma)(S^0+c)} & \frac{\beta S^0}{(\mu+\gamma)(S^0+c)} \\ 0 & 0 \end{bmatrix}.$$

Thus, the basic reproduction number $R_0 = \rho(FM^{-1})$, of the model system (1) is given by

$$\begin{aligned} R_0 &= \frac{\sigma\beta S^0}{\eta_2(\mu+\gamma)(S^0+c)} \\ &= \frac{\sigma\beta(\mu+\theta)\Lambda}{\eta_2(\mu+\gamma)(\Lambda(\mu+\theta)+\eta_1\mu c)}. \end{aligned}$$

Further, the model system (1) also has an interior equilibrium given by $P^* = (S^*, V^*, E^*, I^*)$, where

$$\begin{aligned} S^* &= \frac{I^*+c}{(R_0-1)+c\frac{R_0}{S^0}}, E^* = \frac{\mu+\gamma}{\sigma} I^*, \\ V^* &= \frac{1}{\mu+\theta} \left[\begin{array}{c} \left(\frac{\omega}{(R_0-1)+c\frac{R_0}{S^0}} + \frac{\xi(\mu+\gamma)}{\sigma} + q\gamma \right) I^* \\ + \frac{\omega c}{(R_0-1)+c\frac{R_0}{S^0}} \end{array} \right], \\ I^* &= \frac{\Lambda\left(R_0 - 1 + c\frac{R_0}{S^0}\right)(\theta+\mu)\sigma - \eta_1 c\mu\sigma}{\mu\left(\left(R_0 - 1 + c\frac{R_0}{S^0}\right)\left\{ \begin{array}{l} (\gamma+\mu)(\theta+\mu+\xi) \\ +(q\gamma+\theta+\mu)\sigma \end{array} \right\} + \sigma\eta_1\right)}. \end{aligned}$$

We conclude from the above that the endemic equilibrium point exists if $R_0 > 1$.

4 Stability Analysis of the Model System

We investigate the stability (linear as well as nonlinear) analysis of both the equilibrium points. The reduced limiting dynamical system is given by ([23])

$$\begin{cases} \frac{dS}{dt} = \frac{\Lambda}{\mu}(\mu+\theta) - \eta_1 S - \frac{\beta SI}{S+I+c} - \theta E - (\theta - p\gamma)I, \\ \frac{dE}{dt} = \frac{\beta SI}{S+I+c} - \eta_2 E, \\ \frac{dI}{dt} = \sigma E - (\mu+\gamma)I, \end{cases} \tag{2}$$

with initial conditions: $S(0) = S_0 > 0, E(0) = E_0 > 0, I(0) = I_0 > 0$. All the parameters are positive. Also $\eta_1 = (\mu+\omega+\theta)$, $\eta_2 = (\mu+\xi+\sigma)$ and $p = 1-q$.

Now, the local stability for worm-free equilibrium (WFE) point is established as follows:

Theorem 1. *The WFE point P^0 is*

1) Locally asymptotically stable, if $R_0 < 1$,

2) Unstable, if $R_0 > 1$ and

3) A transcritical bifurcation occurs at $R_0 = 1$.

Proof. The Jacobian matrix J_0 at WFE is given by

$$J_0 = \begin{bmatrix} -\eta_1 & -\theta & -\frac{\beta S^0}{S^0+c} - \theta + p\gamma \\ 0 & -\eta_2 & \frac{\beta S^0}{S^0+c} \\ 0 & \sigma & -(\mu+\gamma) \end{bmatrix}.$$

The characteristic equation of J_0 is given by

$$(\lambda+\eta_1)[\lambda^2 + (\mu+\eta_2+\gamma)\lambda + \eta_2(\mu+\gamma)(1-R_0)] = 0.$$

One eigenvalue is clearly negative; remaining two eigenvalues depends on the sign of the basic reproduction number. If $R_0 < 1$, then remaining two eigenvalues of J_0 have negative real parts and if $R_0 > 1$, then one eigenvalue of J_0 has negative real part and other has positive real part. Hence, WFE is locally asymptotically stable, if $R_0 < 1$ and unstable, if $R_0 > 1$. Now, if $R_0 = 1$, then two eigenvalues of J_0 have negative real parts and one eigenvalue is zero.

Let $x = S - S^0, y = E, z = I$. Then, system (2) reduces to

$$\begin{cases} \frac{dx}{dt} = -Ax - \frac{B(1+r)(x+S^0)z}{x+z+c+S^0} - \theta y - \theta z + (1-q)\gamma z, \\ \frac{dy}{dt} = \frac{B(1+r)(x+S^0)z}{x+z+c+S^0} - (\mu+\xi+\sigma)y, \\ \frac{dz}{dt} = \sigma y - (\mu+\gamma)z, \end{cases} \tag{3}$$

where

$$A = \eta_1, B = \frac{\eta_2(\gamma+\mu)(S^0+c)}{\sigma S^0}, R_0 = 1 + r.$$

For showing the occurrence of transcritical bifurcation at $(R_0, (S,E,I)) = (1, (S^0, 0, 0))$, we write β in terms of R_0 and other parameters. Linearizing system (3) about equilibrium point $(r, (x,y,z)) = (0, (0,0,0))$. we obtain the Jacobian matrix

$$\begin{bmatrix} -A & -\theta & (1-q)\gamma - \theta - \frac{BS^0}{S^0+c} \\ 0 & -\eta_2 & \frac{BS^0}{S^0+c} \\ 0 & \sigma & -\gamma-\mu \end{bmatrix}. \tag{4}$$

The proof is done by projecting the flow onto the extended center manifold [9]. The eigen-vectors corresponding to the eigenvalues $\lambda_1 = 0, \lambda_2 = -A$ and $\lambda_3 = -\gamma-\mu-\eta_2$ when $R_0 = 1 (r = 0)$ are

$$e_1 = \begin{bmatrix} -a_1 \\ a_3 \\ 1 \end{bmatrix}, e_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, e_3 = \begin{bmatrix} a_2 \\ -a_4 \\ 1 \end{bmatrix}$$

respectively, where

$$\begin{aligned} a_1 &= \frac{(\gamma+\mu)(\theta+\mu+\xi) + (q\gamma+\theta+\mu)\sigma}{A\sigma}, \\ a_2 &= \frac{(\gamma-\theta+\mu)(\mu+\xi) + (q\gamma+\mu)\sigma}{\sigma(-A+\gamma+\mu+\eta_2)}, \\ a_3 &= \frac{\gamma+\mu}{\sigma} \text{ and } a_4 = \frac{\eta_2}{\sigma}. \end{aligned}$$

The model matrix P with its column vector as the eigenvector is

$$P = \begin{bmatrix} -a_1 & 1 & a_2 \\ a_3 & 0 & -a_4 \\ 1 & 0 & 1 \end{bmatrix},$$

and hence

$$P^{-1} = \frac{1}{a_3+a_4} \begin{bmatrix} 0 & 1 & a_4 \\ a_3+a_4 & a_1+a_2 & -a_2a_3 + a_1a_4 \\ 0 & -1 & a_3 \end{bmatrix}.$$

Now, we have to find the nature of stability $(x,y,z) = (0,0,0)$ for r near zero. We obtain the transformation using the eigen basis $\{e_1, e_2, e_3\}$,

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = P \begin{bmatrix} u \\ v \\ w \end{bmatrix} \text{ with inverse } \begin{bmatrix} u \\ v \\ w \end{bmatrix} = P^{-1} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

which transform system (3) into

$$\left.\begin{aligned} \begin{bmatrix} \dot{u} \\ \dot{v} \\ \dot{w} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -A & 0 \\ 0 & 0 & -\gamma-\mu-\eta_2 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} \\ + \begin{bmatrix} f(u,v,w,r) \\ g_1(u,v,w,r) \\ g_2(u,v,w,r) \end{bmatrix} \end{aligned}\right\} \tag{5}$$

$$\dot{r} = 0. \tag{6}$$

Here

$$\begin{aligned} f(u,v,w,r) &= l_1u^2 + l_2w^2 + l_3uv + l_4uw + l_5ur + l_6vw \\ &\quad + l_7wr + l_8u^3 + l_9w^3 + l_{10}uvw + l_{11}uw^2 \\ &\quad + l_{12}u^2v + l_{13}u^2w + l_{14}u^2r + l_{15}vw^2 \\ &\quad + l_{16}w^2r + l_{17}wv^2 + l_{18}uvr + l_{19}vwr \\ &\quad + l_{20}uwr + l_{21}uv^2, \end{aligned}$$

$$
\begin{aligned}
g_1(u,v,w,r) &= m_2u^2+m_3w^2+m_4uv+m_5uw \\
&\quad +m_6ur+m_7vw+m_8wr+m_9u^3 \\
&\quad +m_{10}w^3+m_{11}uvw+m_{12}uw^2 \\
&\quad +m_{13}u^2v+m_{14}u^2w+m_{15}u^2r \\
&\quad +m_{16}vw^2+m_{17}w^2r+m_{18}wv^2 \\
&\quad +m_{19}uvr+m_{20}vwr+m_{21}uw \\
&\quad +m_{22}uv^2, \\
g_2(u,v,w,r) &= n_2u^2+n_3w^2+n_4uv+n_5uw+n_6ur \\
&\quad +n_7vw+n_8wr+n_9u^3+n_{10}w^3 \\
&\quad +n_{11}uvw+n_{12}uw^2+n_{13}u^2v+n_{14}u^2w \\
&\quad +n_{15}u^2r+n_{16}vw^2+n_{17}w^2r+n_{18}wv^2 \\
&\quad +n_{19}uvr+n_{20}vwr+n_{21}uwr+n_{22}uv^2.
\end{aligned}
$$

Where

$$
\begin{aligned}
l_1 &= -\frac{B(S^0+ca_1)}{(c+S^0)^2(a_3+a_4)}, \\
l_2 &= \frac{B(-S^0+ca_2)}{(c+S^0)^2(a_3+a_4)}, \\
l_3 &= l_6=\frac{Bc}{(c+S^0)^2(a_3+a_4)}, \\
l_4 &= -\frac{B(2S^0+ca_1-ca_2)}{(c+S^0)^2(a_3+a_4)}, \\
l_5 &= l_7=\frac{BS^0}{(a_3+a_4)(c+S^0)}, \\
l_8 &= -\frac{B(-1+a_1)(S^0+ca_1)}{(c+S^0)^3(a_3+a_4)}, \\
l_9 &= \frac{B(1+a_2)(S^0-ca_2)}{(c+S^0)^3(a_3+a_4)}, \\
l_{10} &= \frac{2B(-c+S^0+ca_1-ca_2)}{(c+S^0)^3(a_3+a_4)}, \\
l_{11} &= \frac{B(3S^0-a_2(2c-2S^0+ca_2)+a_1(c-S^0+2ca_2))}{(c+S^0)^3(a_3+a_4)}, \\
l_{13} &= \frac{B(3S^0-ca_1^2+(-c+S^0)a_2+2a_1(c-S^0+ca_2))}{(c+S^0)^3(a_3+a_4)}, \\
l_{12} &= \frac{B(-c+S^0+2ca_1)}{(c+S^0)^3(a_3+a_4)}, \\
l_{14} &= -\frac{B(S^0+ca_1)}{(c+S^0)^2(a_3+a_4)}, \\
l_{15} &= \frac{B(-c+S^0-2ca_2)}{(c+S^0)^3(a_3+a_4)}, \\
l_{16} &= \frac{B(-S^0+ca_2)}{(c+S^0)^2(a_3+a_4)}, \\
l_{17} &= l_{21}=-\frac{Bc}{(c+S^0)^3(a_3+a_4)}, \\
l_{18} &= l_{19}=\frac{Bc}{(c+S^0)^2(a_3+a_4)}, \\
l_{20} &= -\frac{B(2S^0+ca_1-ca_2)}{(c+S^0)^2(a_3+a_4)},
\end{aligned}
$$

$$
\begin{aligned}
m_1 &= -A, \\
m_2 &= -\frac{B(S^0+ca_1)(a_1+a_2-a_3-a_4)}{(c+S^0)^2(a_3+a_4)}, \\
m_3 &= \frac{B(-S^0+ca_2)(a_1+a_2-a_3-a_4)}{(c+S^0)^2(a_3+a_4)}, \\
m_4 &= m_7=\frac{Bc(a_1+a_2-a_3-a_4)}{(c+S^0)^2(a_3+a_4)}, \\
m_5 &= -\frac{B(2S^0+ca_1-ca_2)(a_1+a_2-a_3-a_4)}{(c+S^0)^2(a_3+a_4)}, \\
m_6 &= m_8=\frac{BS^0(a_1+a_2-a_3-a_4)}{(c+S^0)(a_3+a_4)}, \\
m_9 &= \frac{B(1-a_1)(S^0+ca_1)(a_1+a_2-a_3-a_4)}{(c+S^0)^3(a_3+a_4)}, \\
m_{10} &= \frac{B(1+a_2)(S^0-ca_2)(a_1+a_2-a_3-a_4)}{(c+S^0)^3(a_3+a_4)}, \\
m_{11} &= \frac{2B(-c+S^0+ca_1-ca_2)(a_1+a_2-a_3-a_4)}{(c+S^0)^3(a_3+a_4)}, \\
m_{12} &= \frac{\left(\begin{array}{l} B(3S^0-a_2(2c-2S^0+ca_2) \\ +a_1(c-S^0+2ca_2))(a_1+a_2-a_3-a_4) \end{array}\right)}{(c+S^0)^3(a_3+a_4)}, \\
m_{13} &= \frac{B(-c+S^0+2ca_1)(a_1+a_2-a_3-a_4)}{(c+S^0)^3(a_3+a_4)}, \\
m_{14} &= -\frac{\left(\begin{array}{l} B(-3S^0+ca_1^2+(c-S^0)a_2 \\ -2a_1(c-S^0+ca_2))(a_1+a_2-a_3-a_4) \end{array}\right)}{(c+S^0)^3(a_3+a_4)}, \\
m_{15} &= -\frac{B(S^0+ca_1)(a_1+a_2-a_3-a_4)}{(c+S^0)^2(a_3+a_4)}, \\
m_{16} &= -\frac{B(c-S^0+2ca_2)(a_1+a_2-a_3-a_4)}{(c+S^0)^3(a_3+a_4)}, \\
m_{17} &= \frac{B(-S^0+ca_2)(a_1+a_2-a_3-a_4)}{(c+S^0)^2(a_3+a_4)}, \\
m_{18} &= m_{22}=-\frac{Bc(a_1+a_2-a_3-a_4)}{(c+S^0)^3(a_3+a_4)}, \\
m_{19} &= m_{20}=m_{21}=\frac{Bc(a_1+a_2-a_3-a_4)}{(c+S^0)^2(a_3+a_4)}, \\
n_1 &= -\gamma-2\mu-\xi-\sigma, \\
n_2 &= \frac{B(S^0+ca_1)}{(c+S^0)^2(a_3+a_4)}, \\
n_3 &= \frac{B(S^0-ca_2)}{(c+S^0)^2(a_3+a_4)}, \\
n_4 &= n_7=-\frac{Bc}{(c+S^0)^2(a_3+a_4)}, \\
n_5 &= \frac{B(2S^0+ca_1-ca_2)}{(c+S^0)^2(a_3+a_4)}, \\
n_6 &= n_8=-\frac{BS^0}{(c+S^0)(a_3+a_4)}, \\
n_9 &= \frac{B(-1+a_1)(S^0+ca_1)}{(c+S^0)^3(a_3+a_4)},
\end{aligned}
$$

$$
\begin{aligned}
n_{10} &= \frac{B(1+a_2)(-S^0+ca_2)}{(c+S^0)^3(a_3+a_4)},\\
n_{11} &= -\frac{2B(-c+S^0+ca_1-ca_2)}{(c+S^0)^3(a_3+a_4)},\\
n_{12} &= \frac{\left(\begin{array}{l} B(-3S^0+a_1(-c+S^0-2ca_2)\\ +a_2(2c-2S^0+ca_2)) \end{array}\right)}{(c+S^0)^3(a_3+a_4)},\\
n_{13} &= -\frac{B(-c+S^0+2ca_1)}{(c+S^0)^3(a_3+a_4)},\\
n_{14} &= \frac{B(-3S^0+ca_1^2+(c-S^0)a_2-2a_1(c-S^0+ca_2))}{(c+S^0)^3(a_3+a_4)},\\
n_{15} &= \frac{B(S^0+ca_1)}{(c+S^0)^2(a_3+a_4)},\\
n_{16} &= \frac{B(c-S^0+2ca_2)}{(c+S^0)^3(a_3+a_4)},\\
n_{17} &= \frac{B(S^0-ca_2)}{(c+S^0)^2(a_3+a_4)},\\
n_{18} &= n_{22}=\frac{Bc}{(c+S^0)^3(a_3+a_4)},\\
n_{19} &= n_{20}=-\frac{Bc}{(c+S^0)^2(a_3+a_4)},\\
n_{21} &= \frac{B(2S^0+ca_1-ca_2)}{(c+S^0)^2(a_3+a_4)}.
\end{aligned}
$$

Let

$$
g(u,v,w,r)=\left[\begin{array}{l} \mathrm{g}_1(u,v,w,r)\\ \mathrm{g}_2(u,v,w,r)\end{array}\right].
$$

We have from the existence theorem for center manifolds

$$
W^c(0)=\left\{\begin{array}{l}(u,v,w,r)\in\mathbb{R}^4 \big| v=h_1(u,r), w=h_2(u,r),\\ h_i(0,0)=0, Dh_i(0,0)=0, i=1,2\end{array}\right\}
$$

for u and r sufficiently small.

Let

$$
\delta\equiv u,\zeta\equiv(v,w),h=(h_1,h_2),Q=0
$$

and

$$
R=\left[\begin{array}{ll} -A & 0\\ 0 & -\gamma-2\mu-\xi-\sigma\end{array}\right].
$$

Assume that

$$
\left.\begin{array}{l} h_1(u,r)=c_1u^2+c_2ur+c_3r^2+\cdots,\\ h_2(u,r)=c_4u^2+c_5ur+c_6r^2+\cdots\end{array}\right\}. \tag{7}
$$

Using invariance of the graph of $h(u,r)$ under the dynamics generated by Equation (3), $h(u,r)$ must satisfy

$$
\begin{aligned}
\mathbb{N}(h(\delta,r)) &= D_\delta h(\delta,r)[Q\delta+f(\delta,h(\delta,r)r)]\\
&\quad -Rh(\delta,r)-g(\delta,h(\delta,r)r)\\
&= 0. \qquad (8)
\end{aligned}
$$

Substitute $h=(h_1,h_2)$ from Equation (7) into Equation (8) and then compare the coefficients of u^2, ur and r^2, we obtain

$$
c_1=-\frac{m_2}{m_1},c_2=-\frac{m_6}{m_1},c_3=0,c_4=-\frac{n_2}{n_1},c_5=-\frac{n_6}{n_1},c_6=0.
$$

Hence

$$
h_1(u,r)=-\frac{m_2}{m_1}u^2-\frac{m_6}{m_1}ru, h_2(u,r)=-\frac{n_2}{n_1}u^2-\frac{n_6}{n_1}ru.
$$

Finally substituting the values of $v=h_1, w=h_2$ into Equations (5) and (6) we obtain the vector field reduced to the center manifold

$$
\begin{aligned}
\dot{u} &= r\,u\,l_5+u^2\,l_1+u^3\left(l_8-\frac{l_3\,m_2}{m_1}-\frac{l_4\,n_2}{n_1}\right)\\
&\quad +ru^2\left(l_{14}-\frac{l_3\,m_6}{m_1}-\frac{l_4\,n_6}{n_1}-\frac{l_7\,n_2}{n_1}\right)+\cdots,\\
\dot{r} &= 0.
\end{aligned}
$$

Here we observe that $l_1<0$ and $l_5>0$. On the center manifold, we have

$$
\frac{du}{dt}=G(u,r)=rul_5+u^2l_1,
$$

with

$$
\begin{aligned}
G(0,0) &= G_u(0,0)=G_r(0,0),\\
G_{uu} &= 2l_1,\\
G_{ur} &= l_5,\\
G_{rr} &= 0.
\end{aligned}
$$

Here G_{ur} is positive and G_{uu} is negative. Hence, using transcritical bifurcation and center manifold theorems, worm-free equilibrium point is stable when $R_0<1$(since $r<0$) and there is a separate unstable branch from the endemic equilibrium point and when $R_0>1$(since $r>0$), worm-free equilibrium point becomes unstable while the separating branch becomes stable [9]. When $R_0=1$(since $r=0$), center manifold is approximated by

$$
\frac{du}{dt}\approx l_1u^2+\cdots.
$$

Therefore, the worm-free equilibrium point is stable if it is approached from $u>0$.

Hence, transcritical bifurcation occurs at the bifurcation point $R_0=1$. □

Theorem 2. *The endemic equilibrium point P^* is locally asymptotically stable if $\sigma\beta(S^*+c)S^*\le\eta_2(\gamma+\mu)(S^*+I^*+c)^2$ holds.*

Proof. The Jacobian matrix J^* at endemic equilibrium point P^* is given by

$$
J^*=\left[\begin{array}{ccc} -\eta_1-a_{21} & -\theta & -\theta+p\gamma-a_{23}\\ a_{21} & -\eta_2 & a_{23}\\ 0 & \sigma & -\mu-\gamma\end{array}\right],
$$

where

$$
a_{21}=\frac{\beta(I^*+c)I^*}{(S^*+I^*+c)^2}, a_{23}=\frac{\beta(S^*+c)S^*}{(S^*+I^*+c)^2}.
$$

The characteristic equation of J^* is

$$
\lambda^3+A_1\lambda^2+A_2\lambda+A_3=0,
$$

where

$$\begin{aligned} A_1 &= a_{21}+\gamma+\mu+\eta_1+\eta_2, \\ A_2 &= \theta(\mu+\eta_2)+\mu(\mu+2\eta_2)-a_{23}\sigma \\ &\quad +a_{21}(\gamma+\theta+\mu+\eta_2)+(\mu+\eta_2)\omega \\ &\quad +\gamma(\eta_1+\eta_2), \\ A_3 &= a_{21}(\gamma+\mu)(\theta+\mu+\xi)+a_{21}(q\gamma+\theta+\mu)\sigma \\ &\quad +\eta_1(-a_{23}\sigma+\eta_2(\gamma+\mu)), \end{aligned}$$

and

$$\begin{aligned} A_1A_2-A_3 &= a_{21}^2(\gamma+\theta+\mu+\eta_2) \\ &+(\gamma+\mu+\eta_2)(-a_{23}\sigma+(\gamma+\mu+\eta_1)(\eta_1+\eta_2)) \\ &+a_{21}(\gamma^2+\theta^2-a_{23}\sigma+\theta(2(2\mu+\eta_2)+\xi+\omega) \\ &+(\mu+\eta_2)(3\mu+\eta_2+2\omega)+2\gamma(\mu+\eta_1+\eta_2+\frac{p}{2}\sigma)). \end{aligned}$$

We observe that $A_1>0$ automatically satisfies. Both the conditions $A_3>0$ and $A_1A_2-A_3>0$ satisfies if $\sigma a_{23}\leq(\gamma+\mu)\eta_2$ implies that $\sigma\beta(S^*+c)S^*\leq\eta_2(\gamma+\mu)(S^*+I^*+c)^2$ holds. Hence, by Routh-Hurwitz criterion the endemic equilibrium point P^* is locally asymptotically stable. □

4.1 Global Stability Analysis

We analyze the global dynamics of worm-free and endemic equilibrium points. We find first the global stability of worm-free equilibrium point using the method developed in [3]. Rewrite the model system (2) as

$$\begin{cases} \frac{dY}{dt}=F(Y,Z), \\ \frac{dZ}{dt}=G(Y,Z),\ G(Y,0)=0. \end{cases} \tag{9}$$

where $Y=(S)$ and $Z=(E,I)$, with $Y\in\mathbb{R}$ denoting the number of susceptible node and $Z\in\mathbb{R}^2$ denoting the number of infected nodes (exposed and infectious). The worm-free equilibrium point is denoted by $Q_0=\left(Y^0,0\right)$. The following conditions $(A1)$ and $(A2)$ must give a local asymptotic stability:

(A1) Y^0 is globally asymptotic stable for $\frac{dY}{dt}=F(Y,0)$.

(A2) $G(Y,Z)=BZ-\hat{G}(Y,Z)$ where $\hat{G}(Y,Z)\geq 0$ for $(Y,Z)\in U$,

and $B=D_zG\left(Y^0,0\right)$ is an M-matrix and U is the region of attraction. Then the following lemma holds.

Lemma 1. *The fixed point $Q_0=(Y^0,0)$ is a globally asymptotic stable equilibrium point of (9) if $R_0<1$ and Assumptions (A1) and (A2) are satisfied.*

Theorem 3. *Suppose $R_0<1$, then the worm-free equilibrium point P^0 is globally asymptotically stable.*

Proof. Let $Y=(S)$, $Z=(E,I)$ and $Q_0=(Y^0,0)$, where

$$Y^0=\frac{\Lambda}{\mu}\left(\frac{\mu+\theta}{\eta_1}\right). \tag{10}$$

Then,

$$\begin{aligned} &\frac{dY}{dt}=F(Y,Z) \\ &=\frac{\Lambda}{\mu}(\mu+\theta)-\eta_1S-\frac{\beta SI}{S+I+c}-\theta E-(\theta-(1-q)\gamma)I. \end{aligned}$$

At $S=S^0, F(Y,0)=0$, and

$$\frac{dY}{dt}=F(Y,0)=\frac{\Lambda}{\mu}(\mu+\theta)-\eta_1Y.$$

As $t\to\infty, Y\to Y^0$.

Hence $Y=Y^0(=S^0)$ is globally asymptotically stable. Now,

$$\begin{aligned} G(Y,Z) &= \begin{bmatrix} -\eta_2 & \beta S^0 \\ \sigma & -(\mu+\gamma) \end{bmatrix}\begin{bmatrix} E \\ I \end{bmatrix} \\ &\quad -\begin{bmatrix} \beta S^0I-\frac{\beta SI}{S+I+c} \\ 0 \end{bmatrix}, \\ &= BZ-\hat{G}(Y,Z), \end{aligned}$$

where $B=\begin{bmatrix} -\eta_2 & \beta S^0 \\ \sigma & -(\mu+\gamma) \end{bmatrix}$ and

$\hat{G}(Y,Z)=\begin{bmatrix} \beta S^0I-\frac{\beta SI}{S+I+c} \\ 0 \end{bmatrix}$.

In model system (2), total number of nodes is bounded by $N_1^0=\frac{\Lambda}{\mu}\frac{\mu+\theta}{\mu+\theta+\eta}$ where $\eta=\max\{\omega,\xi,q\gamma\}$, that is, $(S,E,I)\leq N_1^0$.

Since $S^0\geq N_1^0$, we have $S^0\geq N_1^0\geq S\geq\frac{S}{S+I+c}$ and thus, $\hat{G}(Y,Z)\geq 0$. Therefore, B is an M-matrix. Hence $(A1)$ and $(A2)$ are satisfied and by Lemma 1, P^0 is globally asymptotically stable if $R_0<1$. □

Following Li and Muldowney [15], we obtain sufficient conditions for global asymptotic stability of the endemic equilibrium point. Consider the autonomous dynamical system:

$$\dot{x}=f(x)\ \text{ with } x(0,x_0)=x_0 \tag{11}$$

where $f:D\to\mathbb{R}^n, D\subset\mathbb{R}^n$ open set and simply connected and $f\in C^1(D)$. Let x^* be an equilibrium point of Equation (11) that is, $f(x^*)=0$. Assume that the following conditions hold:

(A3) There exists a compact absorbing set $K\subset D$.

(A4) Equation (11) has a unique equilibrium point x^* in D.

We know that if x^* is locally stable and all trajectories in D converges to x^* then it is to be globally stable in D. Bendixon criterion rule out the existence of non-constant periodic solutions of Equation (11) for $n\geq 2$, that conditions satisfied by f. The classical Bendixson's condition $\mathrm{div}f(x)<0$ for $n=2$, is robust under C^1 local perturbations of f.

Lemma 2. *Assume that conditions (A3), (A4) hold and Equation (11) satisfies a Bendixson criterion. Then, x^* is globally stable in D, provided it is stable.*

Let us consider $P(x)$ as $\begin{pmatrix} n \\ 2 \end{pmatrix} \times \begin{pmatrix} n \\ 2 \end{pmatrix}$ matrix-valued function that is C^1 on D and the matrix P_f has components

$$(p_{ij}(x))_f = \left(\frac{\partial p_{ij}(x)}{\partial x}\right)^T .f(x) = \nabla p_{ij}.f(x).$$

Assume that P^{-1} exists and is continuous for $x \in K$ (the compact absorbing set). The matrix $J^{[2]}$ is the second additive compound matrix of the Jacobian matrix J, that is $J(x) = Df(x)$. When $n = 3$, the second additive compound matrix of $J = (a_{ij})$ is given by [20],

$$J^{[2]} = \begin{bmatrix} a_{11}+a_{22} & a_{23} & -a_{13} \\ a_{32} & a_{11}+a_{33} & a_{12} \\ -a_{31} & a_{21} & a_{22}+a_{33} \end{bmatrix}.$$

Let $\mu(B)$ be the Lozinskii measure of B with respect to a vector norm $|.|$ in $\mathbb{R}^N$.$N = \begin{pmatrix} n \\ 2 \end{pmatrix}$, defined by $\mu(B) = \lim_{h\to 0^+} \frac{|I+hB|-1}{h}$. A quantity $\bar{q}$ is defined as

$$\bar{q} = \limsup_{t\to\infty} \sup_{x\in K} \frac{1}{t}\int_0^t \mu\left(B\left(x\left(s, x_0\right)\right)\right) ds,$$

where

$$B = P_f P^{-1} + PJ^{[2]}P^{-1}.$$

If D is simply connected, $\bar{q} < 0$ rules out the presence of any orbit that gives rise to simple closed rectifiable curve that is invariant for Equation (11) and robust under C^1 local perturbations of f near any non-equilibrium point that is non-wandering. The following global stability result is given in Li and Muldowney [15].

Lemma 3. *Assume that D is simply connected and assumptions (A3) and (A4) hold. Then the unique equilibrium point x^* of Equation (11) is globally stable in D if $\bar{q} < 0$.*

Now, we analyze the global stability of the endemic equilibrium point x^*.

Theorem 4. *If $R_0 > 1, \xi + \sigma < \omega$ and $(1-q)\gamma \geq \theta$, then the endemic equilibrium P^* of the model system (2) is globally stable in U.*

Proof. From Theorem 2, if the endemic equilibrium point P^* exists, is locally asymptotically stable. From Theorem 1, when $R_0 > 1$, P^0 is unstable. The instability of P^0 together with $P^0 \in \partial U$ is implies to the uniform persistence [6], that is there exists a constant $C > 0$ such that:

$$\liminf_{t\to\infty} x(t) > C, x = (S, E, I).$$

The uniform persistence together with the boundedness of U, is equivalent to the existence of a compact set in the interior of U which is absorbing for the model system (2) [10]. Thus, $(A3)$ is verified. Now, the second additive compound matrix $J^{[2]}(S, E, I)$ is given by

$$J^{[2]} = \begin{bmatrix} \Phi_1 & \frac{\beta(S+c)S}{(S+I+c)^2} & \frac{\beta(S+c)S}{(S+I+c)^2} + \theta - (1-q)\gamma \\ \sigma & \Phi_2 & -\theta \\ 0 & \frac{\beta(I+c)I}{(S+I+c)^2} & -(\mu+\gamma+\eta_2) \end{bmatrix}$$

where

$$\begin{aligned} \Phi_1 &= -(\eta_1+\eta_2) - \frac{\beta(I+c)I}{(S+I+c)^2}, \\ \Phi_2 &= -(\mu+\gamma+\eta_1) - \frac{\beta(I+c)I}{(S+I+c)^2}. \end{aligned}$$

Let us consider

$$P = P(S, E, I) = diag\left\{1, \frac{E}{I}, \frac{E}{I}\right\}.$$

Therefore,

$$P_f P^{-1} = diag\left\{0, \frac{\dot{E}}{E} - \frac{\dot{I}}{I}, \frac{\dot{E}}{E} - \frac{\dot{I}}{I}\right\}.$$

Then $B = P_f P^{-1} + PJ^{[2]}P^{-1}$

$$= \begin{bmatrix} \Phi_1 & \frac{\beta(S+c)S}{(S+I+c)^2}\frac{I}{E} & \left(\frac{\beta(S+c)S}{(S+I+c)^2} + \theta - (1-q)\gamma\right)\frac{I}{E} \\ \sigma\frac{E}{I} & \frac{\dot{E}}{E} - \frac{\dot{I}}{I} + \Phi_2 & -\theta \\ 0 & \frac{\beta(I+c)I}{(S+I+c)^2} & \frac{\dot{E}}{E} - \frac{\dot{I}}{I} - (\mu+\gamma+\eta_2) \end{bmatrix}.$$

Let

$$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix},$$

where

$$B_{11} = \Phi_1 = -\left(\eta_1+\eta_2+\frac{\beta(I+c)I}{(S+I+c)^2}\right),$$

$$B_{12} = \begin{bmatrix} \frac{\beta(S+c)S}{(S+I+c)^2}\frac{I}{E} & \left(\frac{\beta(S+c)S}{(S+I+c)^2} + \theta - (1-q)\gamma\right)\frac{I}{E} \end{bmatrix},$$

$$B_{21} = \begin{bmatrix} \sigma\frac{E}{I} \\ 0 \end{bmatrix},$$

$$B_{22} = \begin{bmatrix} \frac{\dot{E}}{E} - \frac{\dot{I}}{I} - \frac{\beta(I+c)I}{(S+I+c)^2} & -\theta \\ -(\mu+\gamma+\eta_1) & \\ \frac{\beta(I+c)I}{(S+I+c)^2} & \frac{\dot{E}}{E} - \frac{\dot{I}}{I} - (\mu+\gamma+\eta_2) \end{bmatrix}.$$

Now, consider the norm $|(u_1, u_2, u_3)| = \max\{|u_1|, |u_2|+|u_3|\}$ in $\mathbb{R}^3$, where (u_1, u_2, u_3) denotes vector in $\mathbb{R}^3$ and the Lozinskii measure is denoted by μ with respect to this norm [17].

$$\begin{aligned} \mu(B) &\leq \sup\{g_1, g_2\}, \\ &= \sup\left\{\mu_1(B_{11}) + |(B_{12})|, \mu_1(B_{22}) + |(B_{21})|\right\}. \end{aligned}$$

where $|B_{21}|\,|B_{12}|$ are matrix norms with respect to the L^1 vector norm, and μ_1 denotes the Lozinskii measure with respect to the L^1 norm.

Then

$$
\begin{aligned}
g_1 &= \mu_1(B_{11}) + |B_{12}|, \\
&= -\left(\eta_1+\eta_2+\frac{\beta(I+c)I}{(S+I+c)^2}\right)+\frac{\beta(S+c)S}{(S+I+c)^2}\frac{I}{E} \\
&\quad + \frac{I}{E}\max\{0,\theta-(1-q)\gamma\}, \\
&\le -\left(\eta_1+\eta_2+\frac{\beta(I+c)I}{(S+I+c)^2}\right)+\frac{\beta(S+c)S}{(S+I+c)^2}\frac{I}{E}, \\
&\qquad [\text{when } (1-q)\gamma \ge \theta] \\
&\le -\eta_1+\frac{\dot{E}}{E}-\left[\frac{\beta S}{S+I+c}-\frac{\beta(S+c)S}{(S+I+c)^2}\right]\frac{I}{E}, \\
&[\text{from steady state of second equation of model system (2)}] \\
&= \frac{\dot{E}}{E}-\eta_1-\frac{\beta SI}{(S+I+c)^2}\frac{I}{E}, \\
&\le \frac{\dot{E}}{E}-\mu.
\end{aligned}
$$

Again,

$$
\begin{aligned}
g_2 &= |B_{21}| + \mu_1(B_{22}), \\
&= \sigma\frac{E}{I}+\frac{\dot{E}}{E}-\frac{\dot{I}}{I}-2\mu-\gamma-\theta \\
&\quad +\max\{-\omega,-(\sigma+\xi)\}, \\
&= \frac{\dot{E}}{E}-\mu-\theta+\max\{-\omega,-(\sigma+\xi)\}, \\
&[\text{from steady state of third equation of model system (2)}] \\
&\le \frac{\dot{E}}{E}-\mu-\theta, \quad [\text{when } \xi+\sigma<\omega] \\
&\le \frac{\dot{E}}{E}-\mu.
\end{aligned}
$$

Therefore,

$$\mu(B) \le \sup\{g_1, g_2\} = \frac{\dot{E}}{E}-\mu.$$

Along each solution $(S(t), E(t), I(t))$ of the system with $(S(0), E(0), I(0)) \in K$, where K is the compact absorbing set, we have

$$\frac{1}{t}\int_0^t \mu(B)\,ds \le \frac{1}{t}\log\frac{E(t)}{E(0)}-\mu,$$

which implies that

$$\bar{q} = \limsup_{t\to\infty} \sup_{x_0\in U} \frac{1}{t}\int_0^t \mu(B(x(s,x_0)))\,ds \le -\mu < 0.$$

Therefore, $\bar{q} < 0$ and thus Bendixson criteria is also fulfilled. Hence the global stability of the endemic or worm-induced equilibrium point has established. □

5 Numerical Simulations

Numerically, Runge-Kutta method is used to simulate the model system (2) using MATLAB software. The dynamical behaviors of all the three classes S, E and I are observed by considering a set of parameter values and initial conditions. We have taken initial condition $(22, 20, 20)$ and parametric values

$$\Lambda = 0.4, \mu = \frac{1}{(65*365)}, \theta = \frac{1}{(2*365)}, \gamma = \frac{1}{30},$$
$$\omega = 0.6, \beta = 0.14, c = 10, q = 0.9, \xi = 0.2, \sigma = 0.1. \quad (12)$$

The endemic equilibrium point for the vaccination rate $\omega = 0.1$ is $(102.2687, 10.2488, 30.7065)$ and the worm-free equilibrium point for vaccination rate $\omega = 0.9$ is $(14.8677, 0, 0)$ and other parameter values are same as used in (12). The analysis of Figures 2 and 3 under different vaccination rate shows the stability of both worm-free and endemic equilibrium points that is, for the cases when $R_0 < \text{or} > 1$. We have critically examined the infectious class I for the different values of $\xi(= \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \frac{1}{6}, \frac{1}{7})$ which support the reality that as the self-prevention techniques of nodes ξ decreases, infection increases (Figure 4) and the increment in transmission rate $\beta(= 0.8, 0.12, 0.16, 0.20, 0.24)$ causes increment in infection (Figure 5) that is infected node increases and susceptible node decreases (Figure 6) and will be stable. It is observed from Figure 7, the values of transmission rate β decreases the susceptible nodes attain its saturation value, when $R_0 < 1$. We have also observed the evolutions of susceptible nodes for the fraction of recovered nodes gaining worm-acquired immunity, q (Figure 8) and observation tells that the susceptible node increases when the fraction of recovered nodes gaining worm-acquired immunity, $q(= 0.2, 0.4, 0.6, 0.8, 1.0)$ increases when $\beta = 0.12$ $(R_0 < 1)$ but when $\beta = 0.20$ $(R_0 > 1)$ then the susceptible nodes attain its saturation value. The above parametric values given in (12) satisfies our analytical results.

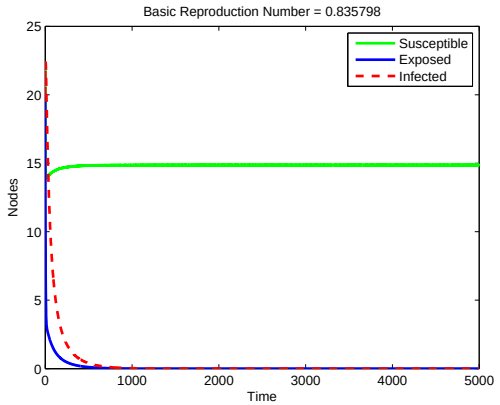

Figure 2: Time series of susceptible, exposed and infected nodes when $\omega = 0.9$

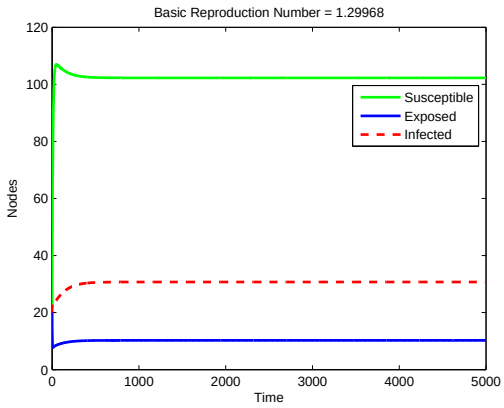

Figure 3: Time series of susceptible, exposed and infected nodes when $\omega = 0.1$

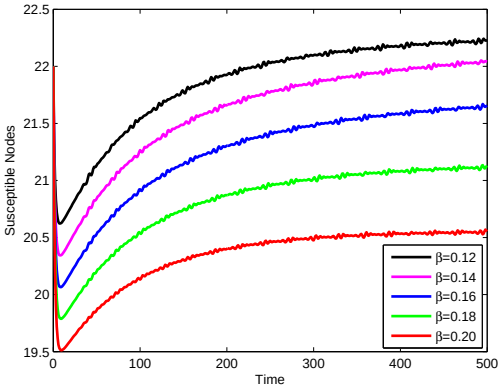

Figure 6: Dynamical behavior of susceptible class for different values of β when $R_0 > 1$

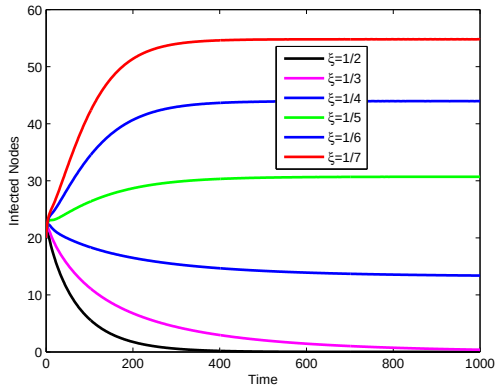

Figure 4: Dynamical behavior of infected class for different values of ξ when $\omega = 0.1$

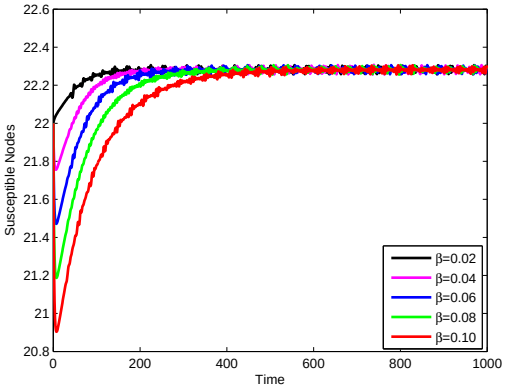

Figure 7: Dynamical behavior of susceptible class for different values of β when $R_0 < 1$

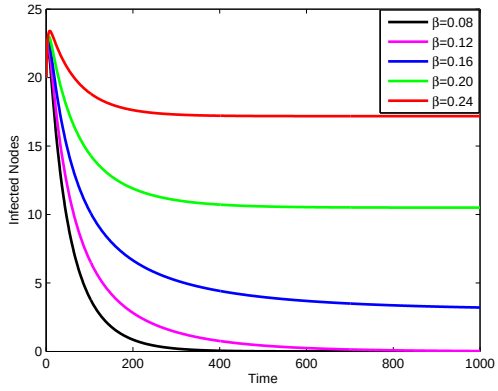

Figure 5: Dynamical behavior of infected class for different values of β

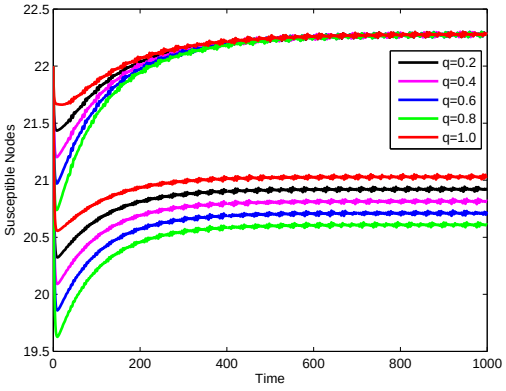

Figure 8: Dynamical behavior of susceptible nodes for different values of q when $\beta = 0.12$ ($R_0 < 1$) and $\beta = 0.20$ ($R_0 > 1$)

6 Discussions and Conclusions

An e-epidemic $SVEIS$ model with nonlinear incidence rate has been proposed for the transmission of worms in the computer network. Stability analysis and behavior of the reduced model system (2) have been investigated for both worm-free and endemic equilibrium points. Local stability analysis is established by using Routh-Hurwitz criterion. Characteristics of basic reproduction number have been discussed and found that if $R_0 < 1$, WFE point P^0 is globally asymptotically stable under certain conditions and worm declines from the computer network, where as if $R_0 > 1$, the worm-free equilibrium point is unstable and worm persists. In Figure 9, the forward transcritical bifurcation occurs at $R_0 = 1$ and it is effectively eradicate the worms. When the bifurcation parameter R_0, crosses the bifurcation threshold $R_0 = 1$, the endemic

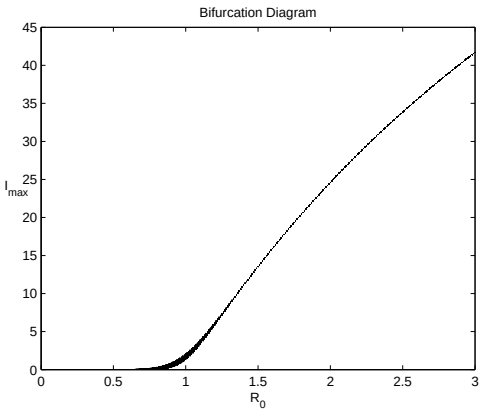

Figure 9: Transcritical Bifurcation diagram in the plane $(R_0, I_{\max})$

equilibrium point enters into the positive orthant. The vaccination rate reaches its critical value

$$\omega_c = \frac{(\mu+\theta)}{\mu c}\left(\frac{\sigma\beta\Lambda}{(\mu+\gamma)(\mu+\xi+\sigma)} - (\Lambda+\mu c)\right)$$

then the basic reproduction number $R_0 = 1$. We have observed the effect of the vaccination rate ω, on the basic reproduction number which ultimately affecting the dynamics of the model system. The optimal vaccine at critical level is most important factor to effectively eradicate the worm. Hence vaccination rate, ω must be greater than critical vaccination rate, $\omega_c = 0.531956$ to control worm from the network otherwise, worm persists in the network.

To study the affect of the parameter q, a fraction of recovered nodes on the dynamics of the model systems we find that it does not appear in the definition of R_0 and w_c. Due to waning of vaccination, a fraction of recovered nodes $(1-q)\gamma I$ moves to the susceptible class directly and rest via vaccinated class [23]. The effect of q on the susceptible node for transmission rate $\beta = 0.12$ and 0.20 is shown in Figure 8. In self-replicating computer worms modeling, the nonlinear incidence rate plays a major role and ensures that the model system can give a reasonable qualitative description of the worm dynamics.

References

[1] K. B. Blyuss and Y. N. Kyrychko, "Stability and bifurcations in an epidemic model with varying immunity period," *Bulletin of Mathematical Biology*, vol. 72, no. 2, pp. 490-505, 2010.

[2] V. Capasso and G. Serio, "A generalization of the Kermack-McKendrick deterministic epidemic model" *Mathematical Biosciences*, vol. 42, no. 1-2, pp. 43-61, 1978.

[3] C. Castillo-Chavez, Z. Feng and W. Huang, *Mathematical Approaches for Emerging and Re-Emerging Infectious Diseases: An Introduction*, Springer Verlag, 2002.

[4] P. V. D. Driessche and J. Watmough, "A simple SIS epidemic model with a backward bifurcation," *Journal of Mathematical Biology*, vol. 40, no. 6, pp. 525-540, 2000.

[5] B. Dubey, P. Dubey and U. S. Dubey, "Dynamics of an SIR model with nonlinear incidence and treatment rate," *Applications and Applied Mathematics: An International Journal (AAM)*, vol. 10, no. 2, pp. 718-737, 2015.

[6] H. I. Freedman, S. Ruan and M. Tang, "Uniform persistence and flows near a closed positively invariant set," *Journal of Dynamics and Differential Equations*, vol. 6, no. 4, pp. 583-600, 1994.

[7] C. Gan, X. Yang, W. Liu and Q. Zhu, "A propagation model of computer virus with nonlinear vaccination probability," *Communications in Nonlinear Science and Numerical Simulation*, vol. 19, no. 1, pp. 92-100, 2014.

[8] C. Gan, X. Yang, W. Liu, Q. Zhu and X. Zhang, "An epidemic model of computer viruses with vaccination and generalized nonlinear incidence rate," *Applied Mathematics and Computation*, vol. 222 pp. 265-274, 2013.

[9] P. Glendinning, *Stability, Instability and Chaos: An Introduction to the Theory of Nonlinear Differential Equations*, vol. 11, *Cambridge University Press*, 1994.

[10] V. Hutson and K. Schmitt, "Permanence and the dynamics of biological systems", *Mathematical Biosciences*, vol. 111, pp. 1-71, 1992.

[11] C. Jin and X. Wang, "Propagation model with varying population size of removable memory device virus," *International Journal of Network Security*, vol. 19, no. 4, pp. 507-516, 2017.

[12] A. Kaddar, "Stability analysis in a delayed SIR epidemic model with a saturated incidence rate," *Nonlinear Analysis: Modelling and Control*, vol. 15, no. 3, pp. 299-306, 2010.

[13] W. O. Kermack and A. G. McKendrick, "A contribution to the mathematical theory of epidemics," *In Proceedings of the Royal Society of London A: mathematical, physical and engineering sciences*, vol. 115, no. 772, pp. 700-721, The Royal Society, 1927.

[14] M. S. S. Khan, "A computer virus propagation model using delay differential equations with probabilistic contagion and immunity," *International Journal of Computer Networks and Communications*, vol. 6, no. 5, pp. 111-128, 2014.

[15] M. Y. Li and J. S. Muldowney, "A geometric approach to global-stability problems," *SIAM Journal on Mathematical Analysis*, vol. 27, no. 4, pp. 1070-1083, 1996.

[16] W. M. Liu, H. W. Hethcote and S. A. Levin, "Dynamical behavior of epidemiological models with nonlinear incidence rates," *Journal of Mathematical Biology*, vol. 25, no. 4, pp. 359-380, 1987.

[17] R. H. Martin, "Logarithmic norms and projections applied to linear differential systems," *Journal of*

Mathematical Analysis and Applications, vol. 45, no. 2, pp. 432-454, 1974.

[18] B. K. Mishra and S. K. Pandey, "Dynamic model of worm propagation in computer network," *Applied Mathematical Modelling*, vol. 38, no. 7, pp. 2173-2179, 2014.

[19] B. K. Mishra and S. K. Pandey, "Effect of anti-virus software on infectious nodes in computer network: a mathematical model," *Physics Letters A*, vol. 376, no. 35, pp. 2389-2393, 2012.

[20] J. Muldowney, "Compound matrices and ordinary differential equations," *Rocky Mountain Journal of Mathematics* vol. 20, no. 4, 1990.

[21] J. R. C. Piqueira, B. F. Navarro and L. H. A. Monteiro, "Epidemiological models applied to viruses in computer networks," *Journal of Computer Science*, vol. 1, no. 1, pp. 31-34, 2005.

[22] J. Ren, X. Yang, Q. Zhu, L. Yang and C. Zhang, "A novel computer virus model and its dynamics," *Nonlinear Analysis: Real World Applications*, vol. 13, no. 1, pp. 376-384, 2012.

[23] G. P. Sahu and J. Dhar, "Analysis of an SVEIS epidemic model with partial temporary immunity and saturation incidence rate," *Applied Mathematical Modelling*, vol. 36, no. 3, pp. 908-923, 2012.

[24] J. M. Tchuenche, S. A. Khamis, F. B. Agusto and S. C. Mpeshe, "Optimal control and sensitivity analysis of an influenza model with treatment and vaccination," *Acta Biotheoretica* vol. 59, no. 1, pp. 1-28, 2011.

[25] R. K. Upadhyay, S. Kumari and A. K. Misra, "Modeling the virus dynamics in computer network with SVEIR model and nonlinear incident rate," *Journal of Applied Mathematics and Computing*, vol. 54, no. 1-2, pp. 485509, 2017.

[26] F. Wang, H. Gao, Y. Yang and C. Wang, "An $SVEIR$ defending model with partial immunization for worms," *International Journal of Network Security*, vol. 19, no. 1, pp. 20-26, 2017.

[27] R. Xu and Z. Ma, "Global stability of a delayed SEIRS epidemic model with saturation incidence rate," *Nonlinear Dynamics*, vol. 61, no. 1, pp. 229-239, 2010.

[28] L. X. Yang, X. Yang, Q. Zhu and L. Wen, "A computer virus model with graded cure rates," *Nonlinear Analysis: Real World Applications*, vol. 14, no. 1, pp. 414-422, 2013.

[29] M. Zhang, G. Song and L. Chen, "A state feedback impulse model for computer worm control," *Nonlinear Dynamics*, vol. 85, no. 03, pp. 1561-1569, 2016.

[30] T. Zhang and Z. Teng, "Pulse vaccination delayed SEIRS epidemic model with saturation incidence," *Applied Mathematical Modelling*, vol. 32, no. 7, pp. 1403-1416, 2008.

Biography

Ranjit Kumar Upadhyay is a professor of Applied Mathematics, Indian Institute of Technology (Indian School of Mines), India. He has published more than 110 papers in different international journals of repute and a book "Introduction to Mathematical Modeling and Chaotic dynamics" from CRC Press, Taylor and Francis. He is in the editorial board and guest editor of the special issues of different international journals. His research areas include chaotic dynamics of real world situations, population dynamics, reaction-diffusion modeling, environmental modeling, virus dynamics and neural modeling. He visited Eötvös Lorand University, Institute of Biology, Department of Plant Taxonomy and Ecology, Isaac Newton Institute for Mathematical Sciences, Cambridge and Applied Mathematics Laboratory, University of Le Havre Normandie, France for academic purposes.

Sangeeta Kumari received her M.Phil degree in Applied Mathematics (2015) from Indian School of Mines, India. Currently she is a research scholar at Indian Institute of Technology (Indian School of Mines), India. Her research interests include Mathematical Modeling, e-Epidemiology and Dynamical System Theory.

A Reusable Multipartite Secret Sharing Scheme Based on Superincreasing Sequence

Putla Harsha, Patibandla Chanakya, and Vadlamudi China Venkaiah
(Corresponding author: Patibandla Chanakya)

School of Computer and Information Sciences & University of Hyderabad
Hyderabad-500046, India
(Email: chanakya.patibandla@gmail.com)
(Received Nov. 1, 2016; revised and accepted Feb. 20, 2017)

Abstract

A new multipartite secret sharing scheme that uses a super increasing sequence is proposed in this paper. Novelty of the scheme is that, apart from being a multipartite scheme, it realizes the level ordered access structure [17]. Also, the proposed scheme is reusable in that the shares of the participants need not be replenished for a new secret after the reconstruction the current secret.

Keywords: Multipartite Secret Sharing; Secret Sharing Schemes; Superincreasing Sequence

1 Introduction

A *Secret Sharing Scheme* consists of two phases: share distribution and secret reconstruction. In share distribution phase, each participant is given a share of the secret. In secret reconstruction phase, authorized subsets of the set of participants collaboratively recover the secret from their shares. A secret sharing scheme is said to be perfect if an unauthorized subset gets no information about the secret and an authorized subset has complete information of the secret in the information theoretic sense. The set of authorized subsets is called an access structure. If the maximal length of the shares and the length of the secret are identical, then such secret sharing scheme is said to be ideal. Several access structures like (t, n)-threshold access structure, Multipartite access structure, Generalized access structure *etc.*, are proposed in the literature.

A $(t, n)-$*Threshold Secret Sharing Scheme* is one in which at least the threshold (t) number of participants participate in secret reconstruction phase to get the secret. In threshold secret sharing scheme, all participants are given equal priority and entrust. Threshold secret sharing schemes were first proposed independently by Adi Shamir [18] and George Blakley [2] in 1979.

In practice, participants may be served with different priorities. So each participant is assigned a level. Based on the level of the participant, the priority, entrust of the participant varies. In general, the participants in the higher level get higher priority and entrust over the participants in the lower level.

In *Multipartite Secret Sharing*, the set of participants are divided into disjoint levels (parts/ compartments) and all participants in a particular level/compartment play the same role as all other participants in that level/compartment. Compartmented threshold secret sharing and multilevel threshold secret sharing are multipartite in nature. Ghodosi *et al.* [9] proposed an ideal and perfect secret sharing schemes for multilevel and compartmented groups. The scheme we are proposing is a multipartite scheme along with a restriction that the low level participants can recover the implicit compartment (level) secret bit assigned to them only if the implicit secret bits assigned to all higher levels are already recovered. The actual secret of our multipartite scheme can be recovered only after completion of the recovery of all the compartment secrets. Compartment secrets can be concurrently reconstructed,but in order to reconstruct the actual secret bit vector/tuple, hierarchy has to be followed. We use Shamir's secret sharing scheme for each compartment to distribute and recover the compartment secret to all the participants of the compartment. Applications of secret sharing can be found in [1, 5, 14, 15, 21] *etc.*

Our scheme supports the property of secret changeability without changing the compartment secret shares. This property makes our scheme secure and robust. All this is possible because the actual secret is not a function of compartment shares alone. That means, the secret can be changed by the dealer any number of times without changing compartment secrets. This property makes our scheme to be reusable in nature and hence avoids communication intense phase of share distribution to the participants whenever the secret is changed. Whenever the secret is changed, the dealer publishes a public value based on the secret and compartment shares. Then with the help of the same compartment secret shares along with the current public value, the compartments together can get the original secret. For each change in secret, the

compartments together has to reconstruct the current secret based on their secret compartment shares and the current public value, since the old secret is no more valid.

2 Related Work

(t, n)-threshold secret sharing schemes proposed by Adi shamir [18] and George Blakley [2] do not support any hierarchy among the participants, but Shamir mentioned that a hierarchical variant of (t, n)-threshold secret sharing scheme can be introduced by assigning larger number of shares to higher level participants. But this will not be ideal. Multipartite secret sharing was introduced by Simmons [3]. Multipartite secret sharing schemes were studied based on different assumptions by different authors *eg* [3, 4, 6–8, 11–13]. Multipartite Secret Sharing schemes by bivariate interpolation are developed by T. Tassa and N. Dyn [19]. Multilevel secret sharing scheme based on the Chinese remainder theorem was discussed in [10]. Hierarchical secret sharing schemes based on MDS codes was proposed in [20].

Secret sharing in multilevel and compartmented groups are discussed in [9]. Our scheme is influenced from the idea of compartmented scheme discussed in [9] in which the global threshold is equal to the sum of all individual compartment thresholds. The scheme, proposed in this paper uses superincreasing sequence, allows concurrency during compartmental share reconstruction. But secret reconstruction is strictly hierarchical in nature, *i.e.*, lower level can get the secret bit(s) information only if all higher levels get their secret bit(s).

The organization of the paper is as follows: In Section 3, the background required like knapsack problem (restricted to our scenario), algorithms related to superincreasing sequence are described and describe concisely the Shamir's secret sharing scheme. In Section 4, multipartite secret sharing scheme based on the superincreasing sequence was proposed. In Section 5, the scheme was explained with an example. In Section 6, the property of secret changeability with an example was explained. In Section 7, a brief note on the security of the scheme in the presence of an intruder and some observations are discussed. Finally concluding remarks are in Section 8.

3 Background

This section reviews the knapsack problem, corresponding instances, algorithms, and the Shamir's threshold secret sharing scheme.

3.1 Knapsack Functions

If the elements inside a knapsack are known, then it is easy to compute the sum of the elements inside the knapsack. But if the sum of the numbers inside the knapsack is only known, it is difficult to list out the numbers inside the knapsack.

Mathematically this observation can be stated as follows:

Let *bagsum* be a function that takes two r-tuples as input and returns an integer value as output. Let $w = [w_1, w_2, w_3, \cdots, w_r]$ and $s = [s_1, s_2, s_3, \cdots, s_r]$ be two r-tuples. The second tuple s contains Boolean elements only *i.e.*, $s_i = 0$ or 1 for $i \in \{1, 2, 3, \cdots, r\}$. $sum = bagsum(w, s) = \sum_{i=1}^{r} w_i s_i = w_1 s_1 + w_2 s_2 + w_3 s_3 + \cdots + w_r s_r$.

Let *bagsum_inverse* be a function that takes the sum value and the first tuple w as input and returns the second Boolean tuple s as output i.e., $s = bagsum\_inverse(sum, w)$.

Given w and s it is easy to find *bagsum*, but given *sum* and w it is difficult to find s. But if the first tuple w contains the superincreasing sequence then it is easy to compute both *bagsum* and *bagsum_inverse* [16].

3.2 Superincreasing Sequence

A sequence is said to be *superincreasing*, if every element (except first) in the sequence is greater than or equal to the sum of all its previous elements. A tuple is said to be superincreasing, if it contains a superincreasing sequence. Thus a tuple $w = [w_1, w_2, w_3, \cdots w_r]$ is superincreasing if and only if $w_j \geq w_1 + w_2 + w_3 + \cdots w_{j-1}$ for $2 \leq j \leq r$.

The pseudo codes of *bagsum* and *bagsum_inverse* for superincreasing sequence are as follows:

Algorithm 1: *bagsum* function

```
1 Function bagsum (w,s);
  Input  : Two tuples: w = [w_1, w_2, w_3, ··· w_r] and
           s = [s_1, s_2, s_3, ··· s_r].
  Output: sum: an integer.
2 sum ← 0;
3 for i = 1 ··· r do
4 |  sum ← sum + w_i × s_i ;
5 end
6 Return sum;
```

The running time of *bagsum* is $\mathcal{O}(r)$. Applied to the superincreasing sequence, in our scheme, the dealer uses this function.

The running time of *bagsum_inverse* applies to the superincreasing sequence (Algorithm 2) is same as *bagsum* *i.e.*, $\mathcal{O}(r)$. In our scheme, this algorithm is implicitly used by compartments during secret reconstruction phase.

Let us define another function *exor*, which takes two binary tuples (bit arrays) of same size as input and performs bit-wise exclusive or of them and outputs the resultant tuple. Let b, b' be two bit tuples of same size, then $c = exor(b, b')$ such that $c_i = b_i \oplus b'_i$, where c_i, b_i, b'_i is the i^{th} bit in the tuples c, b, b' respectively. We use *exor* function to improve security of our scheme by means of hiding the original secret tuple from compartments.

Algorithm 2: *bagsum_inverse* Function for superincreasing tuple

```
Function bagsum_inverse (sum, w);
  Input  : An integer sum and tuple
           w = [w_1, w_2, w_3, ··· w_r].
  Output: Boolean Tuple: s = [s_1, s_2, s_3 ··· s_r].
for i = r ··· 1 do
    if sum ≥ w_i then
        s_i ← 1;
        sum ← sum − w_i;
    else
        s_i ← 0
    end
    Return [s_1, s_2, s_3 ··· s_r];
end
```

3.3 Shamir's Secret Sharing Scheme

Shamir's secret sharing scheme is a threshold secret sharing scheme. It has two phases: share distribution phase and secret reconstruction phase. In share distribution phase the dealer calculates shares and distributes them securely to the participants. In secret reconstruction phase at least threshold number of participants cooperatively participate in interpolation to arrive at the corresponding Lagrange polynomial and recovers the secret. In Shamir's secret sharing scheme dealer uses a polynomial and public identities of the participants to calculate the share values.

Let n be the number, $\{P_1, P_2, P_3, \cdots P_n\}$ be the participants and $t(\leq n)$ be the threshold. Let id_i be the public identity of the participant P_i for $i \in [1, n]$, where $[1, n]$ denote the set $\{1, 2, 3, \cdots, n\}$. Then the Shamir's secret sharing scheme is as follows:

3.3.1 Share Distribution Phase

Dealer performs the following steps:

1) Selects a prime number $q > n$;
2) Selects a secret $s \in \mathbb{Z}_q^*$;
3) Generates a polynomial $f(x) = s + a_1x + a_2x^2 + \cdots + a_{t-1}x^{t-1}$ of degree atmost $t-1$ with coefficients in $a_i \in \mathbb{Z}_q^*$ for $1 \leq i \leq t-1$;
4) Calculates $[s]_i = f(id_i) \mod q$ for $1 \leq i \leq n$;
5) Sends $[s]_i$ securely to the participant P_i for $1 \leq i \leq n$.

3.3.2 Secret Reconstruction Phase

- At least t participants are required to cooperate to reconstruct the secret; which can be done using Lagrange polynomial interpolation formula as follows:

$$s = \sum_{i=1}^{t} \left(\prod_{j \neq i} \frac{id_j}{id_j - id_i} \right) [s]_i \mod q.$$

In our proposed scheme, Shamir's secret sharing scheme is used to share the compartment secret and to reconstruct it.

4 Multipartite Secret Sharing Scheme

Our multipartite secret sharing scheme, based on superincreasing sequence, consists of share distribution phase followed by secret reconstruction phase. After presenting pseudo-code, we explain the algorithm with an example. Let $\mathfrak{P}$ be the set of all participants and let the participants be divided into ℓ disjoint levels. $\mathfrak{P}_i$ be the set of participants at level i. n_i be the total number of participants at level i and $t_i \leq n_i$ be its corresponding compartment (level) threshold. Note that the global threshold t for this scheme is the sum of all the individual level thresholds.

4.1 Share Distribution:

Dealer performs the following steps:

1) Selects a secret $s \in \mathbb{Z}_{2^{\ell-1}} - \{0\}$;
2) Converts the secret s into $\ell - 1$ bit binary number $s_{\ell-1}s_{\ell-2}s_{\ell-3} \cdots s_3s_2s_1$ and forms the secret tuple $st = [s_{\ell-1}, s_{\ell-2}, s_{\ell-3} \cdots s_3, s_2, s_1]$;
3) Generates a random bit tuple $'temp'$ of length $\ell - 1, temp = [e_{\ell-1}, e_{\ell-2}, e_{\ell-3} \cdots e_3, e_2, e_1]$;
4) Calculates $st' = exor(st, temp)$, let $st' = [s'_{\ell-1}, s'_{\ell-2}, s'_{\ell-3} \cdots s'_3, s'_2, s'_1]$;
5) Generates a superincreasing tuple w of size $\ell-1, w = [w_{\ell-1}, w_{\ell-2}, w_{\ell-3} \cdots w_3, w_2, w_1]$;
6) Calculates $u = bagsum(st', w)$ and makes u public;
7) Select a prime q, which is greater than the sum $w_1 + w_2 + w_3 + \cdots w_{\ell-1}$ and $max(n_i)$ for $1 \leq i \leq \ell$, and make q public;
8) Distributes shares of w_i to all the participants in level i using Shamir's (t_i, n_i) share distribution for $1 \leq i \leq \ell - 1$ and distribute the shares of decimal equivalent of bit tuple $temp$ to level ℓ by means of Shamir's (t_ℓ, n_ℓ) share distribution;

Note:

- q, u is public. $s, st, temp, st', w$ are not public, but known to dealer.
- In simple and informal language, the share distribution phase can be summarized as: st is the binary tuple for s. $st' = st \oplus temp$.
- Since $temp$ is distributed to the compartment ℓ, eventhough remaining $\ell - 1$ levels gets the associated bit values, they cannot get secret tuple st.

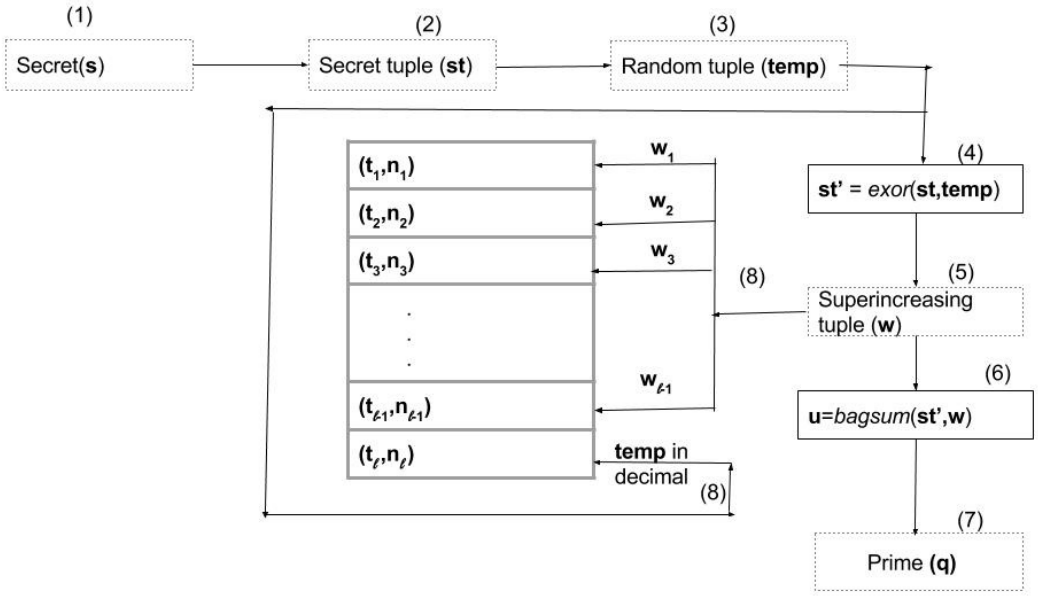

Figure 1: Stepwise pictorial representation of share distribution

4.2 Secret Reconstruction

1) At least t_i participants of level i perform (t_i, n_i) Shamir's secret reconstruction and gets the level share w_i for $1 \leq i \leq \ell - 1$;

2) Level 1 checks whether $u \geq w_1$. If true, then Level 1 outputs bit 1 and sends $u'_{1,2} = u - w_1$ to level 2. Else it outputs bit 0 and sends $u'_{1,2} = u$ to level 2 and appends its output bit (which is s'_1) to an empty tuple (say st'');

3) For $2 \leq i \leq \ell - 1$, Level i checks whether $u'_{i-1,i} \geq w_i$. If true, then level i outputs bit 1 and sends $u'_{i,i+1} = u'_{i-1,i} - w_i$ to level $i+1$. Else outputs bit 0 and sends $u'_{i,i+1} = u'_{i-1,i}$ and appends the output bit (which is s'_i) to the starting index of the tuple st'';

4) Level ℓ performs (t_ℓ, n_ℓ) Shamir's secret reconstruction and converts the result into $\ell - 1$ bit tuple, which is $temp = [e_{\ell-1}, e_{\ell-2}, e_{\ell-3} \cdots e_3, e_2, e_1]$. (Observe that $st' = st''$ and st'' is not public). level ℓ performs $exor(temp, st')$ which results in st;

5) Finally the secret binary tuple st is obtained, which can be converted to decimal to obtain the secret s.

Note: Steps 2 and 3 corresponds to the *bagsum_inverse* algorithm discussed earlier.

5 Explanation with an Example

Let the number of levels (ℓ) be 6. Also let the threshold values and total number of participants for each level be as follows: $(t_1, n_1) = (2, 5)$, $(t_2, n_2) = (3, 5)$, $(t_3, n_3) = (4, 5)$, $(t_4, n_4) = (5, 6)$, $(t_5, n_5) = (3, 6)$, $(t_6, n_6) = (5, 7)$. Let $1, 2, 3 \cdots n_i$ be the public identities of the participants in compartment i.

5.1 Share Distribution:

Dealer performs the following steps:

1) Selects a secret $s \in \mathbb{Z}_{2^{\ell-1}} - \{0\}$;

$$\textbf{Since } \boldsymbol{\ell = 6}\textbf{, } \boldsymbol{\mathbb{Z}_{2^{\ell-1}} - \{0\} = \{1, 2, 3 \cdots 31\}}\textbf{. Let } \boldsymbol{s = 21}\textbf{.}$$

2) Converts the secret s into $\ell - 1$ bit binary number $s_{\ell-1}s_{\ell-2}s_{\ell-3} \cdots s_3 s_2 s_1$ and forms secret tuple $st = [s_{\ell-1}, s_{\ell-2}, s_{\ell-3} \cdots s_3, s_2, s_1]$;

$$\boldsymbol{st = [1, 0, 1, 0, 1]}$$

3) Generates a random bit array(tuple) $'temp'$ of length $\ell - 1, temp = [e_{\ell-1}, e_{\ell-2}, e_{\ell-3} \cdots e_3, e_2, e_1]$;

$$\boldsymbol{temp = [0, 1, 1, 1, 0]}$$

4) Calculates $st' = exor(st, temp)$, let $st' = [s'_{\ell-1}, s'_{\ell-2}, s'_{\ell-3} \cdots s'_3, s'_2, s'_1]$;

$$\boldsymbol{st' = exor(st, temp) =}$$
$$\boldsymbol{exor([1, 0, 1, 0, 1], [0, 1, 1, 1, 0]) = [1, 1, 0, 1, 1]}$$

5) Generates a superincreasing tuple w of size $\ell - 1, w = [w_{\ell-1}, w_{\ell-2}, w_{\ell-3} \cdots w_3, w_2, w_1]$;

$$\textbf{Let } \boldsymbol{w = [10, 11, 30, 60, 130]}$$

Table 1: Shamir's share distribution at compartments

Level number	$(\mathbf{t_i}, \mathbf{n_i})$	Dealer polynomial	Participants share list (Public identity, Share value)
1	(2,5)	$130 + x$	$(1, 131), (2, 132), (3, 133), (4, 134), (5, 135)$
2	(3,5)	$60 + 2x + 3x^2$	$(1, 65), (2, 76), (3, 93), (4, 116), (5, 145)$
3	(4,5)	$30 + 4x + x^3$	$(1, 35), (2, 46), (3, 69), (4, 110), (5, 175)$
4	(5,6)	$11 + 2x + 4x^3 + 3x^4$	$(1, 20), (2, 95), (3, 368), (4, 502), (5, 232), (6, 447)$
5	(3,6)	$10 + x^2$	$(1, 11), (2, 14), (3, 19), (4, 26), (5, 35), (6, 46)$
6	(5,7)	$14 + x + x^2 + x^3 + 16x^4$	$(1, 33), (2, 284), (3, 267), (4, 407), (5, 431), (6, 450), (7, 418)$

6) Calculates $u = bagsum(st', w)$ and makes u public;

$$\boldsymbol{u = bagsum(st', w) = bagsum([1, 1, 0, 1, 1], [10, 11, 30, 60, 130]) = 10 + 11 + 60 + 130 = 211}$$

7) Select a prime q, which is greater than the sum $w_1 + w_2 + w_3 + \cdots w_{\ell-1}$ and $max(n_i)$ for $1 \leq i \leq \ell$ make u as public;

$$\boldsymbol{w_1 + w_2 + w_3 + w_4 + w_5 = 10 + 11 + 30 + 60 + 130 = 241,\ max(n_i) = 7,\ \text{let } q = 541}$$

8) Distributes shares of w_i to all the participants in level i using Shamir's (t_i, n_i) share distribution for $\ell - 1 \leq i \leq 1$ and distribute the shares of decimal equivalent of bit tuple $temp$ to level ℓ by means of Shamir's (t_ℓ, n_ℓ) share distribution;

We have the compartmental shares given in Table 1.

5.2 Secret Reconstruction

1) Performing (t_i, n_i), $1 \leq i \leq \ell - 1$ Shamir's secret reconstruction, we have the compartmental shares given in Table 2.

2) Carrying out the step 2 of secret reconstruction, we have the results given in Table 3.

3) Results of the step 3 of the secret reconstruction are given in Table 4.

4) Step 4 of the reconstruction gives the results given in Table 5. Also

$$\boldsymbol{(14)_{10} = (01110)_2 \implies temp = [0, 1, 1, 1, 0]}$$

$$\boldsymbol{exor(temp, st') = exor([0, 1, 1, 1, 0], [1, 1, 0, 1, 1]) = [1, 0, 1, 0, 1] = st}$$

5) Converting the binary tuple st to decimal, we have $s = 21$.

6 Secret Changeability with an Example

6.1 Secret Changeability

In our scheme, the shares are reusable, That is, the shares of the participants can remain same even for a new secret. The following algorithm is to renew the secret by the dealer.

6.1.1 Secret Renewal

Dealer performs the following steps:

1) Selects a new secret $ns \in \mathbb{Z}_{2^{\ell-1}} - \{0\}$;

2) Converts the secret ns into $\ell - 1$ bit binary number $ns_{\ell-1}ns_{\ell-2}ns_{\ell-3}\cdots ns_3ns_2ns_1$ and forms new secret tuple $nst = [ns_{\ell-1}, ns_{\ell-2}, ns_{\ell-3}\cdots ns_3, ns_2, ns_1]$;

3) Calculates $nst' = exor(nst, temp)$, let $nst' = [ns'_{\ell-1}, ns'_{\ell-2}, ns'_{\ell-3}\cdots ns'_3, ns'_2, ns'_1]$;

4) Calculates $nu = bagsum(nst', w)$ and makes nu public.

Table 6 shows the parameters that change after the secret renewal algorithm.

There is a change in the secret, secret tuple, $exor$ tuple and public value only, all other parameters are intact. Since w contains the compartment secret shares for levels from 1 to $\ell - 1$ and decimal equivalent of $temp$ is the compartment secret share for level ℓ, the compartment shares doesn't change after changing the secret by the dealer. Since there is no change in compartment shares, there is no need for distribution phase after the secret update. There is no change in the secret reconstruction phase. Based on the new public value nu, the compartment can get the actual secret using their corresponding secret shares. We explain the algorithm in detail with an example in next subsection. Pinnacle step for running time of the above algorithm is 4. So the running time of the secret renewal algorithm is $\mathcal{O}(\ell)$.

6.2 Example for Secret Changeability

We use the example explained in section 5 to explain share renewal algorithm. Now, total number of levels (ℓ)=6. Threshold values and total number of participants for

Table 2: Shamir's secret reconstruction at compartments

Level number	$(\mathbf{t_i}, \mathbf{n_i})$	**Interested participants with shares**	**Compartmental share (lagrange interpolation)**
1	(2,5)	$(1, 131), (3, 133)$	$\frac{3}{2}(131) + \frac{-1}{2}(133) = \mathbf{130}$
2	(3,5)	(1,65), (3,93), (5,145)	$\frac{15}{8}(65) + \frac{-5}{4}(93) + \frac{3}{8}(145) = \mathbf{60}$
3	(4,5)	(1,35), (3,69), (4,110), (5,175)	$\frac{5}{2}(35) + (-5)(69) + 5(110) + \frac{-3}{2}(175) = \mathbf{30}$
4	(5,6)	(1,20), (2,95), (3,368), (4,502), (5,232)	$5(20) + (-10)(95) + 10(368) + (-5)(502) + 1(232) = 552 \equiv \mathbf{11} \mod 541$
5	(3,6)	(1,11), (3,19), (6,46)	$\frac{9}{5}(11) + (-1)(19) + \frac{1}{5}(46) = \mathbf{10}$
6	(5,7)	(1,33), (2,284), (3,267), (4,407), (5,431)	$5(33) + (-10)(284) + 10(267) + (-5)(407) + 1(431) = -1609 \equiv \mathbf{14} \mod 541$

Table 3: Output bit from Level 1

Level number	**u**	$\mathbf{w_1}$	$\mathbf{u \geq w_1}$	**Output bit**	$\mathbf{u'_{1,2}}$	$\mathbf{st''}$
1	211	130	True	1	211-130=81	[1]

Table 4: Output bits from Level 2 to Level 5

Level number	$\mathbf{u'_{i-1,i}}$	$\mathbf{w_i}$	$\mathbf{u'_{i-1,i} \geq w_i}$	**Output bit**	$\mathbf{u'_{i,i+1}}$	$\mathbf{st''}$
2	81	60	True	1	81-60=21	[1,1]
3	21	30	False	0	21	[0,1,1]
4	21	11	True	1	21-11=10	[1,0,1,1]
5	10	10	True	1	10-10=0	[1,1,0,1,1]

Table 5: Shamir's secret reconstruction at last compartment

Level number	$(\mathbf{t_i}, \mathbf{n_i})$	**Interested participants with shares**	**Compartmental share (Lagrange interpolation)**
6	(5,7)	(1,33), (2,284), (3,267), (4,407), (5,431)	$5(33) + (-10)(284) + 10(267) + (-5)(407) + 1(431) = -1609 \equiv \mathbf{14} \mod 541$

Table 6: Table showing change in values

Parameter	**Previous value**	**New value**	**Change in parameter**
Secret	s	ns	Yes
Secret tuple	st	nst	Yes
exor tuple	st'	nst'	Yes
Random bit tuple	$temp$	$temp$	No
Super increasing tuple	w	w	No
Public value(u)	u	nu	Yes
prime number	q	q	No

each level are as follows: $(t_1, n_1) = (2, 5)$, $(t_2, n_2) = (3, 5)$, $(t_3, n_3) = (4, 5)$, $(t_4, n_4) = (5, 6)$, $(t_5, n_5) = (3, 6)$, $(t_6, n_6) = (5, 7)$.

Now dealer executes the following algorithm to change secret value from $s = 21$ to $ns = 25$.

6.2.1 Secret Renewal

Dealer performs the following steps:

1) Selects a new secret $ns \in \mathbb{Z}_{2^{\ell-1}} - \{0\}$;

$$\boldsymbol{ns = 25, \text{ Since } \ell = 6,}$$
$$\boldsymbol{\mathbb{Z}_{2^{\ell-1}} - \{0\} = \{1, 2, 3 \cdots 31\}.}$$

2) Converts the secret ns into $\ell - 1$ bit binary number $ns_{\ell-1}ns_{\ell-2}ns_{\ell-3} \cdots ns_3ns_2ns_1$ and forms new secret tuple $nst = [ns_{\ell-1}, ns_{\ell-2}, ns_{\ell-3} \cdots ns_3, ns_2, ns_1]$;

$$\boldsymbol{nst = [1, 1, 0, 0, 1]}$$

3) Calculates $nst' = exor(nst, temp)$, let $nst' = [ns'_{\ell-1}, ns'_{\ell-2}, ns'_{\ell-3} \cdots ns'_3, ns'_2, ns'_1]$;

$$\boldsymbol{nst' = exor(nst, temp) =}$$
$$\boldsymbol{exor([1, 1, 0, 0, 1], [0, 1, 1, 1, 0]) = [1, 0, 1, 1, 1]}$$

4) Calculates $nu = bagsum(nst', w)$ and makes nu public;

$$\boldsymbol{nu = bagsum(nst', w) =}$$
$$\boldsymbol{bagsum([1, 0, 1, 1, 1], [10, 11, 30, 60, 130]) =}$$
$$\boldsymbol{10 + 30 + 60 + 130 = 230}$$

Now, we explain secret reconstruction algorithm with the new values. Instead of repeating compartment share reconstruction, we assume that all applied Lagrange interpolation and obtained their shares as shown in share reconstruction of Section 5.

6.2.2 Secret Reconstruction

1) At least t_i participants of level i performs (t_i, n_i) Shamir's secret reconstruction and gets the compartment/level share w_i for $1 \le i \le \ell - 1$;

Table 7: Compartment secrets

Level number	Compartmental share (Lagrange interpolation)
1	**130**
2	**60**
3	**30**
4	$552 \equiv \mathbf{11} \mod 541$
5	**10**
6	$-1609 \equiv \mathbf{14} \mod 541$

2) Carrying out the step 2 of secret reconstruction, we have the results given in Table 8;

3) Step 3 of the reconstruction gives the results given in Table 9;

4) Level ℓ performs (t_ℓ, n_ℓ) Shamir's secret reconstruction and converts the result in to $\ell - 1$ bit tuple, which is $temp = [e_{\ell-1}, e_{\ell-2}, e_{\ell-3} \cdots e_3, e_2, e_1]$. (Observe that $st' = st''$ and st'' is public). level ℓ performs $exor(temp, st')$ which results in st;

$$\boldsymbol{(14)_{10} = (01110)_2 \implies temp = [0, 1, 1, 1, 0]}$$
$$\boldsymbol{exor(temp, nst') =}$$
$$\boldsymbol{exor([0, 1, 1, 1, 0], [1, 0, 1, 1, 1]) =}$$
$$\boldsymbol{[1, 1, 0, 0, 1] = nst}$$

5) Finally the secret binary tuple nst is obtained, which can be converted to decimal to obtain secret s.

$$\boldsymbol{nst = [1, 1, 0, 0, 1] \implies ns = (11001)_2 = 25.}$$

Thus the new secret can be reconstructed by the compartments using the same shares. Note that for the sake of completeness, we repeated Steps 1 to 4 in the above algorithm, which is not needed since the compartments already have calculated their shares earlier.

7 Brief Note on Security and Observations

7.1 Brief Note on Security

- *Public*: u, q, ℓ;
- *Private to dealer*: $s, st, temp, st', w$, polynomials (use to generate shares);
- *Private to compartment i participant*: Participant share of the compartment secret w_i.

Note that with public values u, q, an intruder (either inside or outside $\mathfrak{P}$) cannot get about st and s because s, st are private to dealer only and w is needed to calculate s along with u. All levels has to intervene to get the secret, because $w, temp$ are available only after the secret share reconstruction of each and every compartment. In our scheme, the compartment with lower level number has highest priority because, other levels cannot get the bit information without the value passed by the higher level.

Assuming that the dealer is honest, it is infeasible to obtain s from u, q, ℓ. Hence our scheme is secure with respect to trusty dealer.

7.2 Observations

The multipartite secret sharing scheme proposed deals with the $\ell - 1$ bit secret, i.e., $1 \le s \le 2^{\ell-1} - 1$. So, in step 1 of share distribution phase, dealer selects secret $s \in \mathbb{Z}_{2^{\ell-1}} - \{0\}$. The selection range of the secret can be increased. Suppose the secret is of $\ell'(> \ell - 1)$ bits, then the following two trivial methods can be employed to handle this issue:

Table 8: Output bit from Level 1

Level number	**nu**	$\mathbf{w_1}$	$\mathbf{nu \geq w_1}$	**Output bit**	$\mathbf{nu'_{1,2}}$	$\mathbf{nst''}$
1	230	130	True	1	230-130=100	[1]

Table 9: Output bits from Level 2 to Level 5

Level number	$\mathbf{nu'_{i-1,i}}$	$\mathbf{w_i}$	$\mathbf{nu'_{i-1,i} \geq w_i}$	**Output bit**	$\mathbf{nu'_{i,i+1}}$	$\mathbf{nst''}$
2	100	60	True	1	100-60=40	[1,1]
3	40	30	True	1	40-30=10	[1,1,1]
4	10	11	False	0	10	[0,1,1,1]
5	10	10	True	1	10-10=0	[1,0,1,1,1]

Table 10: Shamir's secret reconstruction at last compartment

Level number	**Compartmental share (Lagrange interpolation)**
6	$-1609 \equiv \mathbf{14} \mod 541$

Method 1: one of the levels can be used to handle the extra bits. The extra bits can be shared to a compartment as compartment share (may be after *exor* with *temp* tuple) and after reconstruction phase, these bits get pre-appended to the tuple st''.

Method 2: Each level may be given multiple compartment shares. Thus, each level can reconstruct multiple compartment secrets and hence multiple bits related to secret.

The pinnacle step for running time is share distribution by dealer to all the participants using polynomial, which is $\mathcal{O}(\sum_{i=1}^{\ell}(n_i t_i))$. Since compartments can concurrently get their shares during secret reconstruction phase, the running time is $\mathcal{O}((max(n_i))^2 + \ell)$. Note that $\ell - 1$ comparisons and one *exor* is needed in order to obtain the secret.

8 Conclusions

In this paper, we proposed our new multipartite secret sharing scheme, which is based on superincreasing sequence. The property of secret changeability is explained along with an example, listed various observations and discussed briefly about the security of the scheme in the case of trustworthy dealer. Work is underway to extend these ideas to arrive at similar schemes for other access structures.

References

[1] F. Alsolami and T. E. Boult, "Cloudstash: Using secret-sharing scheme to secure data, not keys, in multi-clouds," in *11th International Conference on Information Technology: New Generations (ITNG'14)*, pp. 315–320, 2014.

[2] G. R. Blakley, "Safeguarding cryptographic keys," *Proceeding of the National Computer Conference1979*, vol. 48, pp. 313–317, 1979.

[3] E. F. Brickell, "Some ideal secret sharing schemes," in *Workshop on the Theory and Application of of Cryptographic Techniques*, pp. 468–475, 1989.

[4] E. Dawson and D. Donovan, "The breadth of shamir's secret-sharing scheme," *Computers & Security*, vol. 13, no. 1, pp. 69–78, 1994.

[5] N. A. Ebri, J. Baek and C. Y. Yeun, "Study on secret sharing schemes (SSS) and their applications," in *International Conference for Internet Technology and Secured Transactions (ICITST'11)*, pp. 40–45, 2011.

[6] O. Farràs, J. Martí-Farré, and C. Padró, "Ideal multipartite secret sharing schemes," *Journal of Cryptology*, vol. 25, no. 3, pp. 434–463, 2012.

[7] O. Farras, C. Padró, C. Xing, and A. Yang, "Natural generalizations of threshold secret sharing," *IEEE Transactions on Information Theory*, vol. 60, no. 3, pp. 1652–1664, 2014.

[8] P. M. Fathimal and P. A. J. Rani, "Threshold secret sharing scheme for compartmented access structures," *International Journal of Information Security and Privacy*, vol. 10, no. 3, pp. 1–9, 2016.

[9] H. Ghodosi, J. Pieprzyk, and R. Safavi-Naini, "Secret sharing in multilevel and compartmented groups," in *Australasian Conference on Information Security and Privacy*, pp. 367–378, 1998.

[10] L. Harn and M. Fuyou, "Multilevel threshold secret sharing based on the chinese remainder theorem," *Information Processing Letters*, vol. 114, no. 9, pp. 504–509, 2014.

[11] C. F. Hsu and L. Harn, "Multipartite secret sharing based on crt," *Wireless Personal Communications*, vol. 78, no. 1, pp. 271–282, 2014.

[12] M. Ito, A. Saito, and T. Nishizeki, "Secret sharing scheme realizing general access structure," *Electronics and Communications in Japan (Part III: Fundamental Electronic Science)*, vol. 72, no. 9, pp. 56–64, 1989.

[13] S. C. Kothari, "Generalized linear threshold scheme," in *Workshop on the Theory and Application of Cryptographic Techniques*, pp. 231–241, 1984.

[14] P. S. Kumar, R. R. Kurra, A. N. Tentu, and G. Padmavathi, "Multi-level secret sharing scheme for mobile ad-hoc networks," *International Journal of Advanced Networking and Applications*, vol. 6, no. 2, pp. 2253, 2014.

[15] M. Larson, C. Hu, R. Li, W. Li, and X. Cheng, "Secure auctions without an auctioneer via verifiable secret sharing," in *Proceedings of the Workshop on Privacy-Aware Mobile Computing*, pp. 1–6, 2015.

[16] R. Merkle and M. Hellman, "Hiding information and signatures in trapdoor knapsacks," *IEEE Transactions on Information Theory*, vol. 24, no. 5, pp. 525–530, 1978.

[17] D. K. Pattipati, A. N. Tentu, C. V. Vadlamudi and A. A. Rao, "Sequential secret sharing scheme based on level ordered access structure.," *International Journal of Network Security*, vol. 18, no. 5, pp. 874–881, 2016.

[18] A. Shamir, "How to share a secret," *Communications of the ACM*, vol. 22, no. 11, pp. 612–613, 1979.

[19] T. Tassa and N. Dyn, "Multipartite secret sharing by bivariate interpolation," *Journal of Cryptology*, vol. 22, no. 2, pp. 227–258, 2009.

[20] A. N. Tentu, P. Paul and C. V. Vadlamudi, "Conjunctive hierarchical secret sharing scheme based on mds codes," in *International Workshop on Combinatorial Algorithms*, pp. 463–467, 2013.

[21] Y. Wang and Y. Desmedt, "Efficient secret sharing schemes achieving optimal information rate," in *IEEE Information Theory Workshop (ITW'14)*, pp. 516–520, 2014.

Biography

Putla Harsha received M.Tech from University of Hyderabad, Hyderabad and she did Bachelors degree in computer science. Her research interests include Cryptography, Algorithms, and Computational number theory.

Patibandla Chanakya received M.Tech from University of Hyderabad, Hyderabad. Currently he is pursuing his PhD in Computer Science from University of Hyderabad. His research interests include Computational number theory, Algorithms, and Cryptography.

Vadlamudi China Venkaiah obtained his PhD in 1988 from the Indian Institute of Science (IISc), Bangalore in the area of scientific computing. He worked for several organisations including the Central Research Laboratory of Bharat Electronics, Tata Elxsi India Pvt. Ltd., Motorola India Electronics Limited, all in Bangalore. He then moved onto academics and served IIT, Delhi, IIIT, Hyderabad, and C R Rao Advanced Institute of Mathematics, Statistics, and Computer Science. He is currently serving the Hyderabad Central University. He is a avid researcher. He designed algorithms for linear programming, subspace rotation and direction of arrival estimation, graph colouring, matrix symmetriser, integer factorisation, cryptography, knapsack problem, *etc.*

Secure Time Synchronization Protocol for Wireless Sensor Network Based on uTESLA Broadcasting Protocol

Xiaogang Wang[1,2], Weiren Shi[1]

(Corresponding author: Xiaogang Wang)

College of Automation, Chongqing University[1]
No. 174, Shapingba Road, Shapingba District, Chongqing 400044, China
School of Automation and Information Engineering, Sichuan University of Science and Engineering[2]
No. 108, XueYuan Road, Ziliujing District, Zigong City, Sichuan 643000, China
(Email: wxg_zf@163.com)

(Received July 21, 2016; revised and accepted Feb. 2, 2017)

Abstract

A secure time synchronization protocol (STSP) for wireless sensor network based on uTESLA protocol is proposed according to the issues that DCS algorithm in wireless sensor network time synchronization is limited in the threats of compromised nodes, big communication cost and deficiency in coverage. Firstly, initializing the network topology based on LEACH protocol to balance the network node load and prolong network life cycle. Secondly, excluding the malicious nodes based on uTESLA protocol for making security condition check positively before the network time synchronization. Thirdly, making time synchronization quickly based on TPSN protocol for base station to cluster head node. Fourthly, making time synchronization based on HRTS protocol for cluster head nodes to their own common nodes. Lastly, making reverse authentication based on uTESLA protocol for checking and excluding the compromise nodes after the network time synchronization. The security analysis and simulation show that two times authentication ensure the absolute security of the time synchronization work in STSP, and STSP is much better than DCS in synchronization cycle, precision, ratio and cost.

Keywords: Broadcast Authentication; Time Synchronization; Wireless Sensor Network

1 Introduction

Wireless sensor network (WSN) is a new distributed system in which the nodes are independent and communicate wirelessly [2]. In particular, each node maintains a local clock and the timing signal of each node clock is generally maintained by an inexpensive crystal oscillator, and because of the limitation of crystal oscillator manufacturing process, it is easy to be influenced by the external factors in the process of operation, which leads to the deviation of the time ratio of the node, it's also known as the time out of step [4, 7, 11, 12, 20, 21, 30]. So it is necessary to regularly carry out network time synchronization for maintaining the consistency of the local clock nodes.

Time synchronization is the process of providing a unified time scale for a distributed system by doing some operations on the local clock. Network time protocol (NTP) is the standard of the time synchronization protocol on the internet [23], which is used to synchronize the computer time with universal time coordinated (UTC) and obtain a high precision time by the external connection of a time receiver (such as WWVB, GPS, *etc.*). However, NTP, GPS and other similar traditional time synchronization technology can't be directly applied to WSN because of the following three differences [27]:

1) The sensor nodes in WSN are limited in volume, power supply, computing power, storage space, which causes the NTP protocol can't be run on sensor nodes.

2) There are great differences in bandwidth, anti-interference ability, and the ability to resist weak between them, because WSN uses wireless transmission mode, and the tradition internet mainly uses a reliable cable transmission mode.

3) WSN applications are highly localized or local optimum, while the tradition internet emphasizes the overall optimality.

There have been a number of protocols for WSN time synchronization proposed about the research topic of time synchronization for WSN in recent years. Such as: group authentication and group key distribution for

Ad-Hoc networks(GAGKD) [29],an accurate on-demand time synchronization protocol (AODTSP) [16] and long term and large scale time synchronization (LTLSTS) [17] proposed by Huang Ge, timing-sync protocol for sensor networks (TPSN) [10], improved time synchronization in ML-MAC [19], analysis of quantities influencing the performance of time synchronization(AQIPTS) [3], tiny-sync/mini-sync(TS/MS) [5], hierarchy referencing time synchronization protocol (HRTS) [1], *etc.* Although these time synchronization protocols have achieved good performance from the perspective of each highlighted, they are only applicable to the benign environment without any malicious nodes. While WSN is usually used in some military and commercial areas, it's inevitably that there will be a variety of malicious nodes attacks [13, 15, 22, 26, 31], such as that the malicious nodes can forge the time synchronization message, send the synchronization message containing the error time information, delay sending the synchronization message or not, and destroy the WSN normal time synchronization process [6, 8, 9, 18, 24, 28]. Therefore, the security issues of WSN time synchronization are particularly important.

At present, there are only a few protocols on the security aspects for WSN time synchronization, where diffusion-based clock synchronization (DCS) is the most typical one, which ensures the safety of time synchronization based on receiving 2s+1 synchronization messages from last layer nodes, but it caused a lot of difficulty and communication costs from beginning. In this paper, a secure time synchronization protocol (STSP) for wireless sensor network based on uTESLA protocol is proposed according to the issues that DCS algorithm is limited in the threats of compromised nodes, big communication cost and deficiency in coverage. Firstly, initializing the network topology based on LEACH protocol to balance the network node load and prolong network life cycle. Secondly, excluding the malicious nodes based on uTESLA protocol for making security condition check positively before network time synchronization. Thirdly, making time synchronization quickly based on TPSN protocol for base station to cluster head node. Fourthly, making time synchronization based on HRTS protocol for cluster head nodes to their own common nodes. Lastly, making reverse authentication based on uTESLA protocol for checking and excluding the compromise nodes after network time synchronization. The security analysis and simulation show that two times authentication ensure the absolute security of the time synchronization work in STSP, and STSP is much better than DCS in synchronization cycle, precision, ratio and cost.

This paper is organized as follows. In Section 2, analyze the related work, such as DCS algorithm principle and its issues. In Section 3, discuss the specific principle of STSP, including network model assumptions, initialization and STSP steps. In Section 4, analyze the security of STSP, and make a simulation compared with DCS. Lastly, make a summary in Section 5.

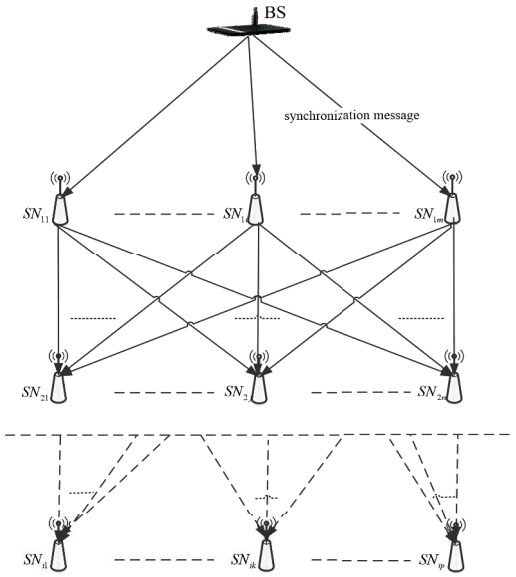

Figure 1: DCS algorithm

2 The Related Work

2.1 DCS Algorithm

DCS algorithm is a cyclical secure time synchronization algorithm for WSN: on the one hand, all nodes in the network can be synchronized to the reference node through the multi-hop manner; on the other hand, providing redundant paths for each receiving node in the multi-hop process, which can offset and tolerate the attacks of malicious nodes on some paths, and achieve the goal of secure time synchronization.

The execution process of DCS algorithm is shown in Figure 1.

- Firstly, the reference node BS(base station) sends the synchronization message to its neighbor nodes SN_{1i},$(1...m)$.
- Secondly, SN_{1i}sends the synchronization message to its neighbor nodesSN_{2j},$(1...n)$ when SN_{1i} completes the calibration of the clock.
- Thirdly, for tolerating the attacks of malicious nodes, the node SN_{ij} $(i > 1)$ needs to receive more than 2s+1 synchronization messages from last layer nodes$SN_{(i-1)j}$, and takes the average of these 2s+1 time messages to calibrate their own local clock.
- Lastly, the synchronous message is passed in turn until the whole network nodes can be synchronized.

2.2 DCS Issues

1) Can not guarantee that all nodes SN_{1i},$(1...m)$ are non-malicious nodes

The nodes SN_{1i},(1...m) do not calibrate the local time by receiving 2S+1 synchronization messages from last layer node BS, so there may be some malicious or compromised nodes in SN_{1i},(1...m), and it caused a lot of difficulty and communication costs from beginning.

2) Serious error accumulation

It's obviously shown in Figure 1 that the more far away from BS, the more error accumulation. We know that the calibrated time of node SN_{ij} $(i > 1)$ is an estimated value which is not precise, because the node SN_{ij} $(i > 1)$ needs to take the average of 2s+1 time messages receiving from last layer nodes $SN_{(i-1)j}$ to calibrate their own local clock, while the time message from $SN_{(i-1)j}$ is also an estimated value because of the same principle. So the error will be accumulated layer by layer.

3) Communication cost

It is bound to increase the redundancy of the message in the network and lead to the increase of the communication cost, because that all nodes need to receive more than 2S+1 synchronization messages from last layer nodes.

4) The condition of 2S+1 is difficult to meet

It is not all the nodes in the network can receive more than 2S+1 synchronization messages from last layer. For instance, the neighbor nodes are less than 2S+1, or the neighbor nodes are malicious nodes. So the DCS algorithm coverage is not good.

5) Compromised node threat

Although the authentication method can be used to defend against the attacks from external malicious nodes, the attacker can still attack the time synchronization process by compromised nodes. Especially for the multi-hop time synchronization process, if the intermediate node is compromised node, this effect is fatal.

3 STSP

3.1 Network Model Assumptions

In order to facilitate the description of STSP, the network is assumed as follows:

1) Assume that the network is isomorphic and static, each sensor node has been uniformly deployed in the target area and has same configure in software and hardware, where the network size is N, including 3 types of nodes: base station BS, cluster head node CH, common sensor node SN, planning network topology by LEACH protocol [14], as shown in Figure 2.

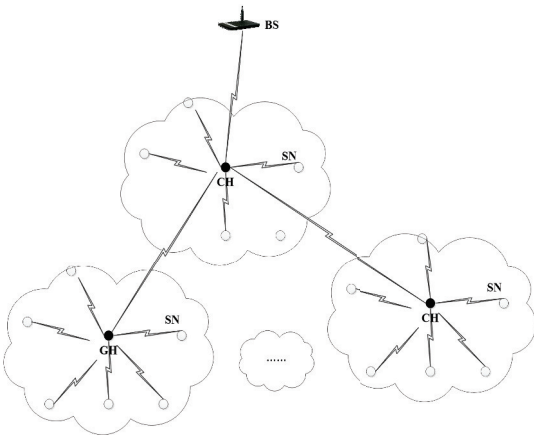

Figure 2: The network topology based on LEACH

2) Assume that base station BS is the time reference node for the network and equipped with abundant software and hardware resources, it is responsible for storing the basic information of all the nodes and has the ability to detect compromised or captured nodes.

3) Assume that common sensor node SN is responsible for collecting environmental data. The ability to process data of sensor node is limited by storage space, energy reserves, and communication distance. The messages between communication nodes which are not in the communication radius should be transferred by their neighbor node.

4) Assume that cluster head node CH is selected from the common nodes based on LEACH and responsible for the data transmission between SN and BS.

The main symbols in the text are shown in Table 1.

Table 1: Explanation of symbols

Symbol	Implication
BS	base station
CH	cluster head
SN	sensor node
h(x)	hash function
K	authentication key
ID_{SN_i}	identity symbol of node SN_i
D	time slice length
δ	key delay time
P	plaintext
L(i)	authentication message of time slice i

3.2 STSP Principle

3.2.1 Initialization

1) Initialization network topology based on LEACH

In order to ensure that each node can obtain the synchronization message from the reference node, synchronous data packet switching as an important process of WSN time synchronization is usually based on the specific network topology path. There is no essential difference between synchronization topology and routing topology except the types of data packet on transmission path, so it can be effectively combined with synchronous topology and routing topology for reducing the energy consumption of the network.

Compared with flooding type network topology of DCS algorithm, the network topology structure based on clustering hierarchical is more suitable for WSN applications (as shown in Figure 2), such as the application of broadcast authentication, reducing the amount of network data traffic, and prolonging the survival time of the network. The most classical clustering protocol LEACH in WSN is chosen to initialize the network topology in this paper, and the selection of cluster head node is the key to LEACH protocol.

Assume that each common sensor node generates a random number between 0 and 1, and it will be the cluster head if some random number is less than a certain threshold value $T(n)$,and set Equation (1):

$$T(n) = \left\{ \begin{array}{ll} 1-p[rmod(1/p)] & n \in G \\ 0 & else \end{array} \right\} \quad (1)$$

Where, p is the percentage of desired cluster head nodes, r is the round, G is the common sensor nodes set in last 1/p round.

2) BS presets the initial parameters of each node SN_i, including the last key $K^0_{SN_i}$ of the key chain, the key delay time δ,the time slice length D, the beginning time T_0, the node identity ID_{SN_i}.

3) SN_i presets the initial parameters of BS, including the last key K^0_{BS} of the key chain, the key delay time δ, time slice length D, the beginning time T_0.

3.2.2 STSP Steps

Step 1. Making security condition check positively based on uTESLA broadcasting protocol.

WSN internal safety hazards are generally caused by the malicious nodes and the compromised nodes, but the malicious nodes cannot access the network without getting the network authentication key because of the local broadcast authentication protocol such as uTESLA, and it can't cause any bad effect for the network time synchronization process.

In uTESLA [25], the asymmetric characteristic of broadcast authentication is realized by using the symmetric encryption mechanism in condition of the loose time synchronization of sending nodes and receiving nodes. The key points of uTESLA protocol are using hash key

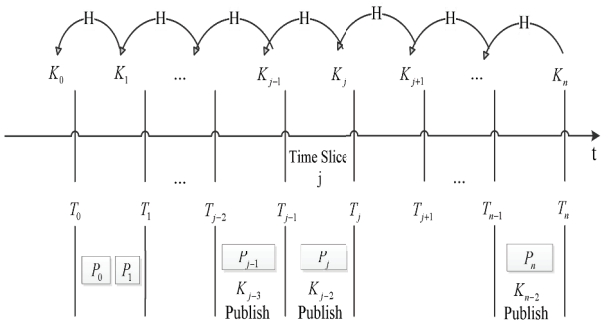

Figure 3: uTESLA protocol

chain and publishing key delayed, as showed in Figure 3, a one-way function key chain is established by the sending node, where the length of key chain is n+1, and the first key K_n of the key chain is generated randomly by the sending node, but the next keys are all generated by the one-way function acting on the last key repeatedly, such as $K_j = H(K_{j+1})$. The sending node divides the communication time into equal time slices, where the length of each time slice is D, and each time slice is assigned a key in order, but the order of the assigned keys is the opposite order of the key chain, and each message P_i of time slice j is encrypted by K_j, such as $MAC_{k_j}(P_i)$. The sending node determines the key delay time δ based on the time slice length, and the key K_j on time slice will be published after δ, such as $\delta = 2$ in Figure 3.

To avoid the additional communication cost, the published key is sent to the receiving nodes by being attached with the data packet. If there is no data packet on some time slice, the key attached with the data packet won't be published, and this key can be calculated by the next keys in one-way function hash. More importantly, the initial parameters K_0, δ, D and starting time T_0 should be sent to receiving nodes before authentication.

The specific steps for making security condition check based on the network topology (as shown in Figure 2) are as follows:

Firstly, BS builds the broadcast authentication information $L(j)$, and assume that $L(j)$ is the broadcast authentication information of the time $j(t)$ on time slice j, where t is a certain time on time slice j, and set Equation (2):

$$L(j) = \left\{ \begin{array}{l} h^j_{BS}(P_{j(t)}, T_{BS}(j(t)), ID_{BS})|| \\ \\ P_{j(t)}||h^{j-2}_{BS}||ID_{BS}||T_{BS}(j(t)) \end{array} \right\} \quad (2)$$

Where, $P_{j(t)}$ is the plaintext message of time $j(t)$, $T_{BS}(j(t))$ is the standard reference time from BS on time $j(t)$, h^{j-2}_{BS} is the published key by BS on time $j(t)$.

Secondly, rebuilding the broadcast authentication information $L_{CH_{1i}}(j)$ when the neighbor cluster head nodes

CH_{1i} obtain $L(j)$, and set Equation (3):

$$L_{CH_{1i}}(j) = \left\{ \begin{array}{l} h^j_{BS}(P_{j(t)}, T_{BS}(j(t)), ID_{BS}) \\ ||h^{j-2}_{BS}||ID_{CH_{1i}}||T_{BS}(j(t)) \\ ||T_{CH_{1i}}||P_{j(t)}||ID_{BS} \end{array} \right\} \quad (3)$$

And then CH_{1i} send $L_{CH_{1i}}(j)$ to their member nodes and neighbor cluster nodes CH_{2i}. Where $T_{CH_{1i}}$ is the local time of CH_{1i}.

Lastly, each node can get the information $L(j)$ based on receiving the broadcast authentication information from their neighbor nodes one by one, and waiting for getting the published key h^j_{BS} by BS after δ to certificate the correctness of $h^j_{BS}(P_{j(t)}, T_{BS}(j(t))$ which can illustrate the identity of BS and the correctness of source attestation. In this time, each node can verify whether the key h^j_{BS} has been published based on $h^{j-2}_{BS} = H^2(h^j_{BS})$: if h^j_{BS} has been published by BS, that each node which has obtained the key h^j_{BS} can forge or tamper with the information $L(j)$, so the information $L(j)$ from these nodes will be judged to be unsafe, and abandon it; if h^j_{BS} has not been published by BS, that each $L(j)$ will be cached until that h^j_{BS} is published.

In this step, on the one hand, the external malicious nodes couldn't get the authentication key to participate in the authentication work, and the external malicious nodes are excluded; on the other hand, if the compromised nodes forge or tamper with the information $L(j)$, that the wrong information $L(j)$ can be detected by the verification step whether the key h^j_{BS} is published based on $h^{j-2}_{BS} = H^2(h^j_{BS})$, and the compromised nodes are detected and excluded.

Therefore, Step 1 is called to be the security condition check because of the excluding of malicious nodes and compromised nodes.

Step 2. The time synchronization of the cluster head nodes.

After the completion of the security condition check in Step 1, it is the time to make time synchronization for the network. In this paper, the first time synchronization work is to realize the time synchronization of the cluster head nodes. Because the proportion of the cluster head nodes in the network is very small, TPSN is the most suitable synchronization protocol, which can avoid large amounts of computation in flooding network topology and keep high precision.

The realization of TPSN is divided into 2 stages: layer discovery and synchronization.

Stage 1. Layer discovery

Assume that BS is the first layer of network called Layer 0, and the cluster head nodes are divided into different layers by receiving the hierarchical data packet from their neighbor cluster nodes. Assume that the hierarchical data packet DP_i includes ID_{BS} and the layer grade L_i, set $DP_i = (ID_{BS}||L_i)$, such as that cluster nodes $CH_{1j}(j > 1)$ receive the packet $DP_0 = (ID_{BS}||L_0)$ from BS, they all needs to reset the layer grade L_1, and broadcast the new packet $DP_1 = (ID_{BS}||L_1)$ to the next neighbor nodes $CH_{2j}(j > 1)$. The layer discovery rule is that each cluster head just receives the first hierarchical data packet from the neighbor cluster heads.

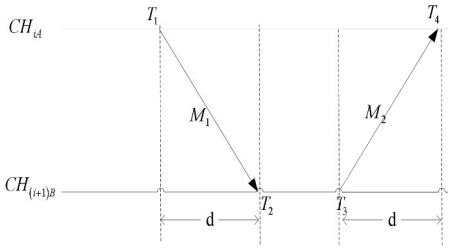

Figure 4: TPSN protocol

Stage 2. Synchronization

After layer discovery phase, BS starts the synchronization work by broadcasting the time synchronization data packet. As shown in Figure 4, there is a mutual information exchange between two neighbor cluster heads CH_{iA} and $CH_{(i+1)B}$, CH_{iA} sends the synchronization information packet at its' local time T_1, $CH_{(i+1)B}$ receives the packet at its' local time T_2, where $T_2 = (T_1 + d + \Delta)$, Δ is the time deviation between CH_{iA} and $CH_{(i+1)B}$, d is the signal propagation delays, $CH_{(i+1)B}$ returns the confirmation packet at its' local time T_3, and CH_{iA} receives confirmation packet at its' local time T_4, where $T_4 = (T_3 + d + \Delta)$, so we can get d and Δ as shown in following Equations (4) and (5):

$$d = \{ [(T_1 - T_2) + (T_4 - T_3)]/2 \} \quad (4)$$

$$\Delta = \{ [(T_1 - T_2) - (T_3 - T_4)]/2 \} \quad (5)$$

So $CH_{(i+1)B}$ can make its local time consistent with last layer node CH_{iA} based on d, and it shows that each cluster node can get the precise time same as time reference node BS.

Step 3. The time synchronization of common sensor nodes.

In Step 2, the cluster head nodes have all completed the synchronization work and become the reference nodes for their own cluster members.

As shown in Figure 2, there are many common sensor nodes belong to the one cluster head based on LEACH, so each common sensor node needs to make time synchronization work consistent with their cluster head based on TPSN, which will cause a large amount of computation and error accumulation. But the all cluster members

can achieve the time synchronization work by three times data communication based on HRTS protocol (as shown in Figure 5), which can reduce the computation greatly and maintain a high precision.

In the first time data communication, the reference node CH_{iA} broadcasts a synchronous request packet F_1 and records the sending time t_1, SN_i is the response node chosen randomly by CH_{iA} . All member nodes record the receiving time, and only SN_i needs to reply the response packet F_2, where SN_j records the receiving time t_{2j}, SN_i records the receiving time t_2, t_3 is the sending time of F_2 recorded by SN_i.

In the second time data communication, F_2 is sent to CH_{iA} by SN_i, where t_2 and t_3 are part of F_2, t_4 is the receiving time of F_2 recorded by CH_{iA}, so CH_{iA} can get t_1 to t_4 after receiving F_2 .

Assume that $\Delta^{'}$ is the time deviation between CH_{iA} and SN_i, $d^{'}$ is the signal propagation delays, T_r is the local time of CH_{iA}, T_p is the local time of SN_i, T_j is the local time of SN_j, so we can getT_r, t_2, t_4 in Equations (6) (7) (8):

$$T_r = \{ T_p - \Delta^{'} \} \quad (6)$$

$$t_2 = \{ t_1 + d^{'} + \Delta^{'} \} \quad (7)$$

$$t_4 = \{ t_3 + d^{'} - \Delta^{'} \} \quad (8)$$

And get $d^{'}$ and $\Delta^{'}$ in Equations (9) and (10):

$$d^{'} = \{ [(t_2 - t_1) + (t_4 - t_3)]/2 \} \quad (9)$$

$$\Delta^{'} = \{ [(t_2 - t_1) + (t_3 - t_4)]/2 \} \quad (10)$$

In the third time data communication, CH_{iA} broadcasts a new synchronous packet F_3 which includes t_2 and $\Delta^{'}$, and SN_j can correct its local time based on the following relationship in Equations (11)and (12):

$$T_p - T_j = \{ t_2 - t_{2j} \} \quad (11)$$

$$T_r = \{ T_j + t_2 - t_{2j} - \Delta^{'} \} \quad (12)$$

As the same way of SN_j, each cluster member node can correct its local time consistent with the reference node CH_{iA} and BS.

Step 4. Reverse authentication.

Although the synchronization work has been completed after Step 2 and Step 3, it's not sure whether all nodes in the network are synchronized, because there are some latent compromised nodes which can make some damage in the time synchronization process. So a reverse authentication method based on uTESLA is proposed for further detecting, and the specific steps are as follows:

Firstly, the cluster member node SN_i builds the reverse authentication information $L(j)^{'}$ (Equation (13)) and sends it to the cluster head CH_{iA} . Assume that $L(j)^{'}$ is the reverse authentication information of the time

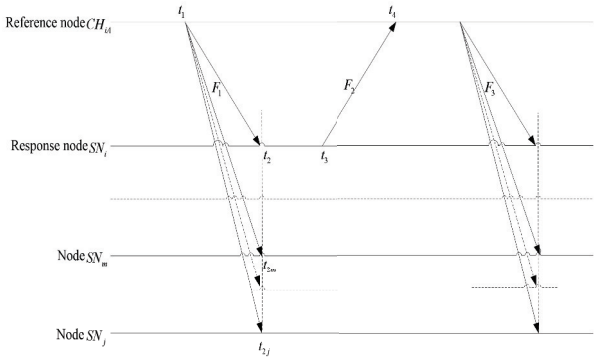

Figure 5: HRTS protocol

$j(t)$on time slice j, where t is a certain time on time slice j.

$$L(j)^{'} = \left\{ \begin{array}{l} h^{j}_{SN_i}(T_{SN_i}(j(t)))||h^{j-2}_{SN_i} \\ \\ ||ID_{SN_i}||T_{SN_i}(j(t)) \end{array} \right\} \quad (13)$$

Where, $T_{SN_i}(j(t))$ is the local time of SN_i, $h^{j-2}_{SN_i}$ is the published key by SN_i at $T_{SN_i}(j(t))$.

Secondly, CH_{iA} rebuilds the reverse authentication information $L_{CH_{iA}}(j)^{'}$ (Equation (14)) and sends it to last layer cluster nodes $CH_{(i-1)B}$ when obtains $L(j)^{'}$.

$$L_{CH_{iA}}(j)^{'} = \left\{ \begin{array}{l} h^{j}_{SN_i}(T_{SN_i}(j(t)))||ID_{SN_i} \\ \\ h^{n}_{CH_{iA}}(h^{j}_{SN_i}(T_{SN_i}(j(t))), \\ \\ T_{SN_{iA}})||h^{j-2}_{SN_i}||h^{n-2}_{SN_{iA}} \\ \\ ||ID_{CH_{iA}}||T_{SN_i}(j(t))||T_{SN_{iA}} \end{array} \right\} \quad (14)$$

Where, $T_{CH_{iA}}$ is the local time of CH_{iA} when broadcasting $L_{CH_{iA}}(j)^{'}$, $h^{n-2}_{SN_{iA}}$ is the published key by CH_{iA} at $T_{CH_{iA}}$.

Lastly, BS can get the information $L(j)^{'}$ based on receiving the reverse authentication information from the cluster nodes one by one, and waiting for getting the published key $h^{j}_{SN_i}$ by SN_i after δ to certificate the correctness of $h^{j}_{SN_i}(T_{SN_i}(j(t)))$ which can illustrate the identity of SN_i and the correctness of source attestation. In this time, each node can verify whether the key $h^{j}_{SN_i}$ is published based on $h^{j-2}_{SN_i} = H^2(h^{j}_{SN_i})$: if $h^{j}_{SN_i}$ has been published by SN_i, that each node which has obtained $h^{j}_{SN_i}$ can forge or tamper with the information $L(j)^{'}$, so the information $L(j)^{'}$ from these nodes will be judged to be unsafe, and abandon it, if $h^{j}_{SN_i}$ has not been published by SN_i, that each $L(j)^{'}$ will be cached until that $h^{j}_{SN_i}$ is published.

Firstly, BS can make a security condition check again by the reverse authentication in Step 4. Secondly, if $j(t)$ is much different with other cluster members after $h^{j}_{SN_i}$ published, it can be judged that SN_i is the compromised

node, and abandon it directly by BS. Thirdly, if the local time of most of the cluster members in CH_{iA} is much different with other cluster heads, it can be judged that CH_{iA} is the compromised node, and all the nodes associated with CH_{iA} are dangerous, so it needs to rebuild the network topology based on LEACH after abandoning CH_{iA}.

Therefore, the latent compromised nodes are detected and excluded in this step, which make the time synchronization work more secure.

The flow chart of STSP is showed in Figure 6.

4 Security Analysis and Simulation

4.1 Security Analysis

The main security problem of WSN time synchronization algorithms is the attack from the malicious nodes and the compromised nodes, the malicious nodes can be excluded by security condition check (such as Step 1 of STSP), but the compromised nodes can be latent down to waiting for an opportunity to attack, such as the attacker can send the cached legitimate data repeatedly to BS, which can cause a large amount of energy consumption. In addition, the false data forged by the compromised nodes in different geographic areas cant be detected and filtered.

The security characteristics of STSP proposed in this paper are analyzed as follows:

1) Excluding the malicious nodes

   Because of the one-way property of the hash key chain, the malicious nodes can't get the unpublished key, such as h_{BS}^{j} in $L(j) = (P_{j(t)}||h_{BS}^{j}(P_{j(t)}, T_{BS}(j(t)), ID_{BS})||h_{BS}^{j-2}||ID_{BS}||T_{BS}(j(t))$, where $h_{BS}^{j-2} = H^2(h_{BS}^{j})$, so that the external malicious nodes couldn't get the authentication key h_{BS}^{j} to participate in the authentication work and make any bad effect. Lastly, the malicious nodes will be detected and excluded by BS.

2) Making security condition check positively based on uTESLA

   Before time synchronization of network, each node can get the information $L(j)$ based on receiving the broadcast authentication information from their neighbor nodes one by one, and waiting for getting the published key h_{BS}^{j} by BS after δ to certificate the correctness of $h_{BS}^{j}(P_{j(t)}, T_{BS}(j(t)))$ which can illustrate the identity of BS and the correctness of source attestation. In addition, Each node can verify whether the key h_{BS}^{j} has been published based on$h_{BS}^{j-2} = H^2(h_{BS}^{j})$: if h_{BS}^{j} has been published by BS, that each node which has obtained the key h_{BS}^{j} can forge or tamper with the information $L(j)$, so the information $L(j)$ from these nodes will be judged to be unsafe, and abandon it; if h_{BS}^{j} has not been published by BS, that each $L(j)$ will be cached until that h_{BS}^{j} is published.

3) Detecting and excluding the latent compromised nodes based on uTESLA

   It's not sure whether all the nodes in the network are synchronized after the time synchronization work, because there are some latent compromised nodes which can make some damage in the time synchronization process, a reverse authentication method based on uTESLA (Step 4 of STSP) is proposed for further detecting: if the local time $j(t)$ of SN_i in CH_{iA} is much different with other cluster members after $h_{SN_i}^{j}$ been published, it can be judged that SN_i is the compromised node, and abandon it directly by BS, if the local time of most of the cluster members in CH_{iA} is much different with other cluster heads, it can be judged that CH_{iA} is the compromised node, and all the nodes associated with CH_{iA} are dangerous, so it needs to rebuild the network topology based on LEACH after abandoning CH_{iA}. Therefore, the latent compromised nodes can be detected and excluded by Step 4.

4) Small network load

   In DCS, it is bound to increase the redundancy of the message in the network and lead to the increase of the communication cost based on the flooding type network topology, because all nodes need to receive more than 2s+1 synchronization messages from last layer nodes.

   In STSP, the most classical clustering protocol LEACH in WSN is chosen to initialize the network topology in STSP, and the network topology structure based on clustering hierarchical is more suitable for WSN applications (as shown in Figure 2), which can avoid the over energy consumption of cluster head nodes, reduce the communication traffic effectively, and extend the life cycle of the network by 15%.

4.2 Simulation

In order to test the validity of STSP, a simulation work is made in the software platform of MATLAB R2014a to compare the difference in synchronization cycle, synchronization precision, synchronization ratio and synchronization cost between STSP and DCS.

The main simulation parameters are shown in Table 2:

Time synchronization cycle is the period that BS initiates the network time synchronization work to the end of the synchronization work, and the shorter the synchronization cycle, the better the convergence of the synchronization algorithm. In STSP, the synchronization cycle is the running time of Step 2 and Step 3. As shown in Figure 7, take the average value of 30 simulation data, it's

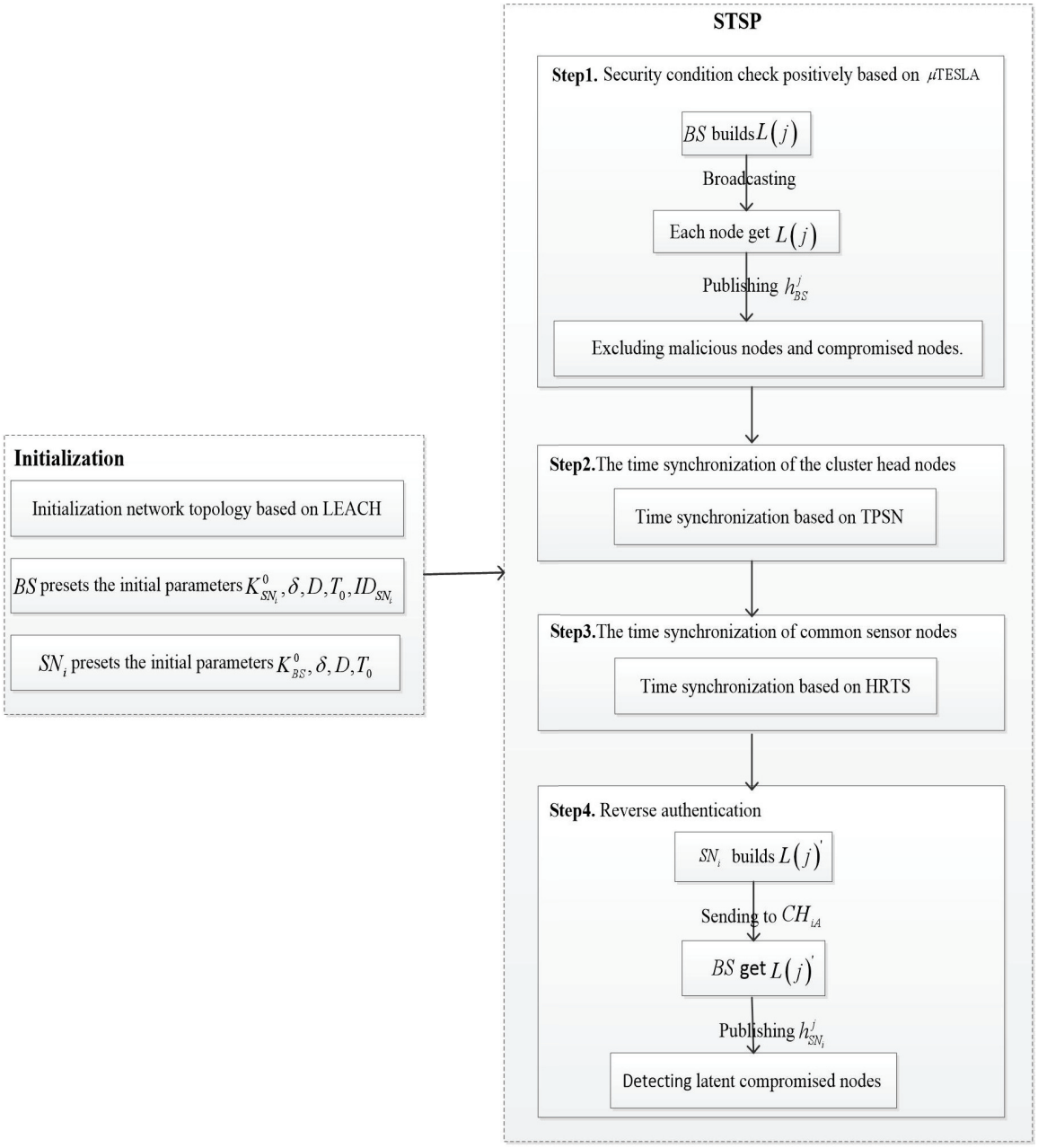

Figure 6: STSP flow chart

Table 2: Simulation parameters

Base station	1
Network size N	[50,100,150,.....,450,500]
Crystal oscillator	32MHZ
Area	1000m*1000m
Radius	500m
Modular	CC2430
Protocol	IEEE802.15.4
Rate	250kb/s
Power	20dBm
Malicious nodes MN	[0, 3, 5]

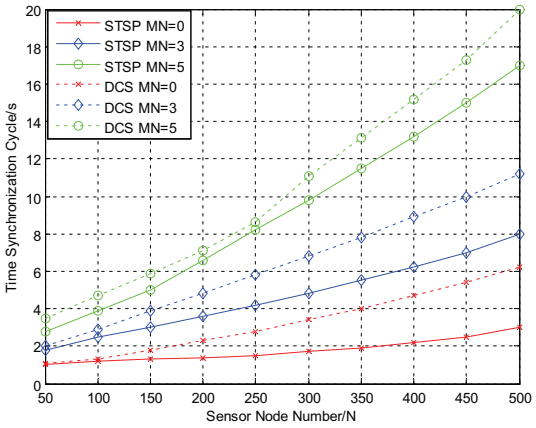

Figure 7: Time synchronization cycle

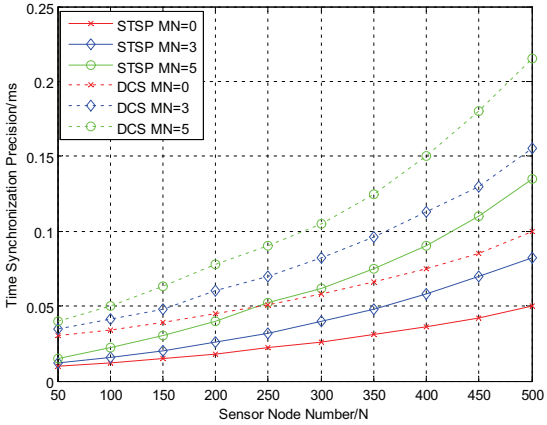

Figure 8: Time synchronization precision

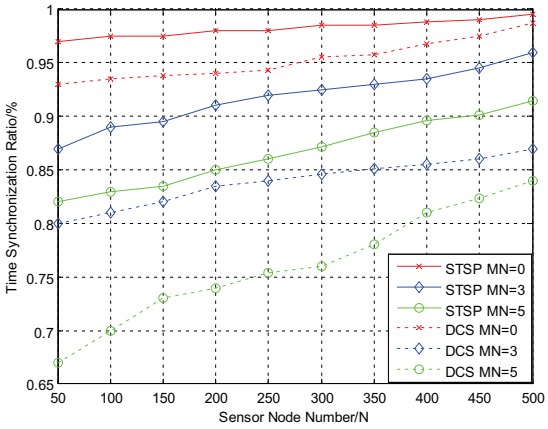

Figure 9: Time synchronization ratio

indicated that the synchronization cycle of these two algorithms will be extended with the increase in the number of network nodes, and the increase in the number of malicious nodes will extend the synchronization cycle too. In addition, the synchronization cycle of STSP is much better than DCS.

Synchronization error is the main characteristic of time synchronization precision, and the synchronization error is the time error between the network nodes and the base station. As shown in Figure 8, take the average value of 30 simulation data, it's indicated that the synchronization error will be increased with the increase in the number of network nodes, and the more the malicious nodes, the worse the synchronization precision. In addition, the synchronization precision of STSP is much better than DCS.

Synchronization ratio is the ratio between the synchronized nodes and the total network nodes, which embodies the security of time synchronization algorithms. As shown in Figure 9, take the average value of 30 simulation data, it's indicated that the synchronization ratio of STSP is much better than DCS. The reason is that not all the nodes in the network can receive more than 2s+1 synchronization messages from last layer in DCS, and it's easy to cause an attack from compromised nodes, but the malicious nodes and the compromised nodes will be detected and excluded by uTESLA in STSP.

Synchronization cost is the number of packets transmitted in once synchronization process. As shown in Figure 10, take the average value of 30 simulation data, it's indicated that the synchronization cost of STSP is much better than DCS. The reason is that it is bound to increase the redundancy of the message in the network and lead to the increase of the communication cost in DCS, that all nodes need to receive more than 2s+1 synchronization messages from last layer nodes, but the synchronization work in STSP only needs three times data communication.

5 Conclusions

A secure time synchronization protocol (STSP) for wireless sensor network based on uTESLA protocol is proposed according to the issues that DCS algorithm in wireless sensor network time synchronization is limited in the threats of compromised nodes, big communication cost and deficiency in coverage. Firstly, initializing the net-

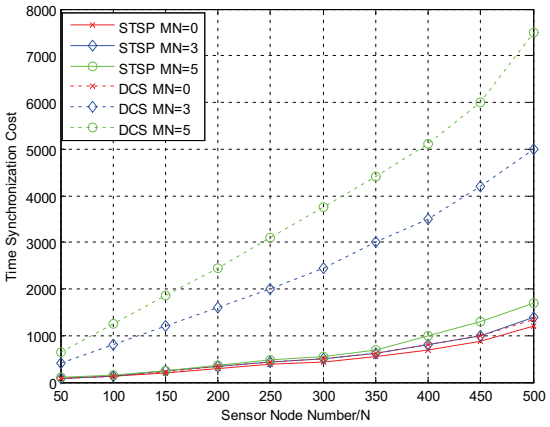

Figure 10: Time synchronization cost

work topology based on LEACH protocol to balance the network node load and prolong network life cycle. Secondly, excluding the malicious nodes based on uTESLA protocol for making security condition check positively before the network time synchronization. Thirdly, making time synchronization quickly based on TPSN protocol for base station to cluster head node. Fourthly, making time synchronization based on HRTS protocol for cluster head nodes to their own common nodes. Lastly, making reverse authentication based on uTESLA protocol for checking and excluding the compromise nodes after the network time synchronization. The security analysis and simulation show that two times authentication ensure the absolute security of the time synchronization work in STSP, and STSP is much better than DCS in synchronization cycle, precision, ratio and cost.

In addition, STSP in this paper still has much room for improvement based on the following reasons:

- Computation cost

  The uTESLA protocol has higher authentication efficiency in the case of sending data packets frequently, but it has a very low sending frequency in some applications, such as fire alarm and other event-driven applications, where the transmission interval of the adjacent data packets may be far greater than the time slice D of uTESLA, and causes lot of keys not used for the data packets authentication, the distance between adjacent keys on the key chain is also increased, and causes a large computation cost and authentication delay.

  Increasing D can alleviate this problem, but it also causes a lot of authentication delay, and the receiving nodes also need more memory space for buffering packets.

- Delay

  In uTESLA, the time interval of sending message $(MAC_{k_i}(P_i)||k_{i-2}||P_i||i(t))$ will be increased gradually, and the time for buffering data packets is also increased because of the authentication delay, which also makes the protocol more vulnerable to be attacked by DoS. Therefore, the authentication mechanism of uTESLA is not suitable for the situation of large sending time interval.

Acknowledgments

This work is funded by the National Science and Technology Planning of China (2015BAG10B00). The authors gratefully acknowledge the anonymous reviewers for their valuable comments.

References

[1] S. Ahmed, F. Xiao, and T. Chen, "Asynchronous consensus-based time synchronisation in wireless sensor networks using unreliable communication links," *IET Control Theory and Applications*, vol. 8, no. 12, pp. 1083–1090, 2014.

[2] I. F. Akyildiz, W. L. Su, and Y. Sankarasubramaniam, "Wireless sensor networks: A survey," *Computer Networks*, vol. 38, no. 1, pp. 393–422, 2002.

[3] D. Capriglione, D. Casinelli, and L. Ferrigno, "Analysis of quantities influencing the performance of time synchronization based on linear regression in low cost WSNs," *Measurement (02632241)*, vol. 77, no. 1, pp. 105–116, 2016.

[4] Z. Chen, D. Li, and Y. Huang, "Vent-triggered communication for time synchronization in WSNs," *Neurocomputing*, vol. 177, no. 2, pp. 416–426, 2016.

[5] H. Dai and R. Han, "Tsync: A lightweight bidirectional time synchronization service for wireless sensor networks," *ACM Mobile Computing and Communications Review*, vol. 1, no. 8, pp. 125–139, 2004.

[6] I. Davut, B. Kemal, and T. Bulent, "Evaluating energy cost of route diversity for security in wireless sensor networks," *Computer Standards and Interfaces*, vol. 39, no. 3, pp. 44–57, 2015.

[7] L. Deng, H. Huang, and Y. Qu, "Identity based proxy signature from rsa without pairings," *International Journal of Network Security*, vol. 19, no. 2, pp. 229–235, 2017.

[8] A. Diaz and P. Sanchez, "Simulation of attacks for security in wireless sensor network," *Sensors(14248220)*, vol. 16, no. 11, pp. 1–27, 2016.

[9] N. S. Fayed and E. M. Daydamoniand A. Atwan, "Efficient combined security system for wireless sensor network," *Egyptian Informatics Journal*, vol. 13, no. 3, pp. 185–190, 2012.

[10] S. Ganeriwal, R. Kumar, and M. B. Srivastava, "Timing-sync protocol for sensor networks," in *Proceeding of the 1st ACM Conference on Embedded Networked Sensor Systems (SenSys'03)*, Los Angeles, California, USA, 2003.

[11] Y. Gao, P. Zeng, K. K. R Choo, and F. Song, "An improved online/offline identity-based signature scheme

for WSNs," *International Journal of Network Security*, vol. 18, no. 6, pp. 1143–1151, 2016.

[12] J. He, J. Chen, and P. Cheng, "Secure time synchronization in wireless sensor networks: A maximum consensus-based approach," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 4, pp. 1055–1065, 2014.

[13] J. P. He, P. Cheng, and L. Hi, "Time synchronization in wsns: A maximum-value-based consensus approach," *IEEE Transactions on Automatic Control*, vol. 59, no. 3, pp. 660–675, 2014.

[14] W. R. Heinzelman, A. Chandrakasanand, and H. Balakrishnan, "Energy-efficient communication protocol for wireless micro-sensor networks," in *Proceeding of the 33rd Annual Hawaii International Conference on System Sciences(AHICSS'00)*, pp. 3005–3014, Maul: IEEE Computer Society, Jan 2000.

[15] W. K. Hu and J. C. Lin, "Ratio-based time synchronization protocol in wireless sensor networks," *Telecommunication Systems*, vol. 39, no. 1, pp. 25–35, 2008.

[16] G. Huang, A. Y. Zomayaand, and F. C. Delicato, "An cccurate on-demand time synchronization protocol for wireless sensor networks," *Journal of Parallel and Distributed Computing*, vol. 72, no. 10, pp. 1332–1346, 2012.

[17] G. Huang, A. Y. Zomayaand, and F. C. Delicato, "Long term and large scale time synchronization in wireless sensor networks," *Computer Communications*, vol. 37, no. 1, pp. 77–91, 2014.

[18] M. Jef, M. Sam, and H. Danny, "A comprehensive security middleware architecture for shared wireless sensor networks," *Ad Hoc Networks*, vol. 25, no. 2, pp. 141–169, 2015.

[19] M. Khurana, R. Thalore, and V. Raina, "Improved time synchronization in ML-MAC for WSN using relay nodes," *International Journal of Electronics and Communications*, vol. 69, no. 11, pp. 1622–1626, 2015.

[20] C. Lan, H. Li, S. Yin, and L. Teng, "A new security cloud storage data encryption scheme based on identity proxy re-encryption," *International Journal of Network Security*, vol. 19, no. 5, pp. 804–810, 2017.

[21] C. H. Ling, C. C. Lee, C. C. Yang, and M. S. Hwang, "A secure and efficient one-time password authentication scheme for WSN," *International Journal of Network Security*, vol. 19, no. 2, pp. 177–181, 2017.

[22] J. Liu, Z. Zhou, and Z. Peng, "Mobi-sync: Efficient time synchronization for mobile underwater sensor networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 2, pp. 406–416, 2013.

[23] D. L. Mills, "Adaptive hybrid clock discipline algorithm for the network time protocol," *IEEE/ACM Transactions on Networking*, vol. 5, no. 6, pp. 505–514, 1998.

[24] K. Pardeep and H. J. Lee, "Security issues in healthcare applications using wireless medical sensor networks: A survey," *Sensors (14248220)*, vol. 12, no. 1, pp. 55–91, 2012.

[25] A. Perrig, R. Szewczyk, J. D. Tygar, V. Wen, and E. C. David, "Spins: Security protocols for sensor networks," *Wireless Networks*, vol. 5, no. 8, pp. 521–534, 2002.

[26] U. A. Selcuk, R. A. Beyah, and J. A. Copeland, "Secure source-based loose synchronization (sobas) for wireless sensor networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 4, pp. 803–813, 2013.

[27] F. Sivrikaya and B. Yener, "Time synchronization in sensor networks:a survey," *IEEE Network*, vol. 4, no. 18, pp. 45–50, 2004.

[28] K. Sun, P. Ning, and C. Wang, "Secure and resilient clock synchronization in wireless sensor network," *IEEE Journal on Selected Areas in Communications*, vol. 2, no. 24, pp. 395–408, 2006.

[29] F. Wang, C. C. Chang, and Y. C. Chou, "Group authentication and group key distribution for ad hoc networks," *International Journal of Network Security*, vol. 17, no. 2, pp. 199–207, 2015.

[30] F. Wang, C. Yu, and X. Wu, "Dual time synchronisation method for wireless sensor networks," *Electronics Letters*, vol. 51, no. 2, pp. 1–2, 2015.

[31] B. S. Yosra and O. Alexis, "A lightweight threat detection system for industrial wireless sensor networks," *International Journal of Network Security*, vol. 18, no. 15, pp. 842–854, 2016.

Biography

Xiaogang Wang: College of Automation, Chongqing University, China. Major in wireless sensor network and security. Room 2508 in the main teaching building of Chongqing University, Shaping Ba distract, Chongqing city, China. (400044). Artificial Intelligence Key Laboratory of Sichuan Province, School of Automation and Information Engineering, Sichuan University of Science and Engineering, Sichuan 643000, China.

Weiren Shi: Prof. College of Automation, Chongqing University, China. Major in wireless sensor network and applications, information control and intelligent systems, embedded systems, pervasive computing, *etc.*

Survey of Peer-to-Peer Botnets and Detection Frameworks

Ramesh Singh Rawat[1,2], Emmanuel S. Pilli[3] and R. C. Joshi[1]
(Corresponding author: Emmanuel S. Pilli)
Department of Computer Science and Engineering, Graphic Era University, Dehradun, India[1]
Department of Computer Science and Engineering, Uttarakhand Technical University, Dehradun, India[2]
Department of Computer Science and Engineering, Malaviya National Institute of Technology, Jaipur, India[3]
(Email: espilli.cse@mnit.ac.in)
(Received Feb. 18, 2017; revised and accepted June 25, 2017)

Abstract

Botnet is a network of compromised computers controlled by the attacker(s) from remote locations via Command and Control (C&C) channels. The botnets are one of the largest global threats to the Internet-based commercial and social world. The decentralized Peer-to-Peer (P2P) botnets have appeared in the recent past and are growing at a faster pace. These P2P botnets are continuously evolving from diverse C&C protocols using hybrid structures and are turning to be more complicated and stealthy. In this paper, we present a comprehensive survey of the evolution, functionalities, modelling and the development life cycle of P2P botnets. Further, we investigate the various P2P botnet detection approaches. Finally, discuss the key research challenges useful for the research initiatives. This paper is useful in understanding the P2P botnets and gives an insight into the usefulness and limitations of the various P2P botnet detection techniques proposed by the researchers. The study will enable the researchers toward proposing the more useful detection techniques.

Keywords: Botnet Architecture; Detection Frameworks; Hybrid Botnets; Peer-to-Peer Botnets; Traffic Analysis

1 Introduction

Botnet is a network of compromised computers that are illicitly controlled and secretly used by attackers for various malicious operations. The attacker controlling the botnet is called bot master or bot herder. The compromised computers in a botnet are called drones or zombies and the malicious software running on them is known as bot. The term "bot" is derived from the word "robot" and it is a program used to automate tasks [8].

Botnets comprise of large pool of thousand to millions of compromised computers which empower the attackers with huge computational power and large band-width to launch attacks at global scale. Botnets are the largest threat to the cyber security of government, industries, academia and critical infrastructure *etc.* [4]. These provide large distributed platform to perform various malicious activities such as distributed denial-of- service (DDoS), spamming, phishing, spying, click-fraud, bitcoin mining, brute force password attacks and compromising social media service [10].

Many papers have surveyed the research literature of botnets [13, 18, 20, 26], but to the best of our knowledge, there is no schematic, analytical & comprehensive survey on the emerging P2P botnets and the detection methods. In this paper, we exclusively discuss various aspects of P2P botnets including: evolution, characteristics, C&C architecture and various detections methods. The paper is useful in understanding the characteristics of P2P botnets and the classification of the various detection techniques. It has the following important contributions:

- The timeline of evolution and development life cycle of P2P botnets is presented;
- The characteristics, architecture and functionalities of P2P botnets are described;
- The taxonomy and analytical survey of the various detection frameworks and models is presented;
- A discussion of the research challenges in detection and defence of these botnets is also included.

The remainder of the paper is organized from Section 2 to Section 7. Section 2 covers the background and related surveys. The botnet architecture and C&C protocols are discussed in Section 3. In Section 4, we have categorized the various detection frameworks. Further, the detection techniques are investigated in Section 5. In Section 6, we have presented the identified research challenges. Finally, Section 7 presents the summary and directions for future research work.

2 Background

Attackers exploit number of vulnerabilities and use social engineering techniques to infect more and more computers, network/IoT devices with the malware to build the botnet [4]. Many robust or stealthy botnets have evolved which wrecked havoc on the cyber security. During the onset of botnets the centralized architecture was popular. This architecture is easy to implement and monitor, but suffers from the limitation of easy detection of the centralized servers. Therefore, attackers adopt the decentralized or hybrid botnet architectures to overcome the limitations of the centralized architecture.

The decentralized P2P or hybrid botnets belong to the third generation. In P2P botnets attackers can directly communicate the commands to any peer bot; who further communicate the commands to other neighboring peers. These botnets have many robust features and stay resilient, *i.e.*, even if a significant part of the botnet is taken down, remaining botnet works under control of bot masters [6, 34]. P2P botnets are the most prevalent in the cyber world since they have many merits in comparison to centralized (IRC or HTTP) botnets.

3 Peer-to-Peer Botnets

Many large P2P botnets have been discovered in the wild. This clearly show a tendency towards the decentralized P2P botnets. These botnets use P2P networks as a vector for infestation and the peer nodes as C&C channels [1]. The P2P botnets form complex overlay networks using either customized or standard P2P protocols [1]. Figure 1 shows the timeline of the P2P botnets' evolution. Table 1 lists the characteristics and applications of the well-known P2P botnets developed over a period of time.

3.1 Botnet Architecture

Botnets are usually characterized on the basis of C&C architecture. The model of a typical botnet can be understood by the analysis of its architecture, C&C mechanism and its development life-cycle. The various phases of the botnets are shown in Figure 2. Cooke *et al.* [8] presented three different botnet communication topologies: centralized (IRC-based and HTTP-based), decentralized (P2P-based) and random. Grizzard *et al.* [13] classified the architecture of decentralized P2P botnets as: structured, unstructured, super-peers and hybrid.

Structured P2P Botnets: Use the overlay structured P2P network and thus employ a globally consistent protocol for efficient routing and search. Storm uses Overnet structured P2P overlay network protocol to build its C&C infrastructure. Most of the P2P botnets are based on structured overlays such as Kademlia. The botnets using public protocols let them mix the C&C traffic with the standards P2P applications.

Unstructured Botnet: Use P2P overlay links established arbitrarily and maintain neighboring peers list to ensure connectivity. This architecture is inherently flexible in the selection of neighboring peers and routing mechanisms. The unstructured botnets are difficult to be crawled and probe, since there is no specific structure that can be exploited.

Superpeer Overlay Botnets: Select some of the globally accessible compromised systems to form the C&C architecture. The compromised peers behind NAT are the normal peer bots and connect to any of the superpeers to pull published commands. Although the design is more scalable, but the superpeers are vulnerable to be detected. Further, the detection and removal of a superpeer does not have any significant impact on the botnet, since communication can be redirected to the new super-peers.

Hybrid P2P Botnets: Overcame the limitations of centralized and decentralized architectures. Many of the real botnets discovered in the wild have multi-layered hybrid P2P architecture. This structure employs top layer bots as master C&C servers. The P2P network serves as relay bots at the second layer communicating with the top servers and the bottom client as peer bots.

3.2 Command and Control Mechanisms

The command and control (C&C) channels are used to command the bots from remote system. The C&C layer forms the multi-tier architecture of the botnets and differentiate them from other malware. The various C&C architectures are based on different C&C protocols including IRC, HTTP, P2P, DNS, Bluetooth, email, social networks and/or other custom protocols [26, 10]. The botnets can select either P2P protocols or non-P2P protocols for C&C communication.

The C&C operations of the botnets are categorized into - **pull** and **push** mechanisms. In pull mechanism botmasters publish commands at certain specific locations for which the peer bots subscribe and actively receive the commands. The bots execute the commands and also forward to other peers in their list [24]. In push mechanism C&C servers push/forward the commands to peer bots for execution. The peer bots passively wait for the commands and also forward the received commands to neighboring peer bots. The bots can receive commands to execute using either a push or pull mechanism.

The C&C activities can be detected by active monitoring of the hosts connecting to the external suspected hosts. Identification of botnet C&C traffic is an important approach to botnet defense, but identification of C&C traffic is difficult since botnets use standard protocols for communications. Moreover, the malicious traffic is similar to legitimate traffic and is fused with the benign P2P communication traffic of the legitimate P2P applications (*eg*, Skype); Trojan.Peacomm botnet and Stormnet are

Table 1: P2P botnets and their characteristics

Botnet Name	Discovered	Architecture [a] / Protocols	Functionalities [b]	Comments
Sinit	Sep. 2003	D/P2P	DD	• Browser exploit, Random scan/probing, • Public key encryption
AgoBot	2003	C/IRC	DD, CS, SP, CF	• Spread using P2P Scanning, Disable AVs, • Uses SSL, Robust, Modular, Flexible
Slapper	2003	D/P2P	DD, CS	• Vulnerability exploit, • Difficult to monitor
Phatbot	Mar. 2004	D/IRC, P2P	DD, SP, MD, CS	• Multi exploit, Polymorphic, • Disable AVs, Not scalable
SpamThru	2006	H/Custom	SP	• Infection via e-mail, Encryption, • Spam templates server, 12,000 bots, 350 m spams/day
Nugache	Apr. 2006	D/Custom	DD	• Use random ports in C&C, E-mail infection, • Encryption, Resilient to take-down
Mega-D	2006	D/Custom	SP	• E-mail attachments, Polymorphic, • 10 billion spams/day, 250,000 infections,
Storm/Peacomm	Jan. 2007	D/P2P-Overnet	DD, SP	• Social engineering, C&C-Fast-flux, • Polymorphic, Hash encrypted, Disable AVs, 85,000-bots, 3 b spams/day
MayDay	Feb. 2008	H/HTTP,P2P	SP, PH	• Random anti-entropy based architecture, Hijacking browser proxy settings, • Encrypted ICMP, Limited C&C traffic, Web proxy,
Waledac	Dec. 2008	D/HTTP, P2P	DD, CS, SP, CS	• E-mail, social engg., Encryption, Packer, • Tunneling, Block DNS look-up, 7000 spams/day, shut down- Mar. 2010
Mariposa/Butterfly	Dec. 2008	D/Custom	DD, CS, SP, MD	• Code injection, Self-propagation, Obfuscation • Anti-debugging, 13,000,000 bots
Conficker C Conficker D, Conficker E,	Feb. 2009, Mar. 2009, Apr. 2009	H/HTTP, P2P	SP, DD	• Exploit-NetBIOS, Block DNS lookups, fast-flux, • Disables AVs & updates, 10,000,000+ bots, 10 b spams/day
Sality (v3, v4)	2009	D/Custom	SP	• Encryption, Resilient, Polymorphic, Disable AVs, • Custom P2P protocol over UDP
Miner	Dec. 2010	H/P2P	DD, MB, MD	• Social engineering, • Conceal C&C servers, Encryption
Kelihos	Dec. 2010	H/P2P-Custom	DD, SP, CS, MB	• Flash-drive, C&C proxies, Hidden-social networks, • 4 billion spams/day, Dismantled in Sep. 2011
ZeroAccess	Jul. 2011	D/P2P-Custom	CF, MB, PH, SP	• Exploit kit, Self healing P2P protocols, • Operation in user-mode, 9 m infections
TDL-4/ TDSS	2011	D/P2P-Kad	MD, CS	• Bootkit- infects MBR, DGA, Encryption, • Removes other malwares, 4.5 m infections
Gameover Zeus	Sept. 2011	H/HTTP, P2P-Kad	CS, SP, PH, DD	• Propagate-spams & phishing, RC4 encryption, • Anti-Crawling Technique, 3.6 m infections
Kelihos.B	Jan. 2012	D/P2P	MB, SB	• Spread via social networks, Encryption, • Sinkholed-March 2012
THOR	Mar. 2012	D/P2P	DD, SP, CS,	• Modules for sale, 256-AES encryption, • 8192-bit RSA instruction signing
Kelihos.C	Apr. 2012	D/P2P		• Stealing Internet browsers passwords, • Sinkholed in 2013
Wordpress/QBot	2013	-	CS, MD	• Crack administrative passwords, • 500,000+ infections, sniffed 800,000 transactions
newGOZ	Jul, 2014	D/DGA	SP,	• DGA generating 1,000 domains per day, • Use spam templates of Cutwail botnet,
Mac OS X botnet	Sep. 2014	-	CS, MD	• Request C&C servers list using MD5 hash of the current date • 17,000 unique IP addresses-Mac hosts

[a]Architecture: C: Centralized, D: Decentralized, H: hybrid,

[b]Functionalities: DD: DDoS, SP: Spamming, CS: Credential Stealing, MD: Malware Distribution, PH: Phishing, CF: Click Fraud, MB: Mining Bitcoins, SB: Stealing Bitcoins

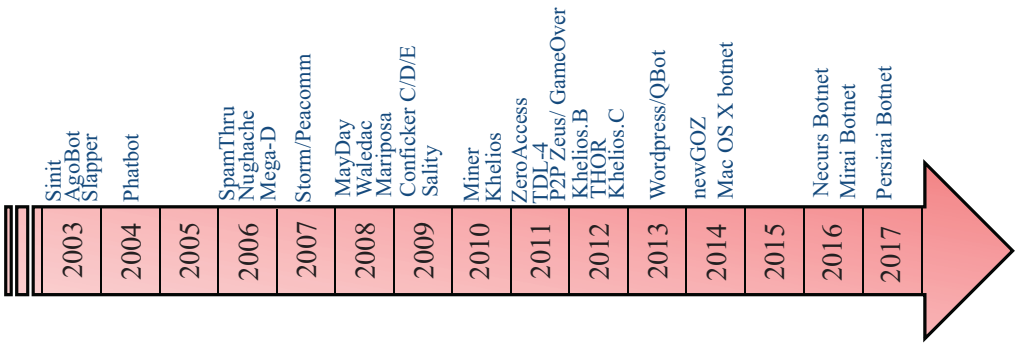

Figure 1: Development timeline of the P2P botnets

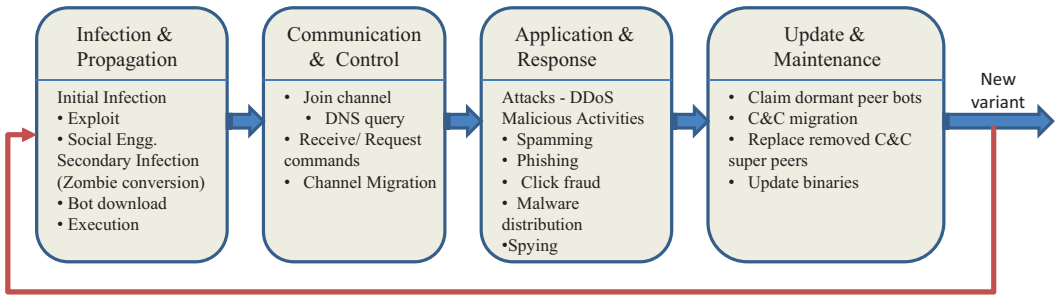

Figure 2: The development lifecycle of the botnets

two such examples [13]. Further, the C&C traffic of the botnets has low volume resulting into lower network latency. This challenges the threshold-based techniques to monitor and detect botnet activities.

4 P2P Botnet Detection Taxonomy

Botnet detection is the important step to combat the these threats. Researchers have analyzed the P2P botnets discovered in the recent past and identified the features which characterize them [13]. This has facilitated to develop various P2P botnet detection and mitigation techniques. Although, the available botnet detection and mitigation techniques have many great thoughts; but, these also suffer from many limitations to combat the real world botnet threats.

In this section, we present the taxonomy of the detection methods. The methods are classified into two major categories: (i) honeypot-based detection and (ii) Intrusion detection system (IDS)-based detection. These are again classified into further subtypes as explained in the subsections. The taxonomy of the detection methods is shown on Figure 3.

4.1 Honeypot-based Detection

A honeypot is a program that appears an attractive service, or an entire operating system to lure an attacker. It is a security resource designed to attract and detect the malicious operations and network attacks [3]. The group of honeypots is known as honeynet and are widely employed by the security communities to log bot activities and monitor the botnets. The logged data is analysed to discover the tools, techniques and motives of the attackers. Further, the obtained information is helpful to design the effective detection and mitigation techniques. This may also help to track the attackers for law enforcement [3].

Many researchers have discussed the use of honeypots/honeynets for botnet detection [8, 12, 26]. Cookie *et al.* [8] presented the use of honeypots to join botnet and monitor the activities. Freiling *et al.* [12] illustrated to prevent the DDoS attacks caused by the use of botnets. Further, honeypots can be deployed to join the botnets and serve as decoy system to provide valuable information about bots behavior and activities. Unfortunately, attackers employ anti-honeypot techniques to prevent any trap by the honeypots.

Honeypots trap the traffic directed to them only and cannot detect the real infected hosts in the enterprise network [3]. Moreover, honeypots need to be monitored to detect any anomalous behavior of the botnet. So, there is a need for developing more advanced honeypots for use

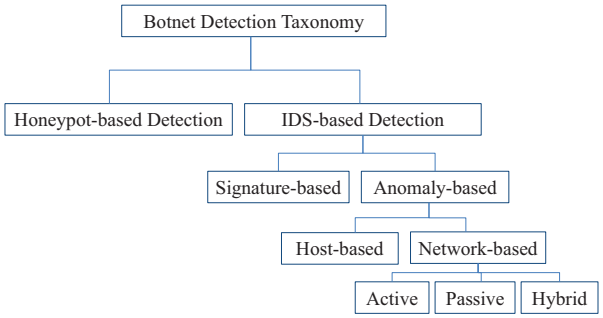

Figure 3: Botnet detection taxonomy

in complement of other botnet detection techniques.

4.2 IDS-based Detection

The IDS-based detection measures rely on the traffic collected at network-level using network sniffing intrusion detection tools and/or the network flow monitors. There are two major types of IDS-based detection techniques: (1) signature-based and (2) anomaly-based. The taxonomy shown in Figure 2.

4.2.1 Signature-based Detection

Signature-based methods use the knowledge of signatures and patterns for accurate botnet detection. The signature-based methods use Deep Packet Inspection (DPI) to inspect the malicious payloads, thus putting too much computational load on the system. These methods only detect the known botnets and are rendered useless by the unknown or even polymorphic botnets. Moreover, the signature-based systems require continuous update of the signature database. Botnet can also employ strong evasion mechanisms such as code obfuscation and encryption to bypass the signature-based detection methods [19, 34]. Therefore, anomaly-based botnet detection methods are proposed to keep up with the changing scenario of polymorphic botnet variants.

4.2.2 Anomaly-based Detection

Anomaly detection refers to finding the patterns that do not conform the normal behavior. These methods are based on the anomalies such as unusual system behavior, high traffic volume, high network latency and traffic on unusual ports. The security professionals and researchers have identified some of the inevitable P2P botnet characteristics such as high connection failure rate, out degree network connection and flow similarity. These characteristics suspect for anomalous behavior and are successfully exploited for the developing detection techniques.

The various anomaly-based botnet detection approaches can be classified into three major categories as: (i) Host-based, (ii) Network-based and (iii) Hybrid approaches.

Host-based Detection: Methods employ system calls monitoring for abnormal activities or data taint analysis techniques for detecting the malicious operations. The system also attempts to access system files to identify the suspected processes [29].

Wurzinger *et al.* [31] presented a host-based anomaly detection system based on aggregate network traffic features. The system inspect packet payloads to search for commonality likely to be C&C instructions from botmaster.

The host-based detection methods mostly inspect the packet payloads to identify a deviation from normal activity; so easily defended using encryption and obfuscation. These systems also suffer from the drawback that they need to be installed on every host. Such detection system demand lot of processing cycles, memory and storage space. Some malware can even disable the anti-malware product on the system it compromised. Therefore, these techniques typically suffer from the performance issues, scalability and effectiveness.

Network-based Detection: The network-based botnet detection methods focus on the network behavior of botnets and detect the C&C traffic between the servers and peers [14]. These detection approaches are generally based on either horizontal correlation (group behavior correlation) or dialog-based correlation (vertical correlation), which mainly utilize either network traffic analysis, aggregating network flows, or network behavior correlation analysis.

The network-based detection methods further employ either active detection or passive detection.

- Active Detection approach participate in the botnet operations. These approaches involve infiltration and C&C server hijack. Such approaches are based on the injection of test packets into the application, server or network for observing the reaction of the network. This produces extra traffic and sometime also violates the privacy.

- Passive Detection approaches silently observe, monitor and analyze the bots for their activities and do not make any efforts to participate in the operation. Passive traffic monitoring includes: behavior-based, DNS-based and data mining detection.

The network-based methods mainly monitor network traffic and can detect known and unknown botnets, but can be evaded by encryption and randomization. Further, these methods utilizing the behavioral characteristics of the bot produce visible network footprints as it works with other peer bots, communicates with botmaster and generates attack traffic. This lends itself exposed to the network-based detection.

Hybrid Detection: The researchers have proposed the host-network cooperative detection methods to overcome the limitations of individual host-level or

network-level detection methods [29]. These methods combine the evidences collected from host-based method with the network-based malicious behaviors method [29].

The different P2P botnet detection approaches usually include two steps: 1) hosts with P2P traffic are identified and separated from the normal traffic and then 2) hosts with P2P botnet traffic are detected and filtered from hosts performing only legal action. The next section groups the various detection proposals according to the detection algorithms used in the techniques.

5 P2P Botnet Detection Techniques

The classical botnets detection techniques have three different points to characterize the attacks: (1) command and control (C&C) server, (2) botnet traffic and (3) bot infected computers. Many detection schemes have been proposed to target either one or more of these points, *i.e.*, to detect either individual bots [29], C&C communication traffic and/or C&C server(s) [5]. We classify the various proposals into the following categories (shown in Figure 3):

5.1 Traffic-based Detection

The P2P bots communicate with many other peer bots to push/pull commands, send harvested information and receive updates; thus continuously generating large traffic [5]. Various traffic-based detection techniques have been proposed, which examine the network traffic and focus to observe the traffic patterns.

Noh *et al.* [23] proposed a botnet detection technique based on multi-phased flow model. The method clusters the flows of communication traffic and then build Markov models. The clustering of TCP/UDP connections form the grouping and track packets to determine if they are normal transmissions or flooding attacks. Further, the approach uses an algorithm to construct transitions based matrix of flow modeling and detection engine. But, the method can only detect P2P C&C traffic similar to the traffic used for training. Moreover, malware may avoid detection by using traffic patterns similar to legitimate P2P network.

Another method proposed by Jiang and Shao [16] detect P2P botnets based on the flow dependency in C&C traffic. The method distinguishes the normal P2P application traffic from P2P botnets by assuming that the normal traffic has heterogeneous short time flow dependency. The method also relies on discovering frequent flow dependencies. If these flows rarely happen, the approach may have difficulty to discover the flow dependency. Moreover, the proposed algorithm extracting flow dependency is based on time information and needs large number of samples and the results are based-on limited synthetic P2P botnet traffic trace samples. Further, the method scales quadratically with flow numbers and hence not scalable.

A large-scale wide-area botnet detection system DISCLOSURE identifies groups of features from the Net Flow records [5]. Authors use the group features to distinguish C&C channels from benign traffic. The system is independent of any knowledge of particular C&C protocols and has the ability to perform real-time detection of both known and unknown C&C servers over large datasets.

Zhang *et al.* [34] proposed a system to detect the stealthy P2P botnets. The system exploits the statistical fingerprints and is scalable by parallelized computation. The approach can also be defeated by advanced evasion and obfuscation techniques.

Kheir *et al.* [17] proposed a behavior-based approach called PeerMinor to detect and classify the P2P bots inside corporate networks. The system combines misuse and anomaly-based detection techniques and uses statistical network features: flow size, chunk rate, periodicities and IP distribution. PeerMinor classify P2P signaling flows and use them for the detection. It detects only P2P bots in monitored network and can be challenged by modifying the statistical consistency in the malware P2P flows.

Table 2 presents the characteristics and limitations/challenges of the various detection proposals. These are listed according to the detection techniques used in each proposal.

5.2 Behavior-based Detection

A comprehensive analysis of botnet measurements by Rajab *et al.* [26] reveals the structural and behavioral properties of botnets. Bots may also possess many inherent features, maintain the persistent connections to communicate with other peer bots and receive the commands from botmaster via C&C server(s). It is observed that the network behavior characteristics of P2P botnets are closely tied to the underlying architecture and operation mechanisms [28]. The bots in network immediately execute received commands and show similar communication behavior unlike human behavior. The traffic-based detection techniques mainly analyze the network behavior characteristics.

The botnet detection systems proposed in [7, 11] focus on the network hosts?behavior analysis using Netflows, which avoid the individual packet or host inspection and do not raise the privacy issues.

Felix *et al.* [11] proposed a scheme to detect P2P botnet based on set of group behavior metrics. The metrics are derived from the three standard network traffic characteristics: topological properties, traffic pattern statistics and protocol sequence for identifying hosts which have similar communication patterns. The approach needs multiple bots to be infected in the monitored network. Moreover, the threshold based filtering in the group behavior graph can be evaded by botmasters launching threshold attack.

Shin *et al.* [29] proposed a host-network cooperative behavior-based bot detection framework called EFFORT. The system relies on the client and network characteristics of bots. Although the method is independence of topology and communication protocol, but can be evaded by choosing suitable evasion techniques.

Rodrguez-Gmez *et al.* [27] proposed a method to detect parasite P2P botnets based on resource sharing. The method relies on the assumptions that the bot peers look for popular resources like the command files issued by the botmaster and further share them with other peer bots in a short period of time. The system only focuses on the parasite P2P botnets?detection in the command communication stage by looking and sharing for popular resources. No, real network attack and any other malicious activity can be detected.

5.3 DNS-based Detection

The bots possess a group activity as a key feature and frequently use DNS to rally C&C servers, launch attacks and update their codes. Bots of same botnet contact the same domain periodically leading to similar DNS traffic which is distinct from legitimate users [7]. In this section, we describe and evaluate the DNS-based anomaly detection techniques.

Choi and Lee [7] proposed an online unsupervised botnet detection technique called BotGAD (Botnet Group Activity Detector). BotGAD is implemented based on DNS traffic similarity and its performance is measured using real-life network traces. Botnets can evade the detection methods by performing DNS queries only once in the lifecycle. Hence, the method can detect only the botnets that perform group activities in DNS traffic. The method can also be evaded by botnets employing multi-C&C servers to separates their domain names.

5.4 Graph-based Detection

The graphical structure is an inherent feature of the botnets and is useful to understand how botnets communicate internally. The graphical analysis of the botnet communication network can be used to find the characteristic patterns of the botnets. The P2P C&C communications graph exhibit the topological features useful for traffic classification and botnet detection.

Ha *et al.* [15] analyzed the structural characteristics of Kademlia-based P2P botnets from a graph-theoretical perspective. The study analyzed the scaling, clustering, reachability and various centrality properties of P2P botnets. The authors also discovered that P2P mechanism helps botnets to hide their communication effectively.

Nagaraja *et al.* [22] proposed BotGrep to detect P2P botnets using the analysis of network flows collected over multiple large networks (eg, ISP networks). BotGrep first identifies groups of hosts that form a P2P network in the global view of Internet traffic. The algorithm is based on the premise that many recent botnets use efficient P2P protocols such as Kademlia for implementing C&C communications. But, BotGrep requires additional information to bootstrap the detection algorithm. Further, acquiring global view of Internet communications to bootstrap the detection algorithm may be very challenging.

Venkatesh *et al.* [30] proposed BotSpot to detect the hosts that are part of the structured P2P botnets. The authors developed algorithms based on the differences in the assortativity and density properties of the structured P2P botnets. BotSpot is based on the analysis of the IP-IP graph. It is complementary to the traffic classification approaches that differentiate between the structured P2P botnets and the legitimate structured P2P applications.

5.5 Data Mining-based Detection

The data mining techniques can be used to detect an anomaly *i.e.*, the unusual or fraudulent behavior. Data mining techniques are used for malicious code detection and intrusion detection. Many authors has used classification and clustering techniques to efficiently detect botnet C&C traffic.

The network traffic is a continuous flow of data stream produced at fast rate, which is not practical to be stored and analyzed entirely. Moreover, botnet codes, features and commands are updated frequently leading to dynamic and temporal behavior. Masud *et al.* [21] proposed stream data classification technique to detect P2P botnets. The authors proposed multi-chunk multi-level ensemble classifier to classify concept-drifting streams data. The approach is tested on limited synthetic data and Nugache botnet data collected in small setup, therefore, does not represent the real characterization of numerous botnets in the malware zoo.

Dietrich *et al.* [9] proposed CoCoSpot to detect botnet C&C channels based on traffic analysis. The system uses the periodicity of messages to suspect for C&C operations. Then, it classifies the P2P applications running on the host as malicious or benign based on payload analysis. The approach will result with false negatives if several messages are sent only in one direction.

Rahbarinia *et al.* [25] proposed a system called Peer-Rush to detect the unwanted P2P traffic. It creates pplication profile?based on the network flow features and inter-packet delays. The system can be evaded by the introduction of noise (random padding and false packets) in communication of bots. Further, the system cannot have much accurate results with the polymorphic and ever evolving P2P botnets.

5.6 Soft Computing-based Detection

Saad *et al.* [28] proposed an technique to detect P2P botnets using network traffic behavior analysis. The authors tested five different machine learning techniques to check for adaptability, novelty and early detection and get promising results using the limited test dataset. The technique is useful only for detection of P2P bots. Further, the

Table 2: Strengths and limitations of detection proposals

Detection Systems	Characteristics	Limitations or Challenges
—Traffic Analysis-based Techniques		
Noh *et al.* [23]	– Behavior & Traffic Analysis, Multi-phased flows model – C&C Traffic detection	– Evasion by using a legitimate P2P network
Jiang and Shao [16]	– Flow dependencies, Independent of malicious traffic,	– High false +ve, Dependency discovery
DISCLOSURE [5]	– Structure & protocol independent, Pattern based features, – Real-time & large scale	– Higher false positives
Zhang *et al.* [34]	– Statistical fingerprints-profile P2P traffic, – Persistent peer clients-similar traffic	– Evasion by blended peer bots and randomization, – Evasion- traffic tunneling through Tor network,
Kheir *et al.* [17]	– Based on P2P Signaling Flow, – Statistical features: Flow size, Chunk rate, periodicities and IP distribution	– Detect P2P bots only in a monitored network,
Behavior Analysis-based Techniques		
Felix *et al.* [11]	– Exploit Traffic pattern, – Bots group behavior	– Multiple bots dependency, Vulnerable to threshold attack
EFFORT [29]	– Host-network cooperation, Independent of topology & protocol, – Resilient to encryption & obfuscation,	– Evasion by bots: using benign domains
Rodrguez-Gmez *et al.* [27]	– Temporal resource sharing model – Monitoring resource sharing behavior	– Used only for parasite P2P botnets, – Source should be popular & short life
DNS-based Techniques		
Choi and Lee [7]	– Group Activity Detector, Online unsupervised known, Unknown, – Scalable, Real-time	– Requires multiple bots
Graph Analysis-based Techniques		
Ha *et al.* [15]	– Reachability & centrality properties – C&C channels detection, Monitoring bot activities	– Vulnerable to random delay, – P2P protocols dependency, False negatives
BotGrep [22]	– C&C patterns in overlay topology – Large-scale, Clustering techniques	– Bootstrap information required
Data Mining-based Techniques		
Masud *et al.* [21]	– Mining Concept-Drifting Data Stream – Packet features are extracted and aggregated into Flow characteristics	– Requires monitoring traffic at each host – Sampling may miss useful communications patterns
CoCoSpot [9]	– Analysis of traffic features – Fingerprint botnet C&C channels	– Evasion by random message padding – Dependency on the dialog-like pattern
PeerRush [25]	– Created application profile from known P2P applications – Based on high-level statistical traffic features	– Deals with the signaling flows as a whole – Evasion by randomization of inter-packet delays
Soft Computing-based Techniques		
Saad *et al.* [28]	– Traffic behavior, Detection in C&C phase – Detection rate 98%	– Dependency on features selection – High computational requirement
Zhao *et al.* [33]	– Anomalous Network traffic, – Real-time detection in C&C phase & attack phase	– Sampling can skip botnet flows, – Vulnerable to obfuscation
General Techniques		
BotMiner [14]	– Anomaly-based-behavior, Traffic-based analysis – Independent to protocol and C&C structure, Real-time	– Detect only active bot(s) – Targets enterprise network only
Wurzinger *et al.* [31]	– Network traffic, Bot behavior, Detect Bots, – No prior information required	– Threshold attack – Content analysis required
PeerPress [32]	– Remote control process- analysis, – Active-informed probing, Fast, Scalable, Real time	– False positives- advanced encryption, – Delayed port binding

technique also needs novel machine learning techniques for more effective results and general botnet detection.

Zhao *et al.* [33] proposed a technique to identify P2P botnet activities. The authors examine the characteristics of the traffic flows in small time windows to achieve the real time detection. But, the reduced time interval to monitor the traffic can skip some flows related to botnets. Further, botnets can also complicate the network flow behavior of the bots and evade detection.

Alauthaman *et al.* [2] proposed a technique using classification and regression tree algorithm and neural network (CART-NN) to detect the P2P bot connections. The technique use the connection-based features extracted from the TCP control packet headers. The method assumes that bots communicate using TCP connections and hence unable to cover the botnet using UDP connections.?

Chen *et al.* [6] proposed a botnet detection framework based on supervised machine learning technique. The framework use the conversation-based features extracted by random forest-based learning. The paper results explain the conversation-based features are better than flow-features, further, random forest is better than other classification algorithms giving the results upto 93.6%.

5.7 Generic Frameworks

A number of general botnet detection frameworks have been proposed based on behavior monitoring and traffic correlation analysis. BotMiner is a general framework for botnet detection [14]. The system detect botnets based on network packets and flow analysis. It relies on behavior monitoring and traffic correlation analysis that is mostly applicable at a small scale and does not scale well, because it requires analysis of vast amounts of fine-grained information. In addition, if there are only small numbers

Table 3: Evaluation of the P2P detection proposals

Detection Proposal	Detection Methodology [a]			Detection Stage [b]			Detection type (KN/UK/B) [c]	Real Time	Scalability [d]
	SB	HB	NB	IP	CC	AT			
BotMiner [14]	✓	✓	✓	✓	✓	✓	B	✓	S/M
Wurzinger *et al.* [31]	✓	✓			✓		KN		S
Ha et el. [15]	✓	✓	✓		✓	✓	KN		S
BotGrep [22]				✓	✓	✓	B		M/L
Saad *et al.* [28]			✓		✓		B	-	M
Choi and Lee [7]		✓	✓		✓	✓	B	✓	M, L
Jiang and Shao [16]			✓		✓		B	-	M
DISCLOSURE [5]			✓		✓		B		L
Felix *et al.* [11]		✓	✓	✓	✓	✓	B		S
Zhao *et al.* [33]			✓		✓	✓	KN		M
PeerPress [32]		✓	✓	✓	✓	✓	B		M, L
EFFORT [29]		✓	✓		✓	✓	B		S, M
PeerRush [25]			✓	✓	✓		KN		S
Rodrguez-Gmez *et al.* [27]			✓	✓	✓		UN	✓	L
Kheir *et al.* [17]			✓	✓	✓		B	✓	S

[a]Detection methodology- SB: Signature-based, HB: Host-based anomaly, NB: Network-based anomaly
[b]Detection phase- IP: Infection/Propagation, CC: Command & Control, AT: Attack
[c]Detection type- KN: Known, UN: Unknown, B: Both
[d]Scalability- S: Small, M: Medium, L: Large

of bots instances in the edge network, it leads to failure of bot coordination and resulting in false negatives. Moreover, BotMiner requires the malicious activities to be visible, thus cannot detect botnets in an early stage and does not work in real-time.

Wurzinger *et al.* [31] proposed a general system to detect bots. The system relies on the characteristics that each bot receives commands and responds in a specific way. It examines the packet payloads to find commonality to be supposed as commands from botmaster. The work complements the existing network-based IDSs by automatically generating the inputs needed by these systems to detect infected machines. But, the approach can be rendered useless by simple encryption in the C&C communication as used by the advanced botnets.

PeerPress is a P2P malware detection framework based on dynamic binary analysis and network level informed probe [32]. First it extracts built-in remotely accessible mechanisms of botnets called Malware Control Birthmarks (MCB) and then performs informed probe at network-level. The framework relies on malware opening service port for communications and provide malicious binary download services on the new infected machines.

The various frameworks use different data-sets of either synthetic or real botnets for evaluation and testing; therefore, the comparative analysis of the techniques is difficult. Table 3 presents the analysis of the promising detection frameworks.

6 Observations and Challenges

Many P2P botnet detection approaches evolved significantly, but many open problems still exist. We have identified the following open problems and research challenges useful for future research:

- The P2P botnets easily evade the port and protocol based detections by using both P2P and non-P2P ports for covert C&C communication. Further, bots traffic may blend with legitimate P2P application traffic.

- The P2P botnets also use fast-flux techniques for anomalous communications and hide the C&C hosts. Further botnets also use Domain Generation Algorithms (DGA) to keep the identity of C&C servers anomalous.

- Botnets also randomize the behavior and stay hidden under the radar to continue the malicious operations. This challenges the early stage detection if few bots exit in monitored network.

- The real-time online botnet detection needs to process huge amount of network traffic streams. Therefore, it is a challenge to develop the fast stream mining and classification algorithms for network traffic analysis to infiltrate the botnet traffic.

- The proliferation of smart-phones and other mobile devices with fast internet access provide new platforms the attackers to build mobile botnets.

- The fast adoption of Cloud computing is likely to attract the development of cloud botnets and is expected to be a big challenge for the security of the cloud computing.

The continuous advances in the botnet technology have enabled the attackers to evade the various detection measures. The exhaustive practices and covert network of the attackers enable them to stay ahead in pursuit of their malicious operations.

7 Conclusions and Future Work

In this paper, we have presented a comprehensive survey of various aspects of P2P botnets. Although the detection techniques have some strengths and scope, but, no single technique can detect such evolving botnets. Further, most detection schemes rely on the offline clustering and classification and does not cope-up the requirements of real time detection. Therefore, there is a requirement to develop a real time clustering and classification of the botnet traffic and on-the-fly mining of the botnet traffic to meet the requirements of real time botnet detections.

There is also a requirement to broaden the scope of detection and cover multiple botnet perspectives and also develop a collaborative detection framework. In our future research work, we would create a model to analyze the latest botnet(s) and develop a generic framework for the detection and mitigation of botnets. Further, extend the model to address the issues of mobile, cloud, social network-based botnets.

References

[1] D. Andriesse, C. Rossow, B. Stone-Gross, D. Plohmann and H. Bos, "Highly resilient peer-to-peer botnets are here: An analysis of gameover zeus," in *8th IEEE International Conference on Malicious and Unwanted Software (MALWARE'13)*, pp. 116–123, 2013.

[2] M. Alauthman, A. Nauman, L. Zhang, R. Alasem and A. Hossain, "A P2P botnet detection scheme based on decision tree and adaptive multi-layer neural networks," *Neural Computing and Applications*, pp. 114, 2016.

[3] P. Bacher, T. Holz, M. Kotter and G. Wicherski. "Know your enemy: Tracking botnets," *Technical Report, The Honeynet Project*, 2005.

[4] E. Bertino and N. Islam, "Botnets and internet of things security," *Computer- IEEE Computer Society*, vol. 50, no. 2, pp. 76–79, 2017.

[5] L. Bilge, D. Balzarotti, W. Robertson, E. Kirda, and C. Kruegel, "Disclosure: Detecting botnet command and control servers through large-scale netflow analysis," in *Annual Computer Security Applications Conference (ACSAC12)*, pp. 129-138, 2012.

[6] R. Chen, W. Niu, X. Zhang, Z. Zhuoand and F. Lv, "An effective conversation-based botnet detection method," *Mathematical Problems in Engineering*, vol. 2017, pp. 9, 2017.

[7] H. Choi and H. Lee, "Identifying botnets by capturing group activities in dns traffic," *Computer Networks*, vol. 56, no. 1, pp. 20–33, 2012.

[8] E. Cooke, F. Jahanian, and D. McPherson, "The zombie roundup: Understanding, detecting and disrupting botnets," in *Proceedings of the USENIX Workshop on Steps to Reducing Unwanted Traffic on the Internet (SRUTI '05)*, pp. 39–44, 2005.

[9] C. J. Dietrich, C. Rossow and N. Pohlmann, "Cocospot: Clustering and recognizing botnet command and control channels using traffic analysis," *Computer Networks*, vol. 57, no. 2, pp. 475–486, 2013.

[10] J. Echeverria and S. Zhou, "Thestar wars' botnet with¿ 350k twitter bots," *ArXiv Preprint ArXiv:1701.02405*, 2017.

[11] J. Felix, C. Joseph, and A. A. Ghorbani, "Group behavior metrics for P2P botnet detection," in *Proceedings of the 14th International Conference on Information and Communications Security (ICICS'12)*, pp. 93–104, Jan. 2012.

[12] F. C. Freiling, T. Holz, and G. Wicherski, "Botnet tracking: Exploring a root-cause methodology to prevent distributed denial-of-service attacks," in *Proceedings of 10th European Symposium on Research in Computer Security (ESORICS'05)*, pp. 319-335, 2005.

[13] J. B. Grizzard, V. Sharma, C. Nunnery, B. B. Kang, and D. Dagon, "Peer-to-peer botnets: Overview and case study," in *Proceedings of First Workshop on Hot Topics in Understanding Botnets (HotBots'07)*, pp. 1–8, Apr. 2007.

[14] G. Gu, R. Perdisci, J. Zhang and W. Lee, "Botminer: Clustering analysis of network traffic for protocol- and structure-independent botnet detection," in *Proceedings of the 17th USENIX Security Symposium (Security'08)*, pp. 139–154, 2008.

[15] D. T. Ha, G. Yan, S. Eidenbenz, and H. Q. Ngo, "On the effectiveness of structural detection and defense against P2P-based botnets," in *Proceedings of 39th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN'09)*, 2009.

[16] H. Jiang and X. Shao, "Detecting P2P botnets by discovering flow dependency in C&C traffic," *Peer-to-Peer Networking and Applications*, pp. 1–12, 2012.

[17] N. Kheir, X. Han and C. Wolley, "Behavioral fine-grained detection and classification of P2P bots," *Journal of Computer Virology and Hacking Techniques*, pp. 1–17, 2014.

[18] C. Y. Liu, C. H. Peng and I. C. Lin, "A survey of botnet architecture and batnet detection techniques," *International Journal of Network Security*, vol. 16, no. 2, pp. 81–89, 2014.

[19] Y. D. Lin, Y. T. Chiang, Y. S. Wu and Y. C. Lai, "Automatic analysis and classification of obfuscated bot binaries," *International Journal of Network Security*, vol. 16, no. 6, pp. 477–486, 2014.

[20] M. Mahmoud, M. Nir and A. Matrawy, "A survey on botnet architectures, detection and defences.," *International Journal of Network Security*, vol. 17, no. 3, pp. 264–281, 2015.

[21] M. M. Masud, J. Gao, L. Khan, J. Han and B. Thuraisingham, "Mining concept-drifting data stream to detect peer to peer botnet traffic," *Technical Report, University of Texas at Dallas, Richardson, Texas, Technical Report UTDCS-05-08*, 2008.

[22] S. Nagaraja, P. Mittal, C. Y. Hong, M. Caesar and N. Borisov, "Botgrep: Finding P2P bots with structured graph analysis," in *Proceedings of 19th USENIX Security Symposium*, pp. 95–110, 2010.

[23] S. K. Noh, J. H. Oh, J. S. Lee, B. N. Noh and H. C. Jeong, "Detecting P2P botnets using a multi-phased flow model," in *Proceedings of the Third International Conference on Digital Society (ICDS'09)*, pp. 247–253, 2009.

[24] P. Porras, H. Saidi, and V. Yegneswaran, "A multi-perspective analysis of the storm (peacomm) worm," *Computer Science Laboratory, SRI International, Technical Report*, 2007.

[25] B. Rahbarinia, R. Perdisci, A. Lanzi and K. Li, "Peerrush: Mining for unwanted P2P traffic," *Journal of Information Security and Applications*, vol. 19, no. 3, pp. 194–208, 2014.

[26] M. A. Rajab, J. Zarfoss, F. Monrose and A. Terzis, "A multifaceted approach to understanding the botnet phenomenon," in *Proceedings of the 6th ACM SIGCOMM Conference on Internet Measurement (IMC'06)*, pp. 41–52, 2006.

[27] R. A. Rodrguez-Gmez, G. Maci-Fernndez, P. Garca-Teodoro, M. Steiner and D. Balzarotti, "Resource monitoring for the detection of parasite P2P botnets," *Computer Networks*, vol. 70, pp. 30–311, 2014.

[28] S. Saad, I. Traore, A. Ghorbani, B. Sayed, D. Zhao, W. Lu, J. Felix and P. Hakimian, "Detecting P2P botnets through network behavior analysis and machine learning," in *Ninth Annual International Conference on Privacy, Security and Trust (PST'11)*, pp. 174–180, 2011.

[29] S. Shin, Z. Xu and G. Gu, "Effort: A new host-network cooperated framework for efficient and effective bot malware detection," *Computer Networks*, vol. 57, no. 13, pp. 2628-2642, 2013.

[30] B. Venkatesh, S. H. Choudhury, S. Nagaraja and N. Balakrishnan, "Botspot: Fast graph based identification of structured P2P bots," *Journal of Computer Virology and Hacking Techniques*, vol. 11, no. 4, pp. 247–261, 2015.

[31] P. Wurzinger, L. Bilge, T. Holz, J. Goebel, C. Kruegel, E. Kirda, M. Backes, and P. Ning, "Automatically generating models for botnet detection," in *Proceedings of the 14th European conference on Research in computer security (ESORICS09)*, pp. 232–249, 2009.

[32] Z. Xu, L. Chen, G. Gu and C. Kruegel, "Peerpress: Utilizing enemies P2P strength against them," in *Proceedings of 19th ACM Conference on Computer and Communications Security (CCS'12)*, pp. 581-592, 2012.

[33] D. Zhao, I. Traore, A. Ghorbani, B. Sayed, S. Saad and W. Lu. "Peer to peer botnet detection based on flow intervals," in *Information Security and Privacy Research*, vol. 376, pp. 87–102, 2012.

[34] J. Zhang, R. Perdisci, W. Lee, X. Luo and U. Sarfraz, "Building a scalable system for stealthy P2P-botnet detection," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 1, pp. 27–38, 2014.

Biography

Mr. Ramesh Singh Rawat received his M. Tech degree in Computer Science and Engineering from Graphic Era University in Dehradun, India in 2010. Afterwards, he is serving as Assistant Professor in Department of Computer Science at Graphic Era University. He is working toward Doctoral degree in Computer Science & Engineering from Uttarakhand Technical University, Dehradun, India. He is a member of IEEE, ACM CSTA, UACEE and CSI India. His research interests include Internet Security and Privacy, IDS and Botnet Defense.

Dr. Emmanuel S. Pilli received his M. Tech (Computer Science) from BIT Ranchi in 2001 and Ph. D (Computer Science) from the Indian Institute of Technology, Roorkee in 2012. He has over 20 years of teaching and research experience and is presently an Assistant Professor in the Department of Computer Science of Engineering, Malaviya National Institute of Technology Jaipur, India. His areas of interest include Security and Privacy, Forensics, Cloud Computing, Big Data and IoT. Dr. Pilli was awarded the ISEA Fellowship in 2011 by the Department of Information Technology, Govt. of India, for his research in Information Security. He is a Senior Member of IEEE and an active participant in professional activities of ACM, CCICI, CSA and CSI.

Prof. R. C. Joshi received M. E., Ph. D from University of Roorkee (now IIT Roorkee) in 1970 & 1980 respectively. He has over 45 years of teaching and research experience. He is presently Chancellor of Graphic Era University Dehradun, India. He was formerly Professor & Head, Electronics & Computer Engineering Department at IIT Roorkee, India. Prof. Joshi has guided 27 Ph.Ds and more mthan 250 M.Tech. dissertations. He has worked as a Principal Investigator in number of Sponsored Projects of Ministry of ICT, DRDO, AICTE, UNDP, ISEA *etc.* He has written 8 Books/Book Chapters and several technical reports. Prof. Joshi has published more than 250 Research Papers in International Journals / Conferences. He was awarded Gold Medal by Institution of India for Best Research Paper.

Bayesian-Boolean Logic Security Assessment Model for Malware-Free Intrusions

Aaron Zimba, Hongsong Chen, Zhaoshun Wang
(Corresponding author: Aaron Zimba)
Department of Computer Science and Technology, University of Science and Technology Beijing
Haidian District 100083, Beijing, China
(Email: gvsfif@gmail.com)
(Received Feb. 22, 2017; revised and accepted June 25, 2017)

Abstract

Attackers have come to leverage exploits precipitated by system vulnerabilities and lapses by using malware which otherwise tends to be noisy as it generates unusual network traffic and system calls. Such noise is usually captured by intrusion detection systems. Therefore, malware-free intrusions which generate little noise if any at all, are especially attractive to APT actors because they covertly use normal applications making it hard for intrusion detection systems. In this paper, we consider malware-free intrusions by formulating representations of system security states using Boolean logic in the scenario of a backdoor attack utilizing system implementation of pre-authentication services. We further derive, from the generated attack scenarios, a Bayesian security assessment model based on the environmental parameters of the experimental test-bed based on the backdoor attack via RDP-based remote access. The malware-free intrusion based on RDP backdoor attack is successfully run on five different versions of operating systems.

Keywords: Advanced Persistent Threat (APT); Bayesian Network; Boolean Logic; Malware-Free Intrusion

1 Introduction

Cyber networks today, the Internet inclusive, are plagued with a myriad of attacks all directed against Confidentiality, Integrity and Availability aspects of security. Attacks against these tenets of security, perpetrated by threat actors in the form of insider or outsider attackers [14], are categorized as either targeted or untargeted [5]. Advanced Persistent Threats (APT) belong to the latter classification and are thus carefully crafted owing to their nature since substantial knowledge of the victim is required prior to a successful attack. APT actors seek to maintain a long stealthy presence to further their attacks. Since the ultimate goal is not just to compromise a system and vacate in the shortest time possible but rather uphold the undetected presence feature, mechanisms and techniques employed to achieve the desired intrusion play a vital role. Attackers therefore use complex techniques to compromise systems which include leveraging vulnerabilities in system software, flaw in design and implementation of a security system as well as ignorance of a benign user. APT attackers use special malware with characteristics different from conventional malware in that such malware may hibernate and remain dormant for long periods before initiating the actual attack or beacon back to the Command and Control (C2) for further directives [22]. But the presence of Intrusion Detection Systems (IDS) present a higher chance of discovery against such malware even as its signature and activities might not escape the eye of the IDS. Since the noise generated by malware in form of network activity, issuance of system calls and signature of the malware itself form the basis upon which IDSs detect such malware [12, 16, 18], an ideal intrusion appealing to the attacker would be one that operates outside the realms of the aforementioned detection parameters.

Malware-free intrusions fit well in the above desirable intrusion requirement because they utilize normal system files and processes to covertly achieve their goal. The network activity generated by normal system processes, associated system calls and even their file signature values all fall under the threshold for normal file classification of the IDS. Attackers have therefore come to exploit shortfalls in the security implementation of systems without using malware to attain the desired malware-free intrusion. One such leveraged security shortfall which has seen considerable usage by various APT actors is the accessibility services backdoor [4, 15]. The attack pursued via this route avails the attacker system level access without logging in at all. Access is achieved by invoking corresponding accessibility keystrokes on a compromised system and this access is also possible over the network through RDP-based remote access. We explore two attack vectors emanating from the attack space cast by the aforementioned

backdoor attack. We employ a conceptual finite state automaton to deduce state equations representative of the various system security states relative to the pursued attack vector and use attack tree analysis for formulation of the security assessment model based on Bayesian inference. We carry out experimental attack tests on an IP network with different versions of operating systems and associated security implementations.

This paper is organized as follows: the second section encompasses state model formulation and analysis whilst the experimental test-bed is discussed in the third section. The security assessment model is derived in section four and we conclude the paper in the fifth section.

2 Model Formulation and Analysis

The accessibility backdoor considered in this paper affects Windows operating systems from Windows XP to Windows 10. Though we do not include server versions of Windows operating system, we contend that the techniques employed herein are applicable thereto provided the server operating system in contention ships with accessibility services and RDP-based remote access, which is the case by default.

2.1 Transitions of System Security States

We construct a conceptual model, as depicted in Figure 1, by employing a finite state machine where the security status of the given system is defined by the state of three attack vectors. The first two attack vectors pursued henceforth are the backdoor implantation variants through which system level access is made available at pre-login before authentication while the third vector is activation of remote access using the system inbuilt Terminal Services.

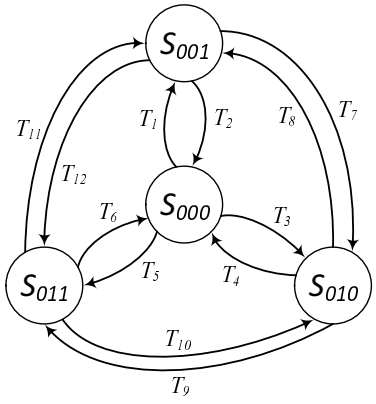

Figure 1: State transition diagram of system security states

The four states of the system are denoted by S_n, where n increments by a single binary unit upon successful implementation of a given attack vector. Transitions from one state to another are denoted by T_n and we use binary state encoding for security state assignment and thus denote the states as follows:

- S_{000} - the initial secure state of the system in the absence of pursuance of all the three attack vectors.
- S_{001} - the state of the system when only one of the three attack vectors is successfully pursued.
- S_{010} - the state of the system when any two of the three attack vectors are successfully pursued.
- S_{011} - the state of the system when all the three attack vectors are successfully pursued.

It is evident from Figure 1 that transitions T_1, T_3 and T_5 are induced by successful pursuance of the associated attack vectors thereby inducing state transitions from S_{000} to S_{001}, S_{010} and S_{011} respectively. On the contrary T_2, T_4 and T_6 are as a result of mitigating the associated security breaches emanating from the corresponding attack vectors. Transitions T_7, T_9 and T_{12} are perturbed by an increment in the number of successfully pursued attack vectors whereas transition T_8, T_{10} and T_{11} reflect the opposite. It should be noted however that the order in which the attack vectors are pursued, hence the corresponding state transition, is not relevant but rather the fact that the attack vector is pursued unto completion.

Let the result of backdoor implantation by the first attack vector via replacement of accessibility executable binary in the %systemroot% \system32 directory be denoted by the binary variable α. And let the result of backdoor implantation by the second attack vector via the system registry be denoted by the binary variable β. Lastly, let the result of activation of RDP-based remote access via system registry alteration be denoted by the binary variable γ. Thus with regards these three binary variables, the security status of the system at any given instance can be formulated as a Boolean function:

$$S_n(\alpha, \beta, \gamma),\ where\ \alpha, \beta, \gamma \in \mathbb{N}_2$$

Consequently since $\alpha, \beta, \gamma \in \{0, 1\}$, it follows henceforth that the false values (binary 0) i.e. when the result of backdoor implantation and RDP-based remote access are unsuccessful thereby inducing no security breach are denoted by way of complementation, hence yielding the variables $\bar{\alpha}$, $\bar{\beta}$ and $\bar{\gamma}$. The possible system security states are given by:

$$q = log_2 n \tag{1}$$

where $q, n \in \mathbb{N}^+$ are the state variables and number of states respectively. Therefore the cardinality S_n for the binary variables α, β and γ using Equation (1) is;

$$|S_n| = 8\ where\ n\ is\ \{n : n \in \mathbb{N}^+,\ 0 \leqslant n \leqslant 3\}$$

The 8 states are spread across S_{000}, S_{001}, S_{010} and S_{011} and we use these to construct the corresponding truth table and k-maps to derive state binary equations representative of the various system security states. Since the input combinations of all states are identical, we use one integrated table instead of four where we only differentiate the output. The resulting truth table of all possible states at any instance is shown in Table 1.

Table 1: Truth table for possible security states

Input Bin. Variables			*Output State* S_n
$Var_1\alpha$	$Var_2\beta$	$Var_3\gamma$	$S_{000}/S_{001}/S_{010}/S_{011}$
F	F	F	*1/0/0/0*
F	F	T	*0/1/0/0*
F	T	F	*0/1/0/0*
F	T	T	*0/0/1/0*
T	F	F	*0/1/0/0*
T	F	T	*0/0/1/0*
T	T	F	*0/0/1/0*
T	T	T	*0/0/0/1*

The security requirement of the initial state S_{000} dictates that there be no breach in the system implying that all input variables be false. This implies a conjunctive Boolean AND operation on all the three complemented input variables. To derive the state equation, we employ the K-map in Figure 2 below and apply the Product of Sums (POS) on the dotted groups and Sum Of Products (SOP) on the solid group for equation validation.

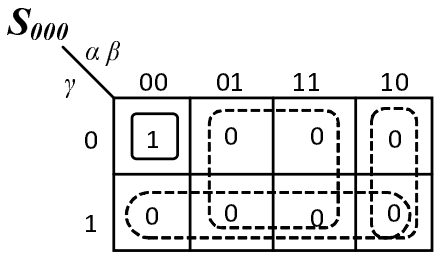

Figure 2: K-map for the single state S_{000}

We use the canonical disjunctive normal form of the maxterms corresponding to the dotted groups of *0*s and thus derive the complemented equation for the initial state:

$$\bar{S}_{000}(\alpha,\beta,\gamma) = \alpha\cdot\bar{\beta}+\beta+\gamma$$

Applying the De Morgan theorem and applicable Boolean identities, we derive the Equation (2) representative of this first secure state:

$$\therefore S_{000}(\alpha,\beta,\gamma) = \bar{\alpha}\cdot\bar{\beta}\cdot\bar{\gamma}\ where\ \alpha,\beta,\gamma\in\mathbb{N}_2 \qquad (2)$$

The state Equation (2) validates with one obtained via SOP of the minterms of the solid group.

When one and only one of the three attack vectors is pursued to fruition in the initial state, the security status of the system transitions via T_1 from state S_{000} to 3 possible states of S_{001}. We use the K-map below in Figure 3 derived from the integrated truth Table 1 to deduce the state equations representative of the new system state.

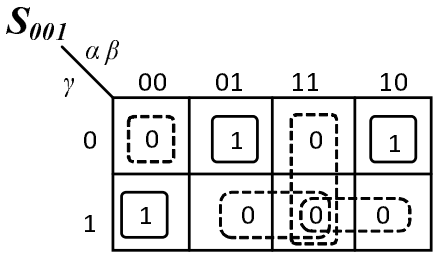

Figure 3: K-map for three possible S_{001} states

We order the (α,β) variable pair in our K-maps in the sequence $01\rightarrow 11$ and not $01\rightarrow 10$ so as to avoid race conditions and static hazards. Therefore, in view of the aforementioned, the isolated groups of *1*s and *0*s in our K-maps do not denote "don't care"entries but are rather inputs of the SOP and POS minterms and maxterms respectively. The equations representative of the new state are thus minimized via SOP as:

$$S_{001}(\alpha,\beta,\gamma) = \bar{\alpha}\cdot\beta\cdot\bar{\gamma}+\bar{\beta}\ (\alpha\oplus\gamma)$$

$$S_{001}(\alpha,\beta,\gamma) = \alpha\cdot\bar{\beta}\cdot\bar{\gamma}+\bar{\alpha}\ (\beta\oplus\gamma)$$

$$S_{001}(\alpha,\beta,\gamma) = \bar{\alpha}\cdot\bar{\beta}\cdot\gamma+\bar{\gamma}\ (\alpha\oplus\beta)$$

$$\therefore S_{001}(\alpha,\beta,\gamma) = \{\alpha,\beta,\gamma \mid \bar{\alpha}\cdot\beta\cdot\bar{\gamma}+\alpha\cdot\bar{\beta}\cdot\bar{\gamma}+\bar{\alpha}\cdot\bar{\beta}\cdot\gamma\} \quad (3)$$

The state Equation (3), depicting the three possible states of S_{001}, correlates with canonical disjunctive normal form of the POS where the maxterms correspond to the dotted groups of *0*s in Figure 3, hence the validation.

If another attack vector of the remaining two is pursued to completion whilst in state S_{001}, the system transitions via T_7 to any of the three possible states of S_{010} bringing the sum of successfully pursued attack vectors to 2. The K-map for these three new possible states is shown below in Figure 4.

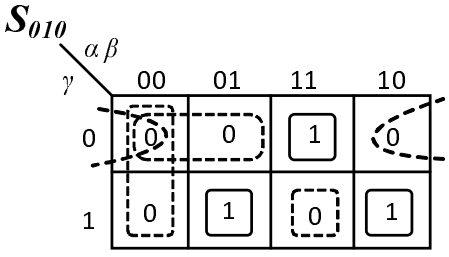

Figure 4: K-map for three possible S_{010} states

The sequence for variable pair ordering on the horizontal axis in Figure 4 likewise follows suit as that of Figure 3. We therefore employ the SOP based on this K-map

to derive the equations representative of the three new possible states:

$$S_{010}(\alpha,\beta,\gamma) = \bar{\alpha}\cdot\beta\cdot\gamma + \alpha\ (\beta\oplus\gamma)$$

$$S_{010}(\alpha,\beta,\gamma) = \alpha\cdot\bar{\beta}\cdot\gamma + \beta\ (\alpha\oplus\gamma)$$

$$S_{010}(\alpha,\beta,\gamma) = \alpha\cdot\beta\cdot\bar{\gamma} + \gamma\ (\alpha\oplus\beta)$$

$$\therefore S_{010}(\alpha,\beta,\gamma) = \{\alpha,\beta,\gamma \mid \bar{\alpha}\cdot\beta\cdot\gamma + \alpha\cdot\bar{\beta}\cdot\gamma + \alpha\cdot\beta\cdot\bar{\gamma}\} \quad (4)$$

The Equation (4) depicts the three possible states with only two attack vectors yielding fruition. Likewise it correlates with its dual obtained by the canonical disjunctive normal form of the POS maxterms, hence the validation.

Now that the system is in a state with two attack vectors pursued to completion hence two system breaches, there only remains one attack vector to be pursued which transitions the system into the final state S_{011}. We yet again use a K-map, Figure 5, to derive the state equation representative of this new state.

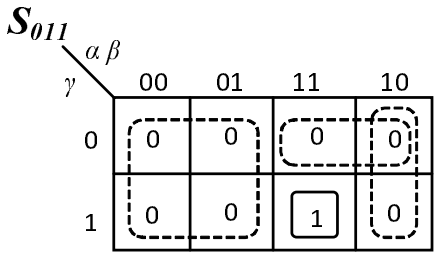

Figure 5: K-map for the final state S_{011}

We this time employ POS of the maxterms of the canonical disjunctive normal form of the dotted groups yielding:

$$\bar{S}_{011}(\alpha,\beta,\gamma) = \alpha\cdot\bar{\beta} + \bar{\alpha} + \bar{\gamma}$$

We further minimize the equation with Boolean identities to find the final compact status equation as:

$$S_{011}(\alpha,\beta,\gamma) = \alpha\cdot\beta\cdot\gamma \ where\ \alpha,\beta,\gamma \in \mathbb{N}_2 \quad (5)$$

This final equation is validated from the K-map by computing the SOP of the single solid group of *1*s. It is apparent from Equation (5) that the system cannot transition into any less secure state than S_{011} because all the attack vectors have been exhausted and the equation itself is a Boolean conjunctive AND operation on all the three variables.

According to the derived four state equations, the eight states from Equation (1) are partitioned as follows:

- One state in S_{000} denoted by Equation (2);
- Three states in S_{001} denoted by Equation (3);
- Three states in S_{010} denoted by Equation (4);
- One state in S_{011} denoted by Equation (5).

Notwithstanding the aforementioned, not all the eight states represent a successful attack though they might imply the presence of a breach due to a specific pursued attack vector. We are now therefore tasked to find out which combinations of pursued attack vectors lead to a successful malware-free intrusion attack. We address this task in the proceeding sub-section where we describe the attack model.

2.2 Attack Modeling and Analysis

We approach attack modeling based on conceptual units [2,8,17] where such discrete units serve as the basic building blocks of the attack. The threat actor who is an attacking agent, executes a series of actions to obtain assets as pivots for reaching the final goal. Therefore our model comprises four units namely assets, actions, agents and goals.

2.2.1 Model Units Formulation

Assets: Assets are anything the attacker needs to acquire not only for optimal output but for actualization of the attack itself as well. In our context, assets include but are not limited to information about a victim host such as IP address, open ports and their associated protocols, underlying operating system, service banner information etc. This is the knowledge domain that the attacking agent has of the target in contention.

Actions: Actions are steps of sequential phases that constitute an attack. Actions have preconditions which foster acquisition of a sought after asset, that is to say that the outcome of an action is whether the asset is acquired or not. The actions in our setting include initiation of the pursuance of the earlier mentioned attack vectors, finding target hosts in the environment, invocation of console system level access via appropriate keystroke combinations etc.

Agents: Agents are the subject of any given attack scenario whose actions are directed towards a specified object. They can broadly be distinguished as human or software actors. The agent in our consideration is a highly skilled technical human actor with a considerable sophistication of stealthiness and non-traceability.

Goals: Goals represents a request knowing all action outcomes which in turn complete the associated assets. Goals are differentiated depending on the context, and in ours, goals include establishment of a remote access connection via the RDP protocol provided that the port scan action returned a value that Terminal Services are available on a target host.

Having defined the components of our model, we now integrate them into attack tress for modeling and analysis.

2.2.2 Attack Tree Modeling Integration and Analysis

The units of the preceding subsection are integrated into an attack tree [13] for modeling and analysis. Here we describe the backdoor attack against a victim host where the nodes, represented by the model units, require complete execution of children nodes to reach the root node which is the sought after system level access via RDP-based remote access. The nodes are either conjunctive AND nodes or disjunctive OR nodes. All children of an AND node need to return a true value if the parent node is to execute successfully while only one or more of an OR node need to be true to accomplish the same. The resultant attack tree is shown in Figure 6 below.

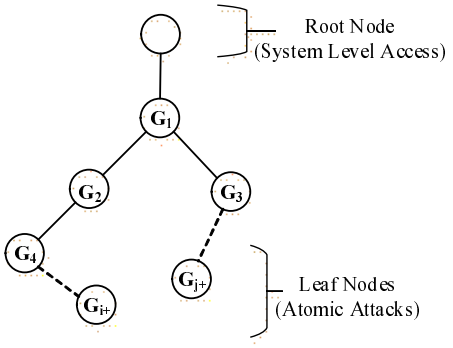

Figure 6: Attack Tree for backdoor attack

The root node is denoted by G_0 which is acquisition of system level access over the network. This is achieved by pursuing the sub-goals originating from the lower leaves. The rest of the nodes are denoted as follows: G_1 –RDP Remote Access Activation, G_2 –SystemRoot Bin. Replacement, G_3 –Registry Debugger Conf., G_4 –Local Privilege Escalation, G_j+ denotes a set of nodes decomposed into conjunctive AND or disjunctive OR representing attack model units needed to be engaged if the pursued attack vector of backdoor implantation is that of registry alteration while G_i+ denotes those where the backdoor is implanted via binary executable replacement. We deduce the adjacency square matrix A_G of the 5th order from the graph after pruning out Gi+ and G_j+:

$$A_G = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \quad (6)$$

We now use Equation (6) of the above matrix rows and columns to derive attack scenarios corresponding to the following paths:

$$P_1 : \{G_4, G_2, G_1, G_0\}$$

$$P_2 : \{G_3, G_1, G_0\}$$

It is evident from the above paths that P_1 is longer than P_2 because pursuance of the former calls for addressing an additional unit of privilege escalation. However, this does not necessarily imply that P_2 is a better path than P_1 because the value of the weights of the edges might tell otherwise depending on the metrics used.

The edge $\{G_1, G_0\}$ represents the isthmus of graph implying that failure to actualize an action associated with this edge thwarts the backdoor attack. The action associated with this edge is activation of RDP-based remote access and this corresponds to the γ binary variable of Equation (2), (3), (4) and (5). Since the value of γ in Equation (2) is definitely false, it follows that the backdoor attack is not feasible and we therefore drop this equation for simulation considerations. Likewise, Equation (3) reduces to the form $S_{001}(\alpha, \beta, \gamma) = \bar{\alpha} \cdot \bar{\beta} \cdot \gamma$ as we drop all the minterms with $\bar{\gamma}$. In the same manner, Equation (4) reduces to $S_{010}(\alpha, \beta, \gamma) = \gamma\ (\alpha \oplus \beta)$ and we use Equation (5) as-is considering that it has no complemented values of γ. This gives us a system of compact equations:

$$S_n(\alpha, \beta, \gamma) \quad = \quad \left\{ \begin{array}{l} \alpha \cdot \beta \cdot \gamma \\ \gamma\ (\alpha \oplus \beta)\ where\ \alpha, \beta, \gamma \in \mathbb{N}_2 \\ \bar{\alpha} \cdot \bar{\beta} \cdot \gamma \end{array} \right\}$$

We therefore base our simulations on this system of equations in the next section. We test the attack on different combinations of literal and complemented values of the binary variables and XORing where applicable according to the equation.

3 Experiment Simulations

The experiment setup consist of two networks; that of the attacker and one where the targeted hosts reside. We build our test-bed environment in Virtual Box where we simulate the internetwork connection and the network attack itself.

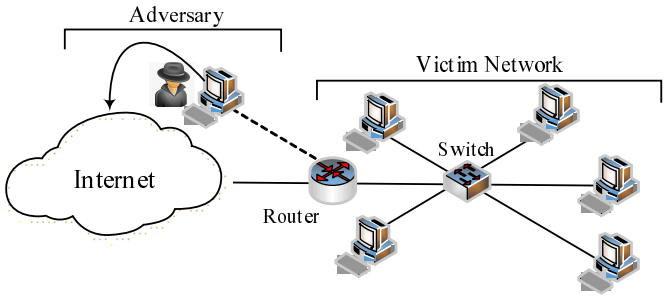

Figure 7: Experiment setup for Malware-free intrusion

The attacking agent defined in the attack model runs on a Linux machine in the first network while the victim hosts running different versions of the Windows operating system reside in the second network. Since Network Level Authentication (NLA) affects the establishment of

RDP session via remote access, we switch it on and off by manipulating the corresponding registry keys. Our experiment setup is shown in Figure 7 above.

We implant the backdoor by file switching method as per α definition and registry debugger configuration as per β definition. We further activate Terminal Services, the service responsible for RDP-based remote access, by changing the hexadecimal value of *fDenyTSConnections* in the registry from 1 to 0 in the registry path $HKEY\_LOCAL\_MACHINE\backslash SYSTEM\backslash CurrentControlSet\backslash Control\backslash TerminalServer$. We likewise switch on and off NLA for different combinations of α, β, γ for registry path HKEY_LOCAL_MACHINE \SYSTEM \CurrentControlSet\Control \Terminal Server\WinStations \RDP-Tcp by changing the key hexadecimal value of UserAuthentication from 0 to 1. We test the attack with different combinations of these parameters and the results are shown in Table 2 below.

Table 2: Attack status with different parameters

S_n	FileSw	RegDebug	RDP Reg	Status
S_{001}	X	X	OFF	*NIL*
S_{001}	0	0	ON	*NIL*
S_{010}	X	X	OFF	*NIL*
S_{010}	1	0	ON	*SUCCESS*
S_{010}	0	1	ON	*SUCCESS*
S_{011}	1	1	ON	*SUCCESS*

It is evident from Table 2 that the core of the success of this malware-free intrusion attack vector is RDP. When RDP is *OFF*, the combination of the rest of the attack parameters is irrelevant as the overall attack does not materialize. The attack in S_{001} does not materialize despite having a state where RDP is ON because the other two parameters which actualize the backdoor are OFF. The "do not care "values are represented by the denotation X to denote that, whether the value is *1* or *0* is actually insignificant as it bears no effect on the end result.

We further observed that with RDP on, only one of the attack vectors of backdoor implantation is required as evidenced by the XOR operation of state S_{010} in Table 2. Furthermore, the presence of two implanted backdoors by the defined attack vectors in the presence of activated RDP adds no difference in that the attack will materialize as if only one backdoor was implanted. However, considering that this is a malware-free intrusion, an attacker might choose to activate both backdoors to increase the level of persistence in the event that one backdoor is detected. Now, having observed that only three states out of the eight actually represent the status of the system where the attack is successful, we finally deduce the overall attack chain for the malware-free intrusion using our defined attack vectors. The attack chain is depicted below in Figure 8.

The chain comprises four steps of which the first and

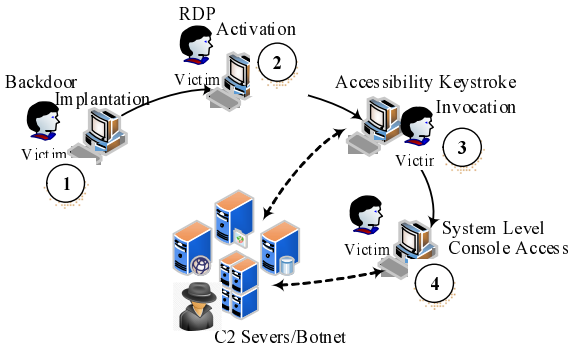

Figure 8: Malware-free intrusion attack chain

second step could come in any order. The attacker can be initiating the attack from Command and Control (C2) servers or botnets and since this backdoor can be persistent, the attacker can include the victim to their list of C2 servers or to their botnet for further compromise. The accessibility keystroke invocation over the network in the third step can only work when the first two steps are completed. Likewise, system level access over RDP-based remote access is only feasible after materialization of the first three steps. Skipping any of these steps thwarts the attack.

It is worth noting that NLA [10], a security feature introduced into later versions of the Windows operating system to thwart Denial of Service (DOS) attacks over an RDP session somewhat inadvertently helps mitigate backdoor attacks discussed herein. Activating NLA after switching on RDP requires that the connecting user avail their authentication credentials before a session is established. We did not include activation of NLA as an attack vector because RDP-based remote access functions without this feature. NLA, however, has its own implementation challenges as later elaborated in this paper.

4 Bayesian and Security Assessment Model

We engage the services of Bayesian network statistics to the attack vectors defined in our attack model and attack chain in the preceding sections. Attack vectors and paths have been employed in the study of vulnerabilities of various information systems [11, 19] and further to develop probabilistic metrics of enterprise networks [6, 21]. We likewise extend and employ this technique to our assessment model via the construction of a directed graph with specific nodes and corresponding edges. We apply the aforementioned attack vectors to the paths and the resultant is a directed graph of an infiltration Bayesian network shown in Figure 9 below.

Our infiltration Bayesian network is presented as a cascaded hierarchical graph of three levels where the topmost level denotes the entry point of the network. The set of

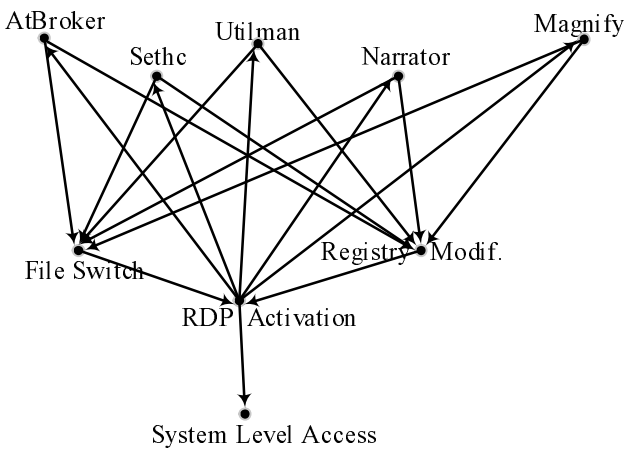

Figure 9: Infiltration Bayesian network

first entry points are attacks on the accessibility suite executable binaries which when switched (second level) with an executable binary (e.g. cmd.exe) capable of availing system level access or where such a file is set as the debugger (second level) for specified accessibility suite executable binary in the registry results into implantation of the malware-free backdoor. We differentiate five accessibility suite binaries namely Atbroker.exe, Sethc.exe, Utilman.exe, Narrator.exe and Magnify.exe which reside in the %systemroot%\system32 directory. The second entry point of the infiltration Bayesian network is RDP activation though the order of these entry points at this level is of no significance as observed in the previous section. The third and final level is system level access after successful traversal of the appropriate paths.

We thus denote the infiltration Bayesian network (BN) with the following parameters:

$$BN^{inf} = (G^{inf}, \Omega^{inf})$$

where G^{inf} is a directed graph containing nodes representative of random variables and the links between the nodes denoting direct dependence relationship, and where Ω^{inf} denotes the set of quantitative parameters ω_i of the given network for $i = 1,...,n$. We represent the sequence of random variables of the given nodes as:

$$S_r^{inf} = \{B_i\}_{i=1}^{n}$$

where the random binary variable B_i dictates if a pursued node $i \in \{1, ..., n\}$ has been accessed through the corresponding attack vector. We define the Bayesian probability b_{ij} as the probability of accessing node i given that node j is accessed by the pursued attack vector:

$$b_{ij} = Pr(B_i|B_j)$$

Therefore, the link set L^{inf}, denoting a collection of nodes through which an attack vector can be pursued with positive probability is defined as:

$$L^{inf} = \{(i,j) : a_{ij} > 0,\ i,j \in \{1,...n\}, i \neq j\}$$

We further define the Bayesian probability of the network parameter ω_i as:

$$\omega_i = Pr(B_i|\Phi_i)$$

where the set of parent nodes $\Phi_i = \{B_j : b_{ij} > 0\}$ and the network parameter is a subset such that $\omega_i \in \Omega^{inf}$.

We now construct a table, Table 3, consisting different security control profiles which can help address the vulnerabilities revealed in our models. We consider four distinct systems with seven security controls partitioned progressively starting with a minimal set of security controls. Enhanced security controls are shaded green while average controls are shaded orange and the least controls not shaded at all.

Security controls in the first system configuration, System 1, are at their minimum. There are no additional controls integrated in the system and those controls that come by default with the setting are left unaltered. This is the least desirable state which would otherwise correspond to state S_{011} of Figure 1 and subsequently Equation (5). Since the attack is a malware-free intrusion attack, there is no high expectation of detection by the IDS that be.

The security controls of the second system configuration, System 2, introduce two security features. These controls are classified as average because file integrity check in the %systemroot%\system32 directory is not complete. This implies the checking mechanism might not be able to detect a backdoor implanted by the unchecked accessibility executable binary. This might be the case of an updated system in which the update introduces a new the set of accessibility features not captured by the integrity check before the update. The registry modification detection will depend on the set of accessibility features present and will likewise in the same manner miss some registry modifications on the same pretext that partial file integrity check fails. This corresponds to state S_{001} of Figure 1 and subsequently Equation (3) where it is possible to pursue either of the two attack vectors to completion but not both.

The set of security controls applied in System 3 include full file integrity check in the %systemroot%\system32 directory and full hash collision detection. This implies that any backdoor implanted by the attack vector of file switch will probably be detected. Furthermore, this security profile includes port obscurity which obscures the RDP port against brute-forcing attacks [24] on the default port 3389 or against the attacks initiated by automated RDP discovery scripts [3].

However, for a targeted attack, the attacker will have to go a step further to probe all ports on the system to overcome this security control. Nonetheless, this security profile lacks NLA which otherwise prevents establishment of an RDP session at pre-authentication. Likewise, it does not employ FDE with key preservation.

Table 3: System security control profiles

System	File Integrity Check	Hash Collision Check	Reg. Modification Detection	NLA	Port Obscurity	FDE	Pre-Authentication Accessibility Features
System 1	Absent	Absent	Not Present	Left OFF	Default RDP	No	Default Settings
System 2	Partial	Absent	Not Present	Turned OFF	Default Port	No	Default Settings
System 3	Complete	Complete	Present	Unchecked	Obscured Port	No	Default Settings
System 4	Complete	Complete	Present	Turned ON	Obscured Port	Yes	All turned off

To this effect it is possible to implant the backdoor but only via setting cmd.exe as the debugger for any of the chosen accessibility suite executable binaries. This, in the event that RDP is activated covertly, would correspond to state S_{010} of Figure 1 denoted by Equation (4) implying the pursuance unto completion two attack vectors which actualize the attack.

The last security profile presents hardened security configurations which provide the highest level security counter measures against this malware-free intrusion backdoor. All binary executables in the %systemroot%\system32 directory are hashed unto completion to verify file integrity and a hash detection computed to determine any indicators of compromise in the event of a hash collision. This entails that the malware-free backdoor cannot be implanted via this attack vector. Furthermore, registry modification detection are carried out for both backdoor implantation of setting a debugger to an accessibility executable binary and also detection for covert activation of RDP through the registry as elaborated in Section 3. In addition to port obscurity, this security profile also enforces the usage of NLA implying that even in an extenuating scenario where a backdoor were to somewhat slip through slip the aforementioned security controls of this profile, interactive console system level access would not be attained as this would require authentication first before session establishment as per NLA requirement. Such a security profile is reflected by the most secure state S_{000} of Figure 1 where Equation (2) represents such as state.

Since the outcome of the attack depends on the conditional probability that the present attack vector can only be pursued unto completion, if the other related attack vector has successfully been actualized, we can employ conditional probabilities of Bayesian inference to evaluate our assessment model. The conditional probabilities are denoted by directed edges in the Bayesian network in Figure 9. So we apply the security controls in Table 3 to these edges in the formulation of our assessment model.

Since Table 3 presents security controls capable of detecting the backdoor, we define a binary random variable Ψ_i^{inf} for $i = 1, ..., n$ to denote detection of infiltration of the corresponding *i-th* node. We therefore compute the probability of undetected infiltration as:

$$Pr(B) = 1 - \prod_{i=1}^{n}[1 - Pr(B|B_i) \cdot Pr(B_i) \cdot [1 - Pr(\Psi_i^{inf})]]$$

and we evaluate the probability of access of the *i-th* node as:

$$Pr(B_i) = 1 - \prod_{j=1}^{n}[1 - Pr(B_i|B_j) \cdot Pr(B_j)]$$

for i = *1,...,n*, we have:

$$Pr(B_i) = 1 - \prod_{j=1}^{n}[1 - b_{ij} \cdot Pr(B_j)]$$

considering that $b_{ij} = Pr(B_i|B_j)$. If the probability of not detecting an attack via a given node is defined as $Pr(B|absence\,of\,detection)$, that is to mean $Pr(\Psi_i^{inf}) = 0$, then we estimate the probability that the implemented security controls can detect the attack as:

$$Pr(\Psi^{inf}) = \frac{Pr(B|absence\,of\,detection) - Pr(B)}{Pr(B)}$$

We assumed the Markov property [7] in our formulations that access to the *i-th* node depends only on its parents and not on the history of the subsequent nodes thereof.

Application of the model requires computation of conditional probabilities for all edges in our Bayesian network. We compute the probability scores of our network using Conditional Probability Tables (CPT), representative of the strength on an influence, as shown in Table 4 below. The CPT likewise represents the probability distribution of possible states of a node of which the preconditions are based on its parents' states.

Table 4: CPT for $Node_i$

Φ_1	Φ_2	Φ_3	Φ_4	*Node i*
T/F	*T/F*	*T/F*	*T/F*	$Pr(B_i\|B_j)$
Otherwise				$1 - Pr(B_i\|B_j)$

Since our malware-free intrusion attack structure is not directly based on software vulnerability exploit, we esti-

mate the probability of success directly in the corresponding knowledge base. Therefore, to derive CPT parameters, we employ the use of discrete levels homologous to those reflected in the Common Vulnerability Scoring System (CVSS) metrics [9, 23]. In view of the aforementioned, we thus assign the following values to the conditional probability: very high – assigned a value of 1, *high* – assigned a value of 0.9, *medium* – assigned a value of 0.5, *low* – assigned a value of 0.1 and *very low* – assigned a value of 0.01. The Table 5 below shows some selected probability scores.

Table 5: Selected probability scores

Attack	System 1	System 2	System 3	System 4
Initial access	1	1	1	1
Reg. modify	1	0.9	0.9	0.01
Sethc switch	0.9	0.5	0.1	0.1
RDP ON	0.9	0.5	0.5	0.01

We assume that access to the entry node is used as given implying that for initial system access: $Pr(B_1) = 1$. This is so because the spectrum of this access is so wide that it most likely encompasses attack vectors that might employ malware like exploit kits which is in conflict with the approach considered in this paper.

Having evaluated the probability scores, we now present security profile assessment results of the security profiles from Table 3 in the following Table 6 below. The probabilities are presented for System Level Access (SLA) with different detection parameters.

Table 6: Security profile assessment results

Security Profile	SLA with detection	SLA minus detection	SLA detection Prob.
System 1	0.88	0.99	0.1
System 2	0.85	0.97	0.12
System 3	0.26	0.49	0.47
System 4	0.002	0.005	0.5

The highest level of infiltration is echoed in system profile 1 and 2 with probability averaging 98%. This is explained by the lack of robust security controls as depicted in Table 3. On the other hand, if the attack is detected before actual SLA, the probability is reduced as evidenced in the third profile with relatively better security controls as opposed to the first and second profile. The fourth profile, with the most hardened security controls, has the highest detection probability and the lowest infiltration cases. This profile extensively employs rigorous integrity checks both in the SystemRoot and Registry whilst counter-checking any indicators of compromise via hash collisions. These security controls mitigate the known attack vectors through which the accessibility backdoor is implanted and subsequently accessed with SLA.

5 Conclusion

We have demonstrated how a security assessment model based on Boolean Logic for security state formulation and Bayesian inference for probability evaluation can be used for the assessment of the security environment of a specified scenario of malware-free intrusion. The number of attack vectors as variables of Boolean functions have a direct influence on the attack paths and the infiltration thereof generated through BN. The assessment model gives insight into the significance of the various security controls meant to counter attacks via malware-free intrusions. Since the characteristics of a malware-free intrusion differs from those IDSs are accustomed to, countering attacks via these attack vectors calls for tailor made solutions which might otherwise not come with the common security products.

Compared to other works on network security assessment based on Bayesian models [1, 20], our work introduces the use of Boolean state machines for precise representation of security states of affected systems. Furthermore, our work encompasses malware-free intrusions which have not been explicitly inferred in Bayesian networks and state machines before.

Inasmuch as the collaboration of Boolean Logic and Bayesian inference sheds more light on the constituents of a malware-free attack, the approach also faces challenges in that there are some components of the Bayesian network which cast a considerable level uncertainty difficult enough to be captured by Boolean Logic reasoning. So regardless of the robustness of the implemented security measures against attacks in this respect, it is not as straightforward to postulate and extrapolate for certain that the attack will not materialize as a binary response.

References

[1] F. X. Aguessy, O. Bettan, G. Blanc, V. Conan, and H. Debar, "Hybrid risk assessment model based on bayesian networks," in *International Workshop on Security*, pp. 21–40, Springer, 2016.

[2] I. Arce and G. Richarte, "State of the art security from an attacker's viewpoint," in *PacSec Conference*, Tokyo, Japan, 2003.

[3] E. Beqiri and D. Campus, *Information and Communication Technology Security Issues*, University of East London, 2010.

[4] M. K. Daly, "Advanced persistent threat," *Usenix*, vol. 4, no. 4, pp. 2013–2016, 2009.

[5] G. Dong, H. Hao, R. Du, and L. Tian, "Attacking mode based on shell structure of complex networks,"

International Journal of Nonlinear Science, vol. 20, no. 3, pp. 148–153, 2015.

[6] M. Frigault, L. Wang, A. Singhal, and S. Jajodia, "Measuring network security using dynamic bayesian network," in *Proceedings of the 4th ACM Workshop on Quality of Protection*, pp. 23–30, 2008.

[7] M. Frydenberg, "The chain graph markov property," *Scandinavian Journal of Statistics*, vol. 17, no. 4, pp. 333–353, 1990.

[8] A. Futoransky, L. Notarfrancesco, G. Richarte, and C. Sarraute, "Building computer network attacks," *arXiv preprint arXiv:1006.1916*, 2010.

[9] L. Glanz, S. Schmidt, S. Wollny, and B. Hermann, "A vulnerability's lifetime: enhancing version information in cve databases," in *Proceedings of the 15th ACM International Conference on Knowledge Technologies and Data-driven Business*, pp. 28, 2015.

[10] K. H. Ho, *Remote Desktop Services in Windows 2008 R2*, 2008. (`http://sharepointgeorge.com/2009/remote-desktop-services-windows-2008-r2-part-1/`)

[11] M. Khosravi-Farmad, R. Rezaee, A. Harati, and A. G. Bafghi, "Network security risk mitigation using bayesian decision networks," in *4th IEEE International eConference on Computer and Knowledge Engineering (ICCKE'14)*, pp. 267–272, 2014.

[12] M. Z. Mas' ud, S. Sahib, M. F. Abdollah, S. R. Selamat, and C. Y. Huoy, "A comparative study on feature selection method for n-gram mobile malware detection," *International Journal of Network Security*, vol. 19, no. 5, pp. 727–733, 2017.

[13] S. Mauw and M. Oostdijk, "Foundations of attack trees," in *International Conference on Information Security and Cryptology (ICISC'05)*, pp. 186–198, 2005.

[14] J. Moon, D. Lee, J. Jung, and D. Won, "Improvement of efficient and secure smart card based password authentication scheme.," *International Journal of Network Security*, vol. 19, no. 6, pp. 1053–1061, 2017.

[15] L. Ray and H. Felch, "Detecting advanced persistent threats in oracle databases: Methods and techniques," in *Strategic Information Systems and Technologies in Modern Organizations*, pp. 71–89, 2017.

[16] K. Rieck, P. Trinius, C. Willems, and T. Holz, "Automatic analysis of malware behavior using machine learning," *Journal of Computer Security*, vol. 19, no. 4, pp. 639–668, 2011.

[17] C. Sarraute, G. Richarte, and J. Lucángeli Obes, "An algorithm to find optimal attack paths in nondeterministic scenarios," in *Proceedings of the 4th ACM Workshop on Security and Artificial Intelligence*, pp. 71–80, 2011.

[18] S. Saxena, "Demystifying malware traffic," *SANS Institute InfoSec*, 2016.

[19] V. Shandilya, C. B. Simmons, and S. Shiva, "Use of attack graphs in security systems," *Journal of Computer Networks and Communications*, vol. 2014, 2014.

[20] J. Shin, H. Son, G. Heo, *et al.*, "Development of a cyber security risk model using bayesian networks," *Reliability Engineering & System Safety*, vol. 134, pp. 208–217, 2015.

[21] A. Singhal and X. Ou, *Security Risk Analysis of Enterprise Networks Using Probabilistic Attack Graphs*, US Department of Commerce, National Institute of Standards and Technology, 2011.

[22] X. Wang, K. Zheng, X. Niu, B. Wu, and C. Wu, "Detection of command and control in advanced persistent threat based on independent access," in *IEEE International Conference on Communications (ICC'16)*, pp. 1–6, 2016.

[23] A. Younis, Y. K. Malaiya, and I. Ray, "Assessing vulnerability exploitability risk using software properties," *Software Quality Journal*, vol. 24, no. 1, p. 159, 2016.

[24] A. Zimba, "Malware-free intrusion: A novel approach to ransomware infection vectors," *International Journal of Computer Science and Information Security*, vol. 15, no. 2, pp. 317, 2017.

Biography

Aaron Zimba received his Master and Bachelor of Science degree from the St Petersburg Electrotechnical University in St Petersburg in 2009 and 2007 respectively. He is currently a PhD student at the University of Science and Technology Beijing in the Department of Computer Science and Technology. He is also a member of the IEEE. His main research interests include Network Security Models, Network and Information Security, Cryptovirology and Cloud Computing Security .

Hongsong Chen received his Ph.D degree in Department of Computer Science from Harbin Institute of Technology, China, in 2006. He was a visiting scholar in Purdue University from 2013-2014. He is currently an associate professor in Department of Computer Science, University of Science and Technology Beijing, China. His current research interests include wireless network security, attack and detection models, and cloud computing security.

Zhaoshun Wang is a Professor and the Associate Head of the Department of Computer Science and Technology at the University of Science and Technology Beijing. He graduated from Department of Mathematics at Beijing Normal University in 1993. He received his PhD from Beijing University of Science and Technology in 2002. He completed postdoctoral research work at the Graduate School of the Chinese Academy of Sciences in 2006. He holds patents and has many awards to his name. His main research areas include Information Security, Computer Architecture and Software Engineering.

Constructing Provably Secure ID-based Beta Cryptographic Scheme in Random Oracle

Chandrashekhar Meshram[1], Sarita Gajbhiye Meshram[1] and Cheng-Chi Lee[2,3]
(Corresponding author: Cheng-Chi Lee)

Department of Mathematics and Computer Science, Rani Durgawati University[1]
Saraswati Vihar, Pachpedi, Jabalpur 482001, India
Department of Library and Information Science,Fu Jen Catholic University[2]
510 Jhongjheng Road, Taipei 24205, Taiwan, R.O.C.
Department of Photonics and Communication Engineering,Asia University[3]
Wufeng Shiang, Taichung, Taiwan 413, R.O.C.
(Email: cs_meshram@rediffmail.com,cclee@blue.lins.fju.edu.tw)
(Received Jan. 24, 2017; revised and accepted June 5, 2017)

Abstract

In this study, we propose a new ID-based beta cryptosystem scheme secure under selective identity adaptive chosen ciphertext security (IND-sID-CCA) under assumption in the random oracle model. We demonstrate that our scheme outperforms the other existing schemes in terms of security, computational cost and the length of public key.

Keywords: Discrete Logarithm Problem; Generalized Discrete Logarithm Problem; ID-based Cryptosystem; Integer Factorization Problem and Beta Cryptosystem; Public key Cryptosystem

1 Introduction

Shamir [24] introduced the idea of ID-based cryptography to simplify the key management problem in 1984. Two efficient ID-based cryptosystem schemes were proposed by Cocks [7] and Boneh and Franklin [6] in 2001. In their seminal paper [5], Boneh and Franklin used a category of bilinear maps as the basis of their construction. This leads a number of ID-based cryptosystem schemes [2, 3, 26], among others based on bilinear maps. Few ID-based cryptosystem schemes [1, 10, 12] have been proposed after 2003. But in these schemes, the public key of each user is not only an identity, but also some random number selected either by the user or by the trusted authority. But which makes the ID-based cryptosystem an active research field in recent years.

The first efficient ID-based cryptosystem scheme was proposed by Boneh and Franklin [5, 6]. The novel approach they use is based on a class of bilinear maps. Following their work, a number of ID-based cryptosystem scheme using bilinear maps were proposed. For example,Waters [12] presented an efficient and secure ID-based cryptosystem scheme without random oracles; Boneh and Boyen [2] designed a secure ID-based cryptosystem scheme without random oracles; Boneh and Boyen [3] gave another efficient ID-based encryption scheme without random oracles, which is secure in the selective identity model.

Meshram *et al.* [17, 20, 22] presented some new efficient ID-based cryptographic schemes and ID-based mechanisms based on discrete logarithm problem, generalized discrete logarithm problem and integer factorization problem. The security of this schemes are solving the hardness of discrete logarithm problem, generalized discrete logarithm problem and integer factorization problem simultaneously. Meshram and Meshram [18, 19] investigate the new variant of ID-based beta cryptographic scheme and transformation process such as public key cryptographic scheme transfer to ID-based cryptographic scheme without developing new ID-based scheme.

Meshram [14, 15, 16] presented new provably secure ID-based cryptographic scheme, new variant of ID-based beta cryptographic scheme, and efficient scheme based on integer factorization problem and discrete logarithm problem. It is as low as ElGamal scheme. Meshram and Obaidat [21] also showed new variant of ID-based cryptographic scheme such as quadratic-exponentiation randomized cryptographic scheme. Recently, Meshram *et al.* [23] projected new ID-based cryptographic scheme based on partial discrete logarithm problem. Liu and Ye [11] presented new variations as homomorphic universal re-encryptor for ID-based cryptography. In similar manner Wang *et al.* [25] presented efficient ID-based proxy multi signature using cubic residues.

As outlined above, unfortunately we found that new cryptographic model always face security challenges and

confidentiality concerns. Therefore, our main contribution of this paper is to fill this gap by proposing a provably secure ID-based beta cryptographic scheme. Specifically, we will show that the security of the proposed scheme is closely, if not tightly, related to difficulty of solving generalized discrete logarithm problem and integer factorization problem. We provided an formal security proof for selective identity adaptive chosen ciphertext security (IND-sID-CCA) in the random oracle, which means that the new scheme offers better security guarantees than existing other ID-based cryptographic schemes. The proposed scheme does not use pairings (bilinear maps), resulting in high efficiency and ease of implementation, neither does it rely on the relatively new and untested hardness assumptions related to pairing-based cryptography. This makes it attractive for application in resource-constrained environments where saving in computation, communication and implementation code area is a premium.

The rest of this paper is organized as follows: The Beta cryptosystem and supporting example for it are demonstrated in Section 2 and 3 respectively. Proposed an ID-based beta cryptographic scheme with chosen ciphertext security is demonstrate in Section 4. The security examination of proposed ID-based cryptographic scheme is presented in Section 5. Discuss comparison with previous ID-based cryptographic schemes in Section 6. Finally, Section 7 concludes the paper.

2 The Beta Cryptosystem

The algorithm consists of three sub-algorithm, Key generation, Encryption and Decryption.

2.1 Key Generation

The key generation algorithm runs as follows (user 1 should do the following).

1) Select arbitrary primes q and p each roughly of the same size.

2) Calculates $N = q \star p$ and Euler-phi function $\varphi(N) = (q-1)(p-1)$.

3) Choose an arbitrary integer $e, 1 \leq e \leq \varphi(N)$ such that gcd $(e, \varphi(N)) = 1$.

4) Choose an arbitrary integer b such that $2 \leq b \leq \varphi(N) - 1$.

5) Choose an element β of the multiplicative group $\mathbb{Z}_N^*$ and calculate $y_1 = \beta^b mod(N)$.

6) By using the extended Euclidean algorithm to calculate the unique integer $d, 1 \leq d \leq \varphi(N)$ such that $ed \equiv 1(mod \varphi(N))$.

The public key is formed by (N, e, β^b) and the corresponding private key is given by (d, b, β).

2.2 Encryption

An user 2 to encrypt a message m to user 1 should do the following:

1) The message is represented as an integer in the interval $[1, N-1]$.

2) The cipher text is given by $C = (m\beta^b)^e mod(N)$.

2.3 Decryption

To recover the plaintext m from the cipher text C, user 1 should do the following:

1) Calculate $y_2 = \beta^{\varphi(N)-b} mod(N) = \beta^{-b} mod(N)$.

2) Then calculate $y_3 = (y_2)^e mod(N)$.

3) Recover the plaintext m by computing $((y_2)^e \star C)^d (mod N)$.

3 Example

To make our construction easy to comprehend, we illustrate an example to show the basic principle of our proposed scheme.

Let the two primes be $q = 29$ and $p = 43$ and set $N = 1247$ and $\varphi(N) = 1176$.

3.1 Key Generation

The key generation algorithm runs as follows.

1) Select an arbitrary integer $e = 11$ and gcd $(11, 1176) = 1$.

2) Select an arbitrary integer $b = 19$.

3) Choose an element $\beta = 10$ of the multiplicative group $\mathbb{Z}_N^*$ and calculate $y_1 = \beta^b mod(N) = (10)^{19} mod\ 1247 = 427$.

4) By using the extended Euclidean algorithm to compute the unique integer $d = 107, 1 \leq d \leq \varphi(N)$ such that $11d \equiv 1(mod\ 1176)$.

The public key is formed by (N, e, β^b) and the corresponding private key is given by (d, b, β).

3.2 Encryption

An user 2 to encrypt a message m to user 1 should do the following:

1) The message $m = 1122$ is represented as an integer in the interval $[1, N-1]$.

2) The cipher text is given by $C = (m\beta^b)^e mod(N) = (479094)^{11} mod\ 1247 = 791$.

3.3 Decryption

To recover the plaintext m from the cipher text C, user 1 should do the following:

1) Calculate $y_2 = \beta^{\varphi(N)-b} mod(N) = \beta^{-b} mod(N) = 917$.
2) Then calculate $y_3 = (y_2)^e mod(N) = 483$.
3) Recover the plaintext m by computing $((y_2)^e \star C)^d (modN) = 1122$.

4 An ID-based Beta Cryptosystem Scheme with Chosen Ciphertext Security

The major contribution of our proposed ID-based beta cryptosystem is the key generation phase. Upon the successful creation of a private key, the scheme concept can be easily implemented in encryption and decryption posses.

4.1 Setup

By taking in security parameter t this algorithm will be carried out by PKG as follows:

1) Let $N = q * p$ be a large prime number,such that $\varphi(N) = (q-1)(p-1)$ and β be an element of order N in $Z_N^\star$, x, y be PKG's secret and public keys respectively, where $y = \beta^x mod\ N$.
2) Select two random integers e and d as $1 \leq e, d \leq \varphi(N)$, such that $gcd(e, \varphi(N)) = 1$ and $ed \equiv 1(mod \varphi(N))$.
3) The PKG chosen randomly secrete information as k_i for $(1 \leq i \leq t)$, where $\Sigma_{i=1}^{t} k_i < \varphi(N)$ and public information K_i, where $K_i = \beta^{k_i} mod\ N$,for $(1 \leq i \leq t)$.
4) Compute the hash function $H : \{0,1\}^t \rightarrow Z_N^*$.

4.2 Exact

For a given user identity $ID \in \{0,1\}^*$, we compute the private key of the user is $\beta^{\theta_A} = vK_A^{K_A} mod\ N$, where $\theta_A = \Sigma_{i=1}^{t} k_i v_{Ai} mod \varphi(N)$, $K_A = \Pi_{i=1}^{t} K_i^{v_{Ai}} mod\ N$ and v_{Ai} is the i^{th} bit of $H(ID_A)$ for $(1 \leq i \leq t)$.

4.3 Encryption

To encrypt a message $M \in \{0,1\}^*$ for ID as follows:

1) Set the public key $V_A = \beta^{\theta_A} = vK_A^{K_A} mod\ N$, where $K_A = \Pi_{i=1}^{t} K_i^{v_{Ai}} mod\ N$.
2) Chosen a random integer e such that $gcd(e, \varphi(N)) = 1$.
3) Compute the ciphertext to be $C = (M\beta^{\theta_A})^e (mod\ N)$.

4.4 Decryption

Let C be the valid ciphertext encrypted by using the public key V_A. The user can decrypt ciphertext using the private key θ_A.

1) Calculate $y_2 = \beta^{-\theta_A} (mod\ N)$.
2) Compute $y_2^e = (\beta^{-\theta_A})^e (mod\ N)$.
3) Out put M as the decryption of C as

$$[(y_2)^e * C]^d (mod\ N) = [\beta^{-\theta_A e} M^e \beta^{\theta_A e}]^d (mod\ N)$$
$$= M^{ed} (mod\ N) = M (mod\ N).$$

5 Security Examination

In this section, we examine the security of ID-based beta cryptosystem scheme. The following theorem shows that ID-based beta cryptosystem scheme is IND-sID-CCA secure, if beta cryptosystem is IND-CCA secure [8] in random oracle model [9].

Definition 1. *An ID-based cryptosystem scheme, E is said to be selective identity, adaptively chosen ciphertext secure (IND-sID-CCA), if no probabilistic polynomial time (PPT) adversary* A *has a non-negligible advantage in the following game in [4].*

Theorem 1. *The identity hash function H be a random oracle. Then ID-based beta cryptosystem scheme is IND-sID-CCA secure, if beta cryptosystem [Section 2] is IND-CCA secure. Concretely, suppose there is an IND-sID-CCA rival R_1 that has advantage $\epsilon(k)$ against ID-based beta cryptosystem. Then there exists an IND-CCA rival R_2 with advantage at least $\epsilon(k)$ against beta cryptosystem. Its running time is rival $O(time(R_1))$.*

Proof. The main idea of this proof is to construct an IND-CCA rival R_2 to gain the advantage against beta cryptosystem in the following IND-CCA game.

At the starting of the game, the IND-CCA challenger generates the public key $K_{pub} = \langle N, \beta, v \rangle$ and a private key x that satisfies $v = \beta^x mod\ N$. The challenger gives K_{pub} to rival R_2, then rival R_2 mounts an IND-CCA attack using the help of algorithm rival R_1 as follows:

Initialization. The rival outputs an identity ID_{ch} which it wishes to be challenged.

Setup. The challenger runs the *setup algorithm*. It gives the rival the resulting system parameters. It keeps the masterkey to itself.

H-queries. To respond to H-query, R_2 maintains a list of tuples $\langle ID_{Ai}, V_{Ai}, \theta_{Ai} \rangle$ which we refer to as H^{list}. The list is initially empty. When R_1 queries H at a point ID_{Ai}, R_2 responds as follows:

1) If the query on ID_{Ai} already appears on the H^{list} in a tuple of the form $\langle ID_{Ai}, V_{Ai}, \theta_{Ai} \rangle$ then R_2 responds with $H(ID_{Ai}) = v_{Ai}$ as a answer.

2) If the query is new to the H oracle, R_2 will pick a random $\theta_{Ai} \in Z_N^*$ and computes $V_{Ai} = \beta^{\theta_{Ai}} mod\ N$, else R_2 sets $\theta_{Ai} = *$ and $V_{Ai} = v$. as a answer. Here $*$ denotes a special symbol.

3) R_2 adds the tuple $\langle ID_{Ai}, V_{Ai}, \theta_{Ai}\rangle$ to H^{list} and gives back V_{Ai} to R_1.

Phase 1-Extraction queries. When R_1 asks for the private key associated to ID_{Ai}, R_2 runs the above algorithm and gets $H(ID_{Ai}) = v_{Ai}$, where $\langle ID_{Ai}, V_{Ai}, \theta_{Ai}\rangle$ is the corresponding entry in H^{list}. As $V_{Ai} = \beta^{\theta_{Ai}} mod\ N$, R_2 can retrieve the legitimate private key θ_{Ai} for ID_{Ai}. The extraction query on ID_{ch} will be denied.

Phase 1-Decryption queries. Let $\langle ID_{Ai}, C_i\rangle$ be a decryption query issued by adversary R_1, where C is ciphertext of beta cryptosystem. R_2 responds to the query as follows:

1) If $\langle ID_{Ai} \neq ID_{ch}\rangle$, then R_2 runs H-query algorithm such that of the form $\langle ID_{Ai}, V_{Ai}, \theta_{Ai}\rangle$ be the corresponding tuple on H^{list}. Next it uses the private key θ_{Ai} to respond to the decryption query.

2) If $\langle ID_{Ai} = ID_{ch}\rangle$, then R_2 forwards the decryption query with $\langle C_i\rangle$ and then relays the challenger's response back to R_1.

Challenge. Once R_1 decides that Phase 1 is over it outputs two messages M_0 and M_1 which it wishes to be challenged on. Algorithm R_2 responds as follows:

1) R_2 gives the challenger M_0 and M_1 as the messages that it wishes to be challenged on. The challenger responds with the beta cryptosystem's ciphertext C such that C is the encryption of M_c for a random coin $c \in \{0, 1\}$.

2) Next, R_2 runs the algorithm for responding H-queries to obtain $v \in Z_N^*$ such that $H(ID_{ch}) = v$ and forwards C to R_1.

Phase 2-Extraction queries. R_2 responds the same as in Phase 1, except for the extraction query on ID_{ch}, which will be rejected.

Phase 2-Decryption queries. R_2 responds the same as in Phase 1 except the decryption query $\langle ID_{ch}, C\rangle$ will be denied.

Guess. Rival R_1 finally outputs a guess c' for c. Rival R_2 outputs c' as its guess for c.

The responses to H-queries are as in the factual attack since each response is uniformly and independent distributed in Z_N^*. All responses to private key extraction queries and decryption queries are valid. So R_2 will not abort during the simulation, the possibility of perfect simulation is 1. From these we can conclude that Rival R_1's view is identical to its view in the factual attack. By the definition of algorithm R_1 we have that $|Pr[c = c'] - 1/2| \geq \epsilon(k)$, thereby R_2 has at least advantage $\epsilon(k)$ against beta cryptosystem. This proves theorem 5.0.2 and terminates the proof.

□

6 Performance Comparison of Other ID-based Cryptographic Schemes

In this section, we have discussed four most wide-used ID-based encryption schemes and compared their performance. These four ID-based cryptographic schemes are: Selective-ID Secure ID-based cryptosystem without Random Oracles [3], Boneh-Franklin ID-based cryptosystem [6], Cocks ID-based cryptosystem [7], Authenticated ID-based cryptosystem [13], and our proposed ID-based beta cryptosystem. These schemes have different performance on server for evaluating Encrypt algorithm performance, decryption algorithm performance, and computational cost. Notations used in this computation are as follows: P = airing operation, M = Modular multiplication, e = Exponentiation in G, m =Scalar or Point Multiplication in G, x = XOR operation, h = Hashing, a = addition modulo, i = inverses modulo , J = Jacobi symbol and $C(\gamma)$ = Computation cost of operation γ.

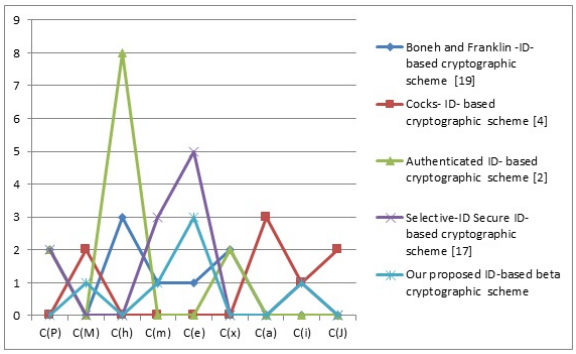

Figure 1: Computational cost

Based on our observation of Figure 1, we have observed that proposed ID-based beta cryptosystem has a better performance than other four schemes [3, 6, 7, 13] in encryption and decryption algorithms. Our proposed scheme is faster than schemes [3, 6, 7, 13] in two aspects. First, our proposed scheme needs no pairing computation in encrypt algorithm and decryption algorithm, because $\tilde{e}(P_1, P_2)$ can be pre-computed. Secondly, in the operation of mapping an identity to an element in G_1 or G_2 , the map-to-point algorithm used by scheme [6] and scheme [3] is not required because simple hash function is used in our scheme to map an identifier to an element in Z_N^*. Our proposed scheme is faster than scheme [7] in one aspect. The size of ciphertext is very large and

consists of two elements of Z_N per bit of the message but the size of our proposed scheme is smaller than scheme [7] and consists of an element of Z_N^* per bit of the message.

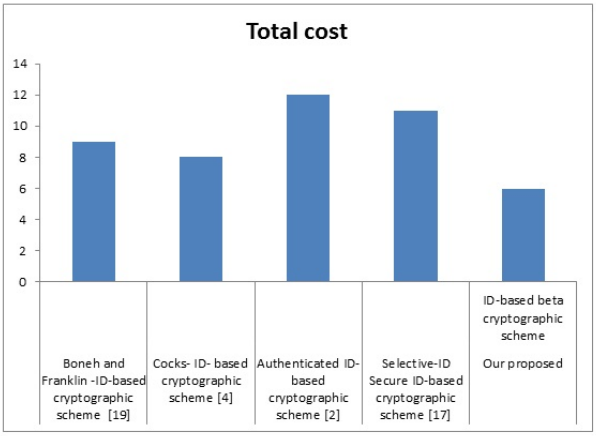

Figure 2: Total computational cost

Also, we evaluated the total computational cost of the four schemes [3, 6, 7, 13] and proposed scheme in Figure 2. We found that the computation cost of scheme [3] is near about the scheme [13] and the computation cost of scheme [7] is near about the scheme [6] and the computation cost of our proposed is much less to other four schemes and half of scheme [13]. As we know that in the Extract algorithm of scheme [6] and scheme [13], an identity string is mapped to a point on an elliptic curve and the corresponding private key is computed by multiplying the mapped point with the master key of public key generator (PKG) and Extract algorithm of our proposed scheme requires much simpler hashing than the schemes [6, 7, 13]. Hence the computational cost will reduce and therefore improves performance.

7 Conclusion

In this article, we deals with new mechanisms for ID-based beta cryptographic scheme, whose unforgeability can be reduced to the hardness of the generalized discrete logarithm problem and integer factorization problem over multiplicative group, which are a fundamental intractable problems in cryptography. It is selective identity adaptive chosen ciphertext security (IND-sID-CCA) under assumption of generalized discrete logarithm problem and integer factorization problem over multiplicative group in random oracle. This scheme is fast than Boneh and Franklin-ID-based cryptographic scheme, Cocks- ID-based cryptographic scheme, Authenticated ID-based cryptographic scheme, Selective-ID Secure ID-based cryptographic scheme and having very low computational cost. Therefore, our new scheme is more practical and has the same security as the original discrete logarithm problem and integer factorization problem-based system.

References

[1] M. Bellare, C. Namprempre and G. Neven, "Security proofs for identity-based identification and signature scheme," *Journal of Cryptology*, vol. 22, pp. 1-61, 2009.

[2] D. Boneh and X. Boyen, "Secure identity based encryption without random oracles," *Advances in Cryptology (CRYPTO'04)*, LNCS 3152, Springer-Verlag, pp. 443-459, 2004.

[3] D. Boneh and X. Boyen, "Efficient selective-id secure identity based encryption without random Oracles," *Advances in Cryptology (Eurocrypt'04)*, LNCS 3027, Springer-Verlag, pp. 223-238, 2004.

[4] D. Boneh, R. Canetti, S. Halevi and J. Katz, "Chosen-ciphertext security from identity-based Encryption," *SIAM Journal on Computing*, vol. 36, no. 5, pp. 1301-1328, 2006.

[5] D. Boneh and M. K. Franklin, "Identity based encryption from the Weil pairing," *SIAM Journal on Computing*, vol. 32, no. 3, pp. 586-615, 2003.

[6] D. Boneh and M. K. Franklin, "Identity based encryption from the Weil pairing," *Advances in Cryptology (CRYPTO'01)*, LNCS 2193, Springer-Verlag, pp. 213-229, 2001.

[7] C. Cocks, "An identity based wncryption scheme based on quadratic residues," in *International Conference on Cryptography and Coding*, LNCS 2260, Springer-Verlag, pp. 360-363, 2001.

[8] M. Hassouna, B. Barry and E. Bashier, "A new level 3 trust hierarchal certificateless public key cryptography scheme in the random oracle model," *International Journal of Network Security*, vol. 19, pp. 551-558, 2017.

[9] J. Kar, "Provably secure identity-based aggregate signcryption scheme in random oracles," *International Journal of Network Security*, vol. 17, no. 5, pp. 580-587, 2015.

[10] E. Kiltz and Y. Vahlis, "CCA2 secure IBE: Standard model efficiency through authenticated symmetric encryption," in *CT-RSA*, LNCS 4964, Springer-Verlag, pp. 221-239, 2008.

[11] L. Liu and J. Ye, "A homomorphic universal re-encryptor for identity-based encryption," *International Journal of Network Security*, vol. 19, no. 1, pp. 11-19, 2017.

[12] W. B. Lee and K. C. Liao, "Constructing identity-based cryptosystems for discrete logarithm based cryptosystems," *Journal of Network and Computer Applications*, vol. 22, pp. 191-199, 2004.

[13] B. Lynn, "Authenticated ID-based encryption," *Cryptology ePrint Archive*, Report 2002/072, 2002.

[14] C. Meshram, "An efficient ID-based cryptographic encryption based on discrete logarithm problem and

integer factorization problem," *Information Processing Letters*, vol. 115, no. 2, pp. 351-358, 2015.
[15] C. Meshram, "An efficient ID-based beta cryptosystem," *International Journal of Security and Its Applications*, vol. 9, no. 2, pp. 189-202, 2015.
[16] C. Meshram, "Factoring and discrete logarithm using IBC," *International Journal of Hybrid Information Technology*, vol. 8, no. 3, pp. 121-132, 2015.
[17] C. Meshram, X. Huang and S. Meshram, "New identity-based cryptographic scheme for IFP and DLP based cryptosystem," *International Journal of Pure and Applied Mathematics*, vol. 81, no. 1, pp. 65-79, 2012.
[18] C. Meshram and S. Meshram, "An identity-based cryptographic model for discrete logarithm and integer factoring based cryptosystem," *Information Processing Letters*, vol. 113, no. 10-11, pp. 375-380, 2013.
[19] C. Meshram and S. Meshram, "An identity based beta cryptosystem," in *IEEE Proceedings of 7th International Conference on Information Assurance and Security (IAS'11)*, pp. 298-303, 2011.
[20] C. Meshram, S. Meshram and M. Zhang, "An ID-based cryptographic mechanisms based on GDLP and IFP," *Information Processing Letters*, vol. 112, no. 19, pp. 753-758, 2012.
[21] C. Meshram and M. S. Obaidat, "An ID-based quadratic-exponentiation randomized cryptographic scheme," in *IEEE Proceeding of International Conference on Computer, Information and Telecommunication Systems*, pp. 1-5, 2015.
[22] C. Meshram and P. L. Powar, "An efficient identity-based QER cryptographic scheme," *Complex and Intelligent Systems*, vol. 2, no. 4, pp. 285-291, 2016.
[23] C. Meshram, P. L. Powar, M. S. Obaidat and C. C. Lee, "An IBE technique using partial discrete logarithm," *Procedia Computer Science*, vol. 93, pp. 735-741, 2016.
[24] A. Shamir, "Identity-based cryptosystems and signature schemes," in *Proceedings of CRYPTO'84*, LNCS 196, Springer-Verlag, pp. 47-53, 1984.
[25] F. Wang, C. C. Chang, C. Lin and S. C. Chang, "Secure and efficient identity-based proxy multisignature using cubic residues," *International Journal of Network Security*, vol. 18, no. 1, pp. 90-98, 2016.
[26] B. Waters, "Efficient identity-based encryption without random oracles,," *Advances in Cryptology (CRYPTO'05)*, LNCS 3494, Springer-Verlag, pp. 114-127, 2005.

Biography

Chandrashekhar Meshram received the PhD from R.T.M. Nagpur University, Nagpur (M.S.) India. Presently he is Post-Doctoral Fellow under Dr. D S Kothari postdoctoral fellow New Delhi, India. His research interested in the field of Cryptography and its Application, Statistics, Raga (Music and Statistics), Neural Network , Ad hoc Network, Number theory, Time Series Analysis and Climate Change, Mathematical modeling and Chaos Theory. He is a member of International Association of Engineers (IAENG), Hong Kong, World Academy of Science, Engineering and Technology (WASET), New Zealand, Computer Science Teachers Association (CSTA), USA, Association for Computing Machinery (ACM), USA, International Association of Computer Science and Information Technology (IACSIT), Singapore, European Association for Theoretical Computer Science (EATCS), Greece, International Association of Railway Operations Research (IAROR), Netherland, International Association for Pattern Recognition (IAPR), New York, International Federation for Information Processing (IFIP), Austria, Association for the Advancement of Computing in Education (AACE), USA, International Mathematical Union (IMU) Berlin, Germany, European Alliance for Innovation (EAI), International Linear Algebra Society (ILAS) Haifa, Israel, Science and Engineering Institute (SCIEI), Machine Intelligence Research Labs (MIR Labs) , USA, Society: Intelligent Systems, KES International Association, United Kingdom, Universal Association of Computer and Electronics Engineers (UACEE), The Society of Digital Information and Wireless Communications (SDIWC) and LifeVtime member of Internet Society (ISOC),USA, Indian Mathematical Society, Cryptology Research Society of India and Ramanujan Mathematical Society of India (RMS). He is regular reviewer of sixty International Journals and International Conferences.

Sarita Gajbhiye Meshram received M. Tech degree in Soil and water engineering in 2009 with gold medal from College of Agricultural Engineering, Jawaharlal Nehru Krishi Vishwa Vidhyalaya, Jabalpur (M.P.); and PhD Degree in Water Resource Development and Management from IIT Roorkee (U.K.) India in 2015. She is currently Dr. D.S. Kothari Post-Doctoral Fellow in the Department of Mathematics and Computer Sciences, Rani Durgawati University, Jabalpur, India. Her current research interests Include Geographical Information Systems, Rainfall-Runoff sediment yield modelling, SCS-CN. She is carrying out her research work in the field of Rainfall-Runoff, Sediment Yield, Water Quality, Application of RS and GIS Water Network and Cryptographic Protocol. He has published more than 50 research papers in refereed journals, conference and workshop proceedings, and books. She is a member of some international society and reviewer of the reputed journal.

Cheng-Chi Lee received the Ph.D. degree in Computer Science from National Chung Hsing University (NCHU), Taiwan, in 2007. He is currently a Professor with the Department of Library and Information Science at Fu Jen Catholic University. Dr. Lee is currently as an editorial board member of International Journal of Network Security, Journal of Computer Science, Cryptography, and International Journal of Internet Technology and Secured Transactions. He also served as a reviewer in many SCI-

index journals, other journals, other conferences. His current research interests include data security, cryptography, network security, mobile communications and computing, wireless communications. Dr. Lee had published over 200 scientific articles on the above research fields in international journals and conferences. He is a member of IEEE, the Chinese Cryptology and Information Security Association (CCISA), the Library Association of The Republic of China, and the ROC Phi Tau Phi Scholastic Honor Society.

Provable Multiple-Replica Dynamic Data Possession for Big Data Storage in Cloud Computing

Huiying Hou[1], Jia Yu[1,2,3], and Rong Hao[1]
(Corresponding author: Rong Hao)
College of Computer Science and Technology, Qingdao University[1]
266071 Qingdao, China
Institute of Big Data Technology and Smart City, Qingdao University[2]
266071 Qingdao, China
State Key Laboratory of Information Security, Institute of Information Engineering[3]
Chinese Academy of Sciences, 100093 Beijing, China
(Email: hr@qdu.edu.cn)
(Received Feb. 21, 2017; revised and accepted June 20, 2017)

Abstract

In order to avoid the data loss in cloud storage, some users prefer to store multiple replicas on the cloud server. Multiple-Replica Provable Data Possession (MR-PDP) schemes are proposed to check the integrity of remote multiple-replica data. In most of the previous schemes, the user has to generate a homomorphism authenticator based on BLS signature or RSA signature for each block of each replica before uploading them to the cloud. This can incur high overhead for the user especially when the data is very big. In this paper, we make use of the algebraic signature technique to generate authenticator for each block of each replica. Because most operations of algebraic signature are XOR operations, it only needs minimal computation and communication cost. Moreover, we design a new data structure named Divided Map-Version Table (DMVT) to efficiently support full dynamic data operations. The performance analysis demonstrates that our scheme is very efficient for verifying the integrity of multiple-replica dynamic big data.

Keywords: Algebraic Signature; Dynamic Data; Multiple Replicas; Provable Data Possession

1 Introduction

Cloud storage brings enormous convenience to the cloud user. However, the user loses direct control of their data in the cloud storage system. In order to detect whether the user's data is unabridged, Provable Data Possession (PDP) schemes have been proposed [9, 10]. In 2007, Ateniese *et al.* [1] firstly presented the definition of Provable Data Possession (PDP). Subsequently, a lot of PDP schemes have been proposed such as [8, 11, 12, 15, 17–24, 26].

The above PDP schemes are designed for single replica of user data. If cloud server suffers from some irresistible disasters, the user may lose some important data permanently. Therefore, it is necessary for the user to store multiple replicas of important data on multiple servers.

In order to check the integrity of multiple copies of data, multiple-replica PDP schemes have been proposed [3]. In most schemes, one replication technology is used to generate multiple replicas in distributed storage system. However, this method cannot resist collusion attack of cloud servers. Cloud servers can make data owner believe that they truly stored all replicas while they only save one replica in fact. In order to solve this problem, Curtmola *et al.* [5] proposed a PDP scheme named Multiple-Replica Provable Data Possession (MR-PDP). In this scheme, multiple replicas of data are stored on multiple servers across multiple data center. Subsequently, other Multiple-Replica PDP schemes were proposed [7, 13, 27]. However, the above schemes can only support static data. In many practical applications, the data owner might frequently update the data stored in the cloud. In order to solve this problem, Ayad *et al.* [2] proposed a provable multi-replica data possession scheme supporting dynamic data. This scheme adopts the homomorphic authenticator to generate the tag for each data block. The data owner needs to generate $m \times n$ homomorphic authenticators if the data file is divided into n blocks and the cloud server stores m replicas. In the process of calculating homomorphic authenticator, there are a lot of modular multiplication operations. This may incur high computation overhead for the data owner. Furthermore,

this scheme cannot efficiently support dynamic data operations, especially for insert or delete operations. Therefore, how to design an efficient Multiple-Replica PDP scheme supporting dynamic data operations is an interesting topic.

The main contribution of this paper can be summarized as follows:

- We propose a provable multiple-replica dynamic data possession scheme for big data storage. We make use of algebraic signature to generate authenticator for each data block, which incurs very low computation overhead for the user.
- We design a new data structure - Divided Map-Version Table (DMVT) - to efficiently support full dynamic data operations (on data block level), such as insertion, deletion, modification, and append. When the large-scale outsourced data are frequently updated, the proposed scheme incurs minimum computation cost.

Organization. The rest of this paper is organized in the following way: In Section 2, we give the definition of system model and present our design goals. Then we introduce the algebraic signature and the proposed data structure DMVT in Section 3. In Section 4, we describe our proposed scheme in detail. The security and performance analysis of the proposed scheme is presented in Section 5. In Section 6, we give the conclusions.

2 Problem Statement

2.1 System Model

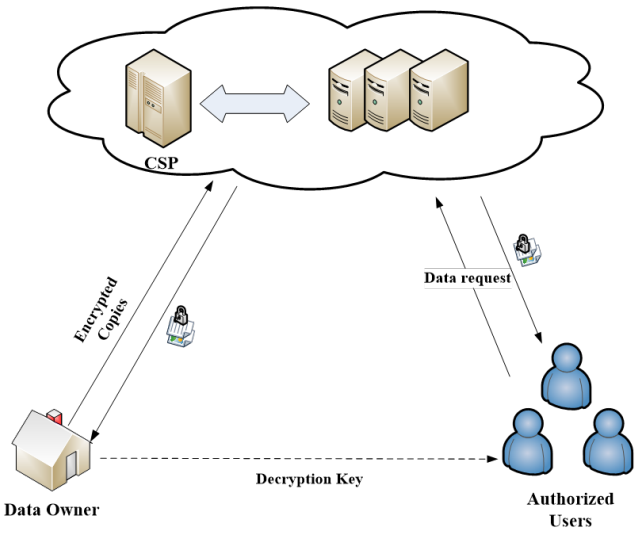

Figure 1: The system model

As shown in Figure 1, there are three types of entities in the system model: (1) Data Owner: an entity who can store large-scale data in the cloud, and then may perform modify, delete, insert, and append operations to update their outsourced data. (2) Cloud Storage Provider(CSP): an entity who provides storage service for the data owner and is in charge of managing the cloud servers. (3) Authorized Users: a collection of clients gain the access authorization from the data owner firstly, then they have the right to access the outsourced data and share the decryption key with the data owner. For the simplicity of description, we assume the data owner is in charge of checking the integrity of multiple replicas of data in our system model.

2.2 Design Goals

A Multiple-Replica PDP scheme should satisfy the following properties: (1) High efficiency: to allow data owner to efficiently check the integrity of multiple replicas of data. (2) Being against collusion attack: to ensure that colluded cloud servers cannot cheat users if they don't have all copies of data. (3) Supporting dynamic operations: to allow data owners to frequently update their outsourced data by performing insert, modify, delete, and append operations.

3 Preliminaries

In this section, we first introduce the algebraic signature technique used in our scheme; and then give the definition of our proposed data structure named DMVT.

3.1 Algebraic Signature

The algebraic signature is a type of hash functions with algebraic properties. The main property of algebraic signature is that the signature of the sum of some random file blocks is equal to the result of the sum of the signatures of the corresponding blocks. Therefore, we can compute the algebraic signature of the data block b_{ij} which is divided into s sectors:

$$S_\alpha(b_{ij}) = \sum_{k=1}^{s} b_{ijk} \cdot \alpha^{k-1}.$$

Note that α is an element in the Galois field. Moreover, b_{ij} denotes the j-th block of the i-th replica. There are some important properties of the algebraic signature [14]:

1) The algebraic signature of concatenation of message m_1 with l length and message m_2 can be computed as follows:

$$S_\alpha(m_1||m_2) = S_\alpha(m_1) \oplus l^\alpha S_\alpha(m_2).$$

2) The signature of the summation of several file blocks is equal to the summation of the signature of each block:

$$S_\alpha(b_{1j}+b_{2j}+...+b_{mj}) = S_\alpha(b_{1j})+S_\alpha(b_{2j})+...+S_\alpha(b_{mj}).$$

3.2 Divided Map-Version Table

The data structure named Divided Map-Version Table(DMVT) consists of two important components: (1) The logical index (L_j): a logical number of file blocks. (2) Version number (V_j): the current version of file blocks. The initial value of V_j is 1. When the data owner updates a data block, the corresponding V_j increases one. Assuming the file F is divided into 15 blocks, and 3 DMVTs are used to support dynamic operations, we display these 3 DMVTs in Figure 2. The DMVTs are stored on the data owner side.

DMVT1

Lj	Vj
1	1
2	1
3	1
4	1
5	1

{1,2,3,4,5}

DMVT2

Lj	Vj
6	1
7	1
8	1
9	1
10	1

{6,7,8,9,10}

DMVT3

Lj	Vj
11	1
12	1
13	1
14	1
15	1

{11,12,13,14,15}

Figure 2: An example of three DMVTs

4 The Proposed Scheme

4.1 Common Notations

$\phi : Z_q^* \times Z_q^* \to Z_q^*$, a pseudo-random function (PRF).

$\pi : Z_q^* \times \{1, 2, ..., n\} \to \{1, 2, ..., n\}$, a pseudo-random permutation (PRP).

$E_k(\cdot), D_k(\cdot)$: the encryption algorithm and the decryption algorithm of a symmetric cryptosystem with symmetric key k.

4.2 Scheme Description

The proposed scheme consists of six algorithms (*Setup*, *ReplicaGen*, *TagBlock*, *DataUpdate*, *Challenge*, *ProofGen*, *ProofVerify*).

Setup: Let G_1 be a multiplicative cyclic group generated by g with prime order q. The data owner randomly selects a secret key $k \xleftarrow{R} Z_q^*$, and computes a public key $y = g^x \in G_1$.

ReplicaGen: Assume that the file F is divided into n blocks $\{b_1, b_2, ..., b_n\}$. The data owner creates m differentiable replicas $\hat{F} = \{\hat{F}_i\}_{1 \le i \le m}$ that are stored on m cloud servers. A replica $\hat{F}_i$ is divided into n blocks $\hat{F}_i = \{\hat{b}_{ij}\}_{1 \le j \le n}$, where $\hat{b}_{ij} = E_k(i||b_j)$. Furthermore, the block $\hat{b}_{ij}$ is fragmented into s sectors with the same length. Denote block $\hat{b}_{ij} = \{\hat{b}_{ij1}, \hat{b}_{ij2}, ..., \hat{b}_{ijs}\}_{\substack{1 \le i \le m, \\ 1 \le j \le n}}$. So replica $\hat{F}_i = \{\hat{b}_{ijk}\}_{\substack{1 \le j \le n, , \\ 1 \le k \le s}}$, where each sector $\hat{b}_{ijk} \in Z_q$.

TagBlock: Given the distinct data file replicas $\hat{F} = \{\hat{F}_i\}_{1 \le i \le m}$, where $\hat{F}_i = \{\hat{b}_{ij}\}_{1 \le j \le n}$, the data owner generates a tag T_{ij} for each block $\hat{b}_{ij}$ by computing:

$$\begin{aligned} T_{ij} &= S_\alpha(\hat{b}_{ijk}||F_{id}||j||L_j||V_j) \\ &= \sum_{k=1}^{s} (\hat{b}_{ijk}||F_{id}||j||L_j||V_j) \cdot \alpha^{j-1} \end{aligned}$$

where L_j is the logical number of the block at physical position j, V_j is the current version of the block, and F_{id} is the unique name of the file F. In order to avoid the replay attack, the data owner computes:

$$\begin{aligned} C_{ij} &= S_\alpha(F_{id}||j||L_j||V_j) \\ &= \sum_{k=1}^{s} (F_{id}||j||L_j||V_j) \cdot \alpha^{j-1} \end{aligned}$$

The data owner computes $T_j = \sum_{i=1}^{m} T_{ij}$ and $C_j = \sum_{i=1}^{m} C_{ij}$ to reduce the storage overhead and the communication overhead of cloud servers. Hence, the CSP only needs to store n tags for the replicas $\hat{F} = \{F_i\}_{1 \le i \le m}$. Denote the set C as $\{C_j\}_{1 \le j \le n}$ and the set T as $\{T_j\}_{1 \le j \le n}$. The data owner sends $\{C, \hat{F}, T\}$ to the CSP, and deletes the local replicas and tags.

DataUpdate: The dynamic operations include Block Modification (denoted by **BM**), Block Insertion (denoted by **BI**), Block Append, and Block Deletion (denoted by **BD**).

- **Block Modification:** Assume that the data owner wants to modify a block b_j with b'_j in file $F = \{b_1, b_2, b_3, ..., b_n\}$ for all file replicas $\hat{F} = \{F_i\}_{1 \le i \le m}$. The data owner does as follows:
  1) Finds the corresponding V_j, then updates $V_j = V_j + 1$. The data owner recomputes C_{ij}.
  2) Generates m differentiable blocks $\{\hat{b}'_{ij}\}$. Divides $\hat{b}'_{ij} = E_k(i||b'_j)$ intossectors $\{\hat{b}'_{ij1}, \hat{b}'_{ij2}, ..., \hat{b}'_{ijs}\}$.
  3) Computes a new tag for each block $\hat{b}'_{ij}$ as follows:

$$\begin{aligned} T'_{ij} &= \sum_{k=1}^{s} S_\alpha(\hat{b}'_{ijk} \,\|F_{id}\, \|j\, \|L_j\, \|V_j\,) \\ &= \sum_{k=1}^{s} (\hat{b}'_{ijk} \,\|F_{id}\, \|j\, \|L_j\, \|V_j\,) \cdot \alpha^{j-1} \end{aligned}$$

And computes an aggregated tag $T'_j = \sum_{i=i}^{m} T'_{ij}$.

4) Sends a modification message $< F_{id}, BM, j, \{\hat{b}'_{ij}\}_{\substack{1\le i\le m,\\ 1\le j\le n}}, T'_j >$ to the CSP.

When the CSP receives the modification message from the data owner, he replaces the block $\hat{b}_{ij}$ with $\hat{b}'_{ij}$ for $1 \le i \le m$, and replaces T_j in the set T with T'_j.

- **Block Insertion:** Assume that the data owner wants to insert a new block $\hat{b}$ after the j-th block in file $F = \{b_1, b_2, ..., b_n\}$, and the new file is $F' = \{b_1, b_2, ..., b_j, \hat{b}, ..., b_{n+1}\}$. The data owner does as follows:

1) Finds the location of the j-th block in the DMVT table according to the tuples of the range of L_j.
2) Inserts a new table entry $< L_{j+1}, V_{j+1} >=< n+1, 1 >$ in the DMVT after position j. And recomputes C_{ij}.
3) Generates m distinct blocks $\{\hat{b}_i\}_{1\le i\le m}$, where $\hat{b}_i = E_k(i||\hat{b})$. Each block of $\{\hat{b}_i\}_{1\le i\le m}$ is fragmented into s sectors.
4) Computes a new tag $\hat{T}_i$ for each block $\hat{b}_i$ as follows:

$$\hat{T}_i = \sum_{k=1}^{s} S_\alpha(\hat{b}_{ik}||F_{id}||j+1||L_{j+1}||V_{j+1}) = \sum_{k=1}^{s} (\hat{b}_{ik}||F_{id}||j+1||L_{j+1}||V_{j+1}) \cdot \alpha^{j-1}$$

And computes an aggregated tag $T' = \sum_{i=1}^{m} \hat{T}_i$.

5) Sends an insert message $< F_{id}, BI, j, \{\hat{b}_i\}_{1\le i\le m}, T' >$ to the CSP.

Upon receiving the insert message, the CSP inserts the new block $\hat{b}$ for each file replica to generate new file replicas $\{\hat{F}_i\}_{1\le i\le m}$, and constructs the new file replicas $\{\hat{F}'_i\}_{1\le i\le m}$. The CSP also inserts T' after the position j of aggregated tags.

- **Block Append:** The append operation of the data block is equivalent to performing an insert operation after the last block of the file.
- **Block Deletion:** If the data owner hopes to remove the block at position j from all replicas, and he deletes the entry at position j from the DMVT. Meanwhile the number of elements in L_i decreases one. He sends the deletion request $< F_{id}, BD, j, null, null >$ to the CSP.

Upon receiving the deletion request, the CSP deletes the blocks $\{\hat{b}_{ij}\}_{1\le i\le m}$ and T_j from T. The CSP outputs the new file replicas $\hat{F}' = \{\hat{F}'_i\}_{1\le i\le m}$ and a new set $T' = \{T_1, T_2, ..., T_{j-1}, T_{j+1}, ..., T_{n-1}\}$.

Figure 3 shows the changes of the DMVTs for different dynamic operations. The file F is divided into 15 blocks, and 3 DMVTs are used to support dynamic operations. As shown in Figure 3(a), the initial value of Vi is 1. As shown in Figure 3(b), when the data owner modifies f[9], $V9$ is incremented by 1. To insert a new block after f[9], Figure 3c shows that a new entry $< 16, 1 >$ is inserted after the position 9, where 16 is the logical index the new inserted block, and 1 is the version number of the new inserted block. As shown in Figure 3(d), the append operation of the block is equivalent to performing an insert operation after the last block of the file. Deleting a block f[3] requires deleting the table entry $< L3, V3 >$ and shifting all subsequent entries one position up (as shown in Figure 3(e)).

Challenge:

1) The data owner chooses a random value c as the number of the challenged blocks.
2) And then picks two random numbers $k_1 \underleftarrow{R} Z_q^*, k_2 \underleftarrow{R} Z_q^*$.
3) Finally, the data owner sends the challenge $chal = (c, k_1, k_2)$to the CSP.

ProofGen:

1) After receiving the challenge from the data owner, the CSP computes $l_t = \pi_{k_1}(t)$ and $a_t = \phi_{k_2}(t)$ for $1 \le t \le c$.
2) And then the CSP computes $\mu = \sum_{t=1}^{c} T_{l_t} \oplus C_{l_t}$ and $\sigma_k = \sum_{t=1}^{c} \hat{b}_{i,l_t,k}$ for $1 \le k \le s$.
3) Finally, the CSP returns a proof (σ, μ).

ProofVerify: Upon receiving the proof from the CSP, the data owner does as follows:

Firstly, computes $l_t = \pi_{k_1}(t), a_t = \phi_{k_2}(t)$ for $1 \le t \le c$. And then checks whether the following verification equation holds or not.

$$S_\alpha(\sigma) \stackrel{?}{=} \mu.$$

If this equation holds, it means that the CSP properly stores all the replicas of the file. Otherwise, not.

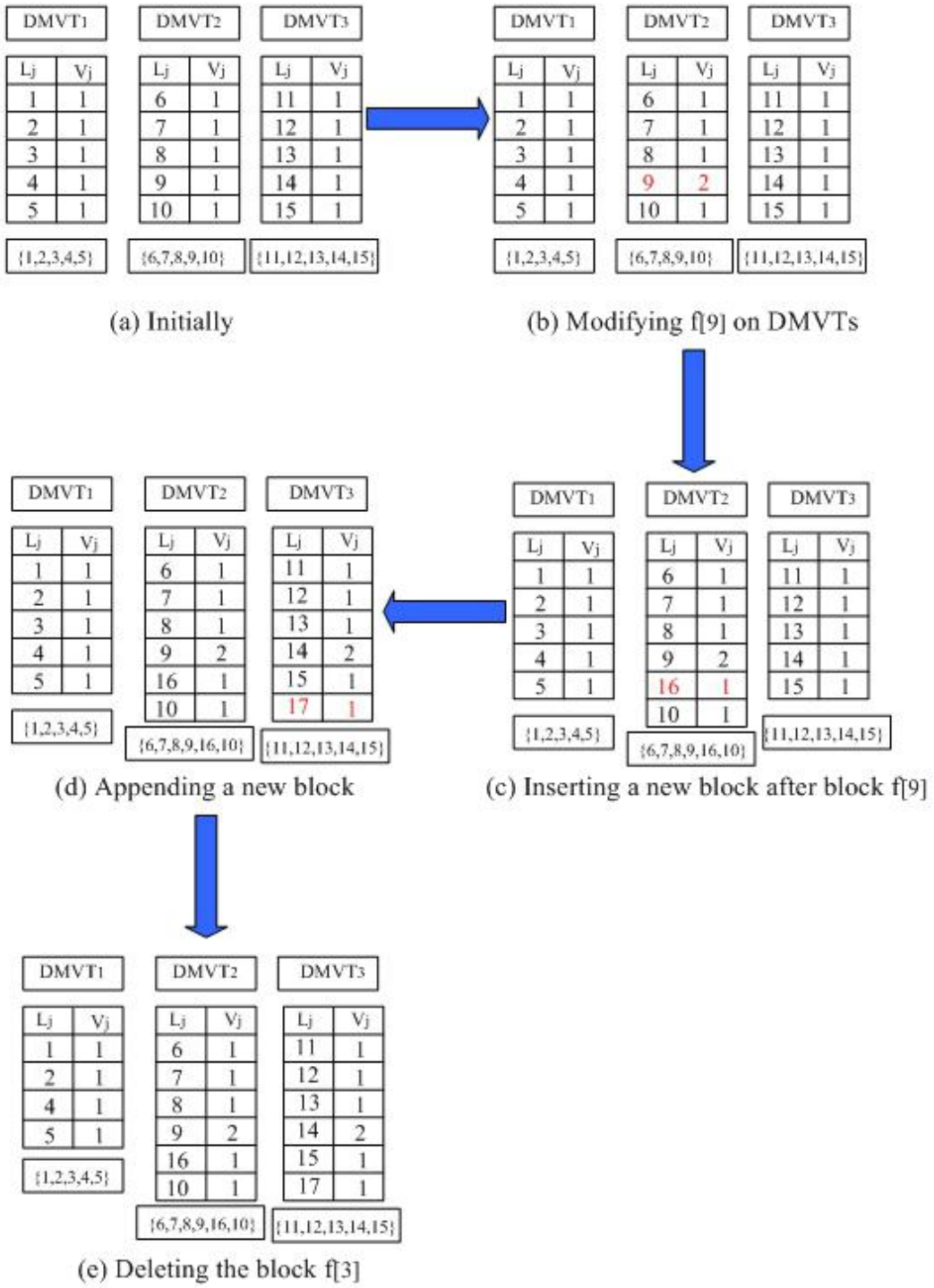

Figure 3: The changes of the DMVTs for different dynamic operations

5 The Security and Performance Analysis

5.1 The Correctness and Security Analysis

5.1.1 The Correctness Analysis

We analyze the correctness and the security of the proposed scheme. In the ProofVerify stage, we firstly extend by using the properties of the algebraic signature as follows:

$$\begin{aligned}\mu &= \sum_{t=1}^{c} T_j \oplus C_j \\ &= \sum_{t=1}^{c}\sum_{k=1}^{s}(S_\alpha(\hat{b}_{ijk}||F_{id}||j||L_j||V_j) \oplus S_\alpha(F_{id}||j||L_j||V_j)) \\ &= \sum_{t=1}^{c}\sum_{k=1}^{s} S_\alpha(\hat{b}_{ijk}) \oplus l^\alpha \oplus S_\alpha(F_{id}||j||L_j||V_j) \\ &\quad \oplus S_\alpha(F_{id}||j||L_j||V_j)) \\ &= \sum_{t=1}^{c}\sum_{k=1}^{s} S_\alpha(\hat{b}_{ijk}).\end{aligned} \tag{1}$$

And then we demonstrate the correctness of the above verification equation as follows:

$$\begin{aligned}S_\alpha(\sigma) &= S_\alpha(\sum_{t=1}^{c}\sum_{k=1}^{s}\hat{b}_{ijk}) \\ &= \sum_{t=1}^{c} S_\alpha(\sum_{k=1}^{s}\hat{b}_{ijk}) \\ &= \sum_{t=1}^{c}\sum_{k=1}^{s} S_\alpha(\hat{b}_{ijk}) \\ &= \mu.\end{aligned}$$

5.1.2 The Security Analysis

Theorem 1. *(The auditing soundness)* In the proposed scheme, the cloud can pass the verification only if it actually stores intact data.

Proof. If the cloud passes the verification but does not possess the intact data, it means that the cloud can forge the valid algebraic signature for any message. Algebraic signature condenses a large block into a bit string. The bit string can be made long enough to make an accidental almost impossible to happen. For example, a 64 bits signature will suffer a collision with probability 2^{-64} and a 256 bits signature with probability 2^{-256}. It is probabilistically impossible for a site that does not know any secret to generate a coherent set of signatures. As a result, the algebraic signature is secure enough for checking the integrity of multiple replicas. □

Theorem 2. *(The resisting collusion attack of cloud servers)* In the proposed scheme, the cloud cannot make data owner believe that they truly stored all replicas, but in real they only save one replica.

Proof. In our scheme, the data owner creates m differentiable replicas $\hat{F} = \{\hat{F}_i\}_{1\le i\le m}$ that are stored on m cloud servers. The cloud sever can not know the content of replicas stored on other cloud servers. The proof (σ, μ) generated by the CSP will be valid and will pass the verification equation $S_\alpha(\sigma) \stackrel{?}{=} \mu$ only if all copies are intact. Thus, when there is one or more corrupted copies, the whole auditing procedure fails. So, the cloud cannot make data owner believe that they truly stored all replicas, but in real they only save one replica. □

Theorem 3. *(Detectability)* Our proposed auditing scheme is $(\frac{m}{n}, 1-(\frac{n-1}{n})^c)$ detectable if the cloud stores a file with n blocks including m bad (deleted or modified) blocks, and c blocks are challenged.

Proof. Assume that the cloud stores a file with total n blocks including m bad (deleted or modified) blocks. The number of challenged blocks is c. Thus, the bad blocks can be found out if and only if at least one of the challenged blocks chosen by the verifier matches the bad blocks. We use a discrete random variable X to denote the number of blocks selected by the challenger that matches the block-tag pairs changed by the adversary. We use PX to denote the probability that at least one block chosen by the challenger matches the blocks changed by the adversary. So

$$\begin{aligned}P_X &= P\{X \ge 1\} \\ &= 1 - P\{X = 0\} \\ &= 1 - \frac{n-m}{n}\frac{n-1-m}{n-1} \times ... \times \frac{n-c+1-m}{n-c+1}.\end{aligned}$$

We can get $P_X \ge 1-(\frac{n-m}{n})^c$. Thus, the proposed auditing scheme is $(\frac{m}{n}, 1-(\frac{n-m}{n})^c)$ detectable if the cloud stores a file with n blocks including m bad (deleted or modified) blocks, and c blocks are challenged. □

5.2 Performance Analysis

Compared with schemes [2,6,13,25], the proposed scheme is more efficient and has two advantages. In the following paragraphs, we will thoroughly explain why our scheme has two advantages over these schemes.

1) In this work, the computation overhead on the data owner is greatly reduced. In schemes [2, 6, 13, 25], the data owner uses BLS signature to generate a homomorphism authenticator for each block of each replica. In the process of calculating homomorphic authenticator, there are a lot of modular multiplication operations. This incurs high computation overhead for the data owner. So the data owner needs to

be powerful enough to perform these costly computations when the data are outsourced. However, in the real world, the data owner (*eg*, using PDAs and mobile phones) may possess low computation capabilities. Observing this fact, we select to use the algebraic signature technique to generate authenticator for each block of each replica in this paper. Because most operations of algebraic signature are XOR operations, the computation overhead on the data owner is minimal.

2) Our scheme efficiently supports full dynamic data operations. As shown in Figure 5, our proposed scheme is more efficient than the scheme [2] when the data owner frequently performs data update. In the scheme [2], the computation overhead during the insert and delete operations is O(n), where n is the number of the file blocks. In our scheme, the data owner only needs to shift a part of the outsourced data blocks ($\frac{n}{k} - i$) that incurs only O($\frac{n}{k}$) computation overhead on the data owner side when he inserts or deletes a data block. In the scheme [13], the data owner uses fully homomorphic encryption algorithm to generate multiple copies and to support data block dynamic operations. We know that fully homomorphic encryption algorithm can incur heavy computing burden on data owner and is inefficiency according to the scheme [4]. So, our proposed scheme is more efficient than the scheme [13]. The scheme [25] supports batch verification based on identity but not allows data owner stores multiple replicas on the cloud and not supports data block dynamic operations. In the scheme [6], the cloud uses skip list to support data block dynamic operations and the computation overhead during the insert and delete operations is O(n). Moreover, in this scheme, data file F is split into blocks, and each data block is split into sectors. In the scheme [6], data file F is split into blocks, but the data block is not split into sectors. In fact, the fragment operation of data blocks can reduce the number of data authenticators. Obviously, our scheme is more efficient than the scheme [6].

Here we analyze the storage and communication overhead of our proposed scheme. For a concrete example of using our scheme, we consider an algebraic signature is 256 bits, a 102MB file F is divided into 125,00 blocks (each block is 8KB) and the file F has 10 replicas. In the Setup stage, the data owner only needs to store one secret key k (160 bits). During the *TagBlock* stage, the data owner stores the file and its tags on cloud servers. The additional storage is less than 4MB. In the *challenge* phase, the data owner sends challenge message to server, and the size of this message is about 480 bits. If the server deletes at least 1% of F, the data owner can detect server misbehavior with probability over 99% by asking proof for 460 blocks. The response of server is less than 8KB.

5.2.1 Experiment Results

With the help of Pairing-Based Cryptography (PBC) library [16], we evaluate the proposed scheme in several experiments. We conduct these experiments on a Linux server with Intel processor running at 2.70 GHz and 4 GB memory. We choose a bilinear map that uses a supersingular curve to achieve the fast pairing operations. Therefore, the base field is 160bits, the size of an element in Z_q^* is 20 bytes, and the size of an element is 128 bytes. In our experiments, the file is set to 102MB consisting of 125,00 blocks, and has 10 replicas.

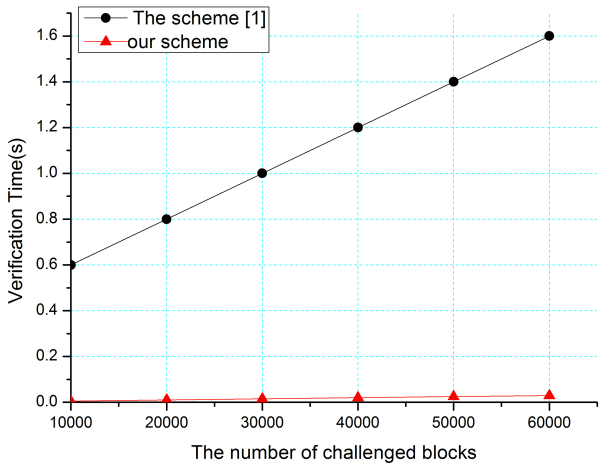

Figure 4: Computation overhead of data owner in verification phase

In our scheme, the most of operations in verification phase are XOR operations. In scheme [2], the most of operations are multiplication operations. We show these two schemes verification time with different number of the challenged blocks in Figure 4. We can see that the verification time in scheme [2] is about 0.6s with the 10,000 challenged blocks. In contrast, the verification time in our scheme is only about 5ms with the 10,000 challenged blocks. Therefore, the verification time in our scheme is remarkable efficient than that in scheme [2].

In Figure 5, we demonstrate the efficiency of the scheme [2] and our scheme when the data owner frequently performs data update. In our scheme, 10 DMVTs are used to support insert or delete operation. In the experiment, we consider computation time of inserting or deleting a block(i) with the number of updated blocks increasing from 100 to 1000. When insert or delete a block(i) in the scheme [2], the data owner firstly looks for the precise position of the block(i); and then shifts (n-i) blocks. This process will incur high computation cost, when the data owner frequently performs data update. Our scheme overcomes this weakness because 10 DMVTs

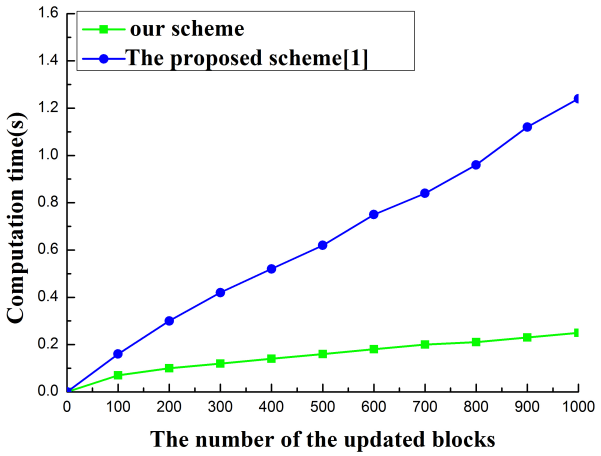

Figure 5: Comparison of computation cost of frequent data update

are used to enhance the efficiency. When the data owner frequently performs data update, our proposed scheme is more efficient than the scheme [2].

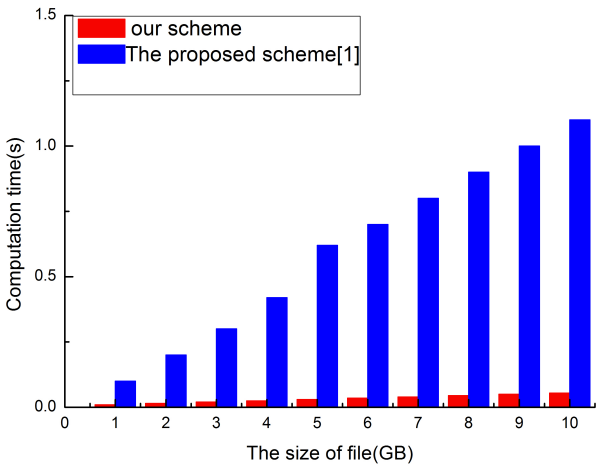

Figure 6: Comparison of computation cost when number of update requests is 100

In Figure 6, we show the computation cost of dynamic data update with the file size from 1GB to 10GB. Assume that 100 blocks are inserted or deleted. In scheme [2], the data owner needs to shift a large number of data blocks, so it will incur high computation overhead. When the data size increases from 1GB to 10GB, the computation time increases from 0.1s to 1.1s. In contrast, our scheme can remarkably reduce the computation overhead. For a 10GB file, the computation time is only 0.05s in our scheme. Therefore, our scheme is very efficient for large-scale files.

As shown in Table 1, the storage space of file copies

Table 1: Storage and communication overheads in our scheme and schemes [2, 3, 26]

Costs		Our Scheme	The Scheme[1]	The Scheme [2]	The Scheme[17]
Storage	File Copies	$m\|F\|$	$m\|F\|$	$m\|F\|$	$m\|F\|$
	CSP Overhead	$m\|F\|+2n\times256$ bits	$m\|F\|+(n+s)\times256+64$ bits	$m\|F\|+(n+s)\times256+64$ bits	$(4m+1)\times n\times256$ bits
	Verifier Overhead	2×256 bits	2×256 bits	2×256 bits	2×256 bits
Communi-cation	Challenge	$256+\log_2(c)$ bits	$256+\log_2(c)$ bits	$256\times\log_2(c)$ bits	512 bits
	Response	2×256 bits	$257+256sm$ bits	$257+256sm$ bits	$256+1024$ bits

($|F|$ is the size of the file F; m is the number of file copies; n is the number of data blocks; s is the number of sectors of one data block; c is the number of the challenged blocks.)

in our scheme is equal to that in schemes [2, 3, 26]. The CSP overheads in both our scheme and schemes [2, 3, 13] are linear in n. The verifier overheads and the size of the challenge message in our scheme are equal to those in the scheme [2, 3]. Especially, the size of response message in our scheme is almost constant and far less than that in schemes [2, 3]. Compared with the scheme [13], the size of challenge message and response message are less than that in scheme [13].

6 Conclusion

In this paper, we propose a provable multiple-replica dynamic data possession for big data storage in cloud computing. In our scheme, we use the algebraic signature to reduce the computation and communication overhead on the data owner side. Meanwhile, in order to achieve efficient dynamic operation, we design a new data structure DMVT. The experimental results demonstrate that our scheme is efficient.

Acknowledgments

This research is supported by National Natural Science Foundation of China (61572267, 61272425, 61402245), the Open Project of the State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences(2017-MS-21, 2016-MS-23).

References

[1] G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson and D. Song, "Provable data possession at untrusted stored", in *Proceeding of ACM CCS*, pp. 598-609, 2007.

[2] A. F. Barsoum and M. A. Hasan, "Provable Multireplica Dynamic Data Possession in Cloud Computing Systems," *IEEE Transactions on Information Forensics and Security*, pp. 485-497, 2015.

[3] A. F. Barsoum and M. A. Hasan, *Provable Possession and Replication of Data over Cloud Servers*, Cryptographic Research, The University of Waterloo, USA, 2010.

[4] G. Craig, "Computing on encrypted data," *Lecture Notes in Computer Science*, vol. 5888, pp. 477-477, 2009.

[5] R. Curtmola, O. Khan, R. Burns and G. Ateniese, "MR-PDP: Multiple-replica provable data possession," in *Proceedings of the 28th International Conference on Distributed Computing Systems*, pp. 411-420, 2008.

[6] M. Etemad, A. Kupcu, "Transparent, distributed, and replicated dynamic provable data possession," in *International Conference on Applied Cryptography and Network Security*, pp. 1-18, 2013.

[7] Z. Hao and N. Yu, "A multiple-replica remote data possession checking protocol with public verifiability," in *Proceedings of 2nd International Symposium Data*, Privacy, E-Commerce, Sep., pp. 84-89, 2010.

[8] W. Hsien, C. Yang and M. S. Hwang, "A Survey of Public Auditing for Secure Data Storage in Cloud Computing," *International Journal of Network Security*, vol. 18, no. 1, pp. 133-142, 2016.

[9] M. S. Hwang, C. C. Lee, T. H. Sun, "Data error locations reported by public auditing in cloud storage service," *Automated Software Engineering*, vol. 21, no. 3, pp. 373–390, Sep. 2014.

[10] M. S. Hwang, T. H. Sun, C. C. Lee, "Achieving dynamic data guarantee and data confidentiality of public auditing in cloud storage service," *Journal of Circuits, Systems, and Computers*, vol. 26, no. 5, 2017.

[11] C. Liu, R. Ranjian, X. Zhang, C. Yang, D. Georgakopoulos and J. Chen, "Public Auditing for Big Data Storage in Cloud Computing–A Survey," in *Proceeding of 16th IEEE International Conference Computational Science and Engineering (CSE'13)*, pp. 1128-1135, 2013.

[12] C. Liu, W. Hsien, C. Yang and M. S. Hwang, "A survey of public auditing for shared data storage with user revocation in cloud computing," *International Journal of Network Security*, vol. 18, no. 4, pp. 650-666, 2016.

[13] M. Li, L. Wang and J. Wei, "Distributed data possession provable in cloud," *Distributed Parallel Databases*, vol. 35, pp. 1-21, 2017.

[14] W. Litwin and T. Schwarz, "Algebraic signatures for scalable distributed data structures," in *Twentieth IEEE International Conference on Data Engineering*, pp. 412-423, 2004.

[15] Y. Ming and Y. Wang, "On the security of three public auditing schemes in cloud computing," *International Journal of Network Security*, vol. 17, no. 6, pp. 795-802, 2015.

[16] Pairing-Based Cryptography(PBC) library, Feb. 9, 2018. (`https://crypto.stanford.edu/pbc/howto.html`)

[17] W. Shen, J. Yu, R. Hao and X. Wang, "A public cloud storage auditing scheme with lightweight authenticator generation," in *IEEE International Conference on P2P, Parallel, Grid, Cloud and Internet Computing*, pp. 36-39, 2015.

[18] C. Wang, Q. Wang, K. Ren and W. Lou, "Privacy-preserving public auditing for data storage security in cloud computing," in *The Proceedings of IEEE INFOCOM'10*, pp. 525-533, 2010.

[19] C. Wang, S. Chow, Q. Wang, K. Ren and W. Lou, "Privacy preserving public auditing for secure cloud storage," *IEEE Transactions on Computers*, vol. 62, no. 2, pp. 362-375, 2013.

[20] H. Wang, D. He and S. Tang, "Identity-based proxy-oriented data uploading and remote data integrity checking in public cloud," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 6, pp. 1165-1176, 2016.

[21] G. Yang, J. Yu, W. Shen, Q. Su, F. Zhang and R. Hao, "Enabling public auditing for shared data in cloud storage supporting identity privacy and traceability," *Journal of Systems and Software*, vol. 113, pp. 130-139, 2016.

[22] J. Yu, K. Ren, and C. Wang, "Enabling cloud storage auditing with verifiable outsourcing of key updates," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 6, pp. 1362-1375, 2016.

[23] J. Yu, K. Ren, C. Wang and V. Varadharajan, "Enabling cloud storage auditing with key-exposure resistance," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 6, pp. 1167-1179, 2015.

[24] J. Yuan and S. Yu, "Efficient public integrity checking for cloud data sharing with multi-user modification," in *Proceedings of IEEE INFOCOM*, pp. 2121-2129, 2014.

[25] F. Zhou, S. Peng, J. Xu and Z. Xu, "Identity-based batch probable data possession," in *International Conference on Provable Security*, pp. 112-129, 2016.

[26] J. Zhang, P. Li and M. Xu, "On the security of an mutual verifiable provable data auditing in public cloud storage," *International Journal of Network Security*, vol. 19, no. 4, pp. 605-612, 2017.

[27] Y. Zhang, J. Ni, X. Tao, Y. Wang and Y. Yu, "Provable multiple replication data possession with full dynamics for secure cloud storage," *Concurrency Computation*, vol. 28, no. 4, pp. 1161-1173, 2016.

Biography

HuiYing Hou received B.S. degrees in School of Computer Science and Technology from Qingdao University, China, in 2013. She will receive M.S. degree in the college of Computer Science and Technology from Qingdao University, China, in 2016. Her research is cloud computing security.

Rong Hao received Master degree in Institute of Network Security from Shandong University. She is working in the College of Computer Science and Technology, Qingdao University. Her research interest is information security.

Jia Yu received the M.S. and B.S. degrees in School of Computer Science and Technology from Shandong University, China, in 2003 and 2000, respectively. He received Ph. D. degree in Institute of Network Security from Shandong University, China, in 2006. Since 2012, he has been a full professor and the department director of information security, at Qingdao University, China. He was a visiting professor with the Department of Computer Science and Engineering, the State University of New York at Buffalo from 2013 to 2014. His research interests include cloud computing security, key evolving cryptography, digital signature, and network security. He has published over 100 academic papers. He is the reviewer of more than 30 international academic journals.

IDS Against Black-Hole Attack for MANET

Mohamed Abd-El-Azim, Hossam EL-Din Salah, and Menas Ebrahim
(Corresponding author: Menas Ebrahim)
Electronics and Communications Engineering Department, Mansoura University
Mansoura 35516, Egypt
(Email: menasebrahim@gmail.com)
(Received Feb. 13, 2017; revised and accepted June 25, 2017)

Abstract

Black-Hole and Gray-Hole attack considers two of the most affected kind of attacks on the Mobile Ad-Hoc Network (MANET). Therefore, the use of intrusion detection system (IDS) has a major importance in the MANET protection. In this paper, a proposed optimized fuzzy based intrusion detection system is presented with an automation process of producing a fuzzy system by using an Adaptive Neuro-Fuzzy Inference System (ANFIS) for the initialization of the FIS and then optimize this initialized system by using Genetic Algorithm (GA).

Keywords: ANFIS; Black-Hole Attack; FIS; GA

1 Introduction

Mobile ad-hoc network (MANET) is a new and evolving area of interests, which used in many different applications [6]. The black-hole attack is a Denial of Service (DoS) attack; it works by drawdown packets in the network to a malicious node and then drops, alters the content of the packets, or even passes the packets to another malicious node [1]. Another effective attack is gray-hole attack [7], MANET works under an assumption that all nodes in the network are collaborating to forward packets [3], which is not true as there are selfish nodes that refuse to forward the packets to reserve its energy and other resources, also there are attack nodes, which drop packet to harm the network. One of the important parts in utilizing and deploying the MANET is securing it. Achieving a secure MANET helps this kind of network to achieve its full potential, which is to be used not only in military and crises situation applications but also in a commercial way. A certain level of security can accomplish by using the existing security solution [25]. However, because of the nature of the MANET, it has its own vulnerabilities coupled with the normal vulnerabilities of the wireless networks [5]. Therefore, these solutions cannot provide a sufficient security level. Intrusion detection systems [2] with the traditional security solutions can accomplish a sufficient security level. In this paper, a proposed intrusion detection system (IDS) introduced against the black-hole and gray-Hole attack where an adaptive neuro-fuzzy inference system (ANFIS) used to automate the process of producing a fuzzy system and then optimizes this system using the genetic algorithm (GA). The system tested in the presence of black-Hole attack and the presence of both black and gray-hole attack.

The rest of this paper is organized as follow: Section 2: literature survey; Section 3: problem statement; Section 4: proposed systems; Section 5: performance evaluation; Section 6: results are discussions, and Section 7: conclusions.

2 Literature Survey

There are many techniques used to detect the packet drop attack (mainly black-Hole attack and gray-attack) some of these techniques are presented below [8, 24].

2.1 PDRR Based Detection Method

The packet Drop Ratio (PDRR) used in [19] to detect the behavior of black-Hole nodes. The PDRR calculated from the Packet Delivery Ratio (PDR) where it is used as a performance metric. The maximum PDRR calculated in an attack free network and then sets as a threshold value.

In network exhibiting attacks the PDR calculated for each node, the node that has a PDR exceeds the threshold value consider malicious otherwise, its behavior consider normal.

2.2 Promiscuous Mode Detection Method

A secure routing protocol presented in [21], which is a secure modified version of the ad-hoc on demand distance vector (AODV) routing protocol. In this modified routing protocol, the promiscuous mode used to detect malicious nodes. The promiscuous mode allows any node to overhear the communication of its neighbors. As soon as the

tested node sends a route reply (RREP) message to the source node replying to the route request (RREQ) message, its neighbor promiscuously hears this RREP. So, it sent a plane packet to the tested node to see if it forwarded it to the destination if it does the testing node is considered normal and if not malicious.

2.3 Additional RREP Detection Method

In [13], a secure AODV routing protocol presented to detect the malicious nodes by adding a preprocessing stage called preprocess RREP. In this method, all the RREP messages received in a predefined time slot is stored to find the freshest RREP, analyze the data, and secure route to the destination. When the source node receives RREP messages it stores them and compares its destination sequence number, whenever a RREP message received with a much higher destination sequence number the RREP will be discarded and the black-Hole attack is detected.

2.4 Watchdog and Pathrater Detection Method

The watchdog and pathrater technique was introduced by Marti *et al.* in [15]; it was added on top of the standard routing protocol to increase the throughput of the network when malicious nodes appear in the network. This method is divided into two parts: watchdog part and Pathrater part. The watchdog part works as IDS for MANETs to prevent malicious nodes. This is done by promiscuously listening to its next hop's transmission, if the node does not transmit the packet within a predefined time the watchdog increases its failure counter, Whenever a node's failure counter surpasses a predefined limit, the Watchdog hub reports it as getting out of hand. When a node reported as a misbehaving node the pathrater which is the other part of the technique work with the routing protocol to avoid the misbehaving nodes in the future transmission. This technique proved itself efficient; it is also a node detection technique rather than link detection technique.

2.5 Permutation-Based ACK Detection Method

In [9], Dave proposed an Ad-hoc On-demand Multipath Secure Routing (AOMSR), which is an improvement of the AODV routing protocol with a security mechanism based on the adaptive acknowledgment (AACK) and TWO-ACK security mechanism. In this protocol, the source node stores all the paths to the destination that came from the RREP message. After detecting many routes to the destination, the source node sends the same packet throughout those paths. Every time the destination node receives this packet from any of these paths, the destination node sends back a permutated acknowledgment. If one of these paths does not send back a permutated acknowledgment, a black-Hole attack can be detected.

2.6 Using Fuzzy Logic Approach

The fuzzy logic used in intrusion detection since 90's [16] because it is able to deal with uncertainty and complexity, which derived from human reasoning. By the help of fuzzy variables or linguistic terms, intrusion detection features can be viewed easily and the decision of normal and abnormal activity in the network is based on its fuzziness nature that can identify the degree of maliciousness of a node instead of yes or no conditions. IF-then-else based fuzzy rules are used to define all situations in the network for identifying the attacks or intrusions. The fuzzy rule-based system is known as fuzzy interference system (FIS) that is responsible for taking decisions.

Method (1): Ramkumar in [17], proposed a fuzzy based IDS where forward packet ratio and the average destination sequence number is used to distinguish normal from malicious. The system is divided into four parts: (1) Fuzzy factor withdrawal, (2) Fuzzy calculation, (3) Fuzzy confirmation module and (4) Alarm packet generation module.

Method (2): Balan in [4], proposed a fuzzy based IDS for black-Hole and gray-hole attack. The proposed system consists of three main blocks they are: attack categorization, fuzzy implementation, and fuzzy estimation. The number of packets dropped by the node is used in the fuzzy implementation module.

Method (3): Wahengbam in [29] proposed a fuzzy based IDS. The parameters used in work were the number of packets lost and the number of packets forwarded by the node.

Method (4): Sengar *et al.* in [20] proposed a fuzzy based where a trust level is calculated by a proposed formula. Three ranges to this trust level used to categorize the nodes and differ the normal from abnormal behavior.

Method (5): Vydeki *et al.* used in [28] a Sugeno type-2 FIS to detect the black-Hole attack. It is proven to have 97% detection rate.

3 Problem Statement

The black-hole attack is a denial of service attack which drawdown the network traffic to a specific malicious node. The attack node in this type of attack act maliciously in the route discovery process [14], this is done by sending a fake route reply message to a requesting source node when it sends a route request message with a fake destination sequence number to fool the source node that it is the shortest path to the destination. Then it drops, alters the content of the packets, or even passes the packets

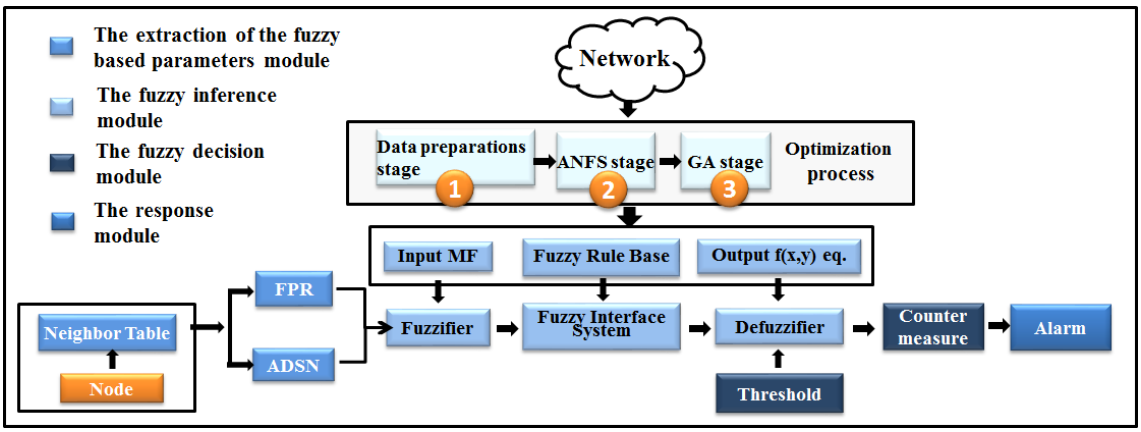

Figure 1: The normal proposed system in detail

to another malicious node. The gray-hole attack is also a denial of service attack like the black-Hole attack. The big difference between the two is that the black-Hole attack act maliciously from the beginning at the route discovery process but the gray-hole attack does not. In gray-hole attack, the malicious node acts legitimately in the route discovery process by sending a true RREP message, but if it is chosen to forward packets is act maliciously [11]. The malicious node can accomplish the gray-hole attack selectively. This can be done by dropping packets for a specific destination, or at a defined part of the day, or by dropping a packet every t seconds or every n packets, or even a randomly selected portion of the packets [18]. To detect the behavior of this attack and prevent it from affecting the network an IDS mechanism must be used.

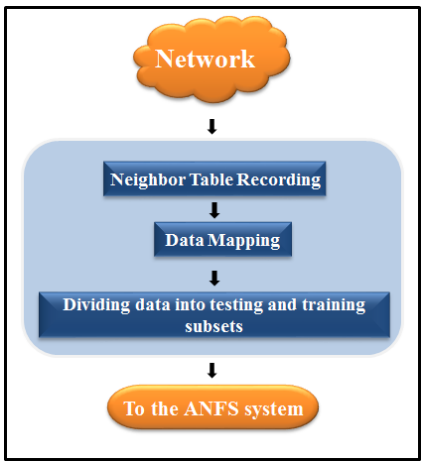

Figure 2: The data preparation stage

4 Proposed Systems

In many types of research, the solution of detection the black and gray-Hole attack comes with the use of FIS, which relies on the researcher experience to understand the system very well in order to choose the number of the membership function for each fuzzy set, the shape, and the position of each one. In addition, it requires an effort from the researcher's hand to set the rule base for that fuzzy system (noticing that even with a high expert researcher these parameters are difficult to be optimized). In order to see the effectiveness of the optimization process in discovering the black-hole attack and gray-Hole attack a fuzzy based IDS is introduced. A similar optimization process used for grade estimation in [26].

The proposed intrusion detection system illustrated in Figure 1 consists of four main modules: (1) Extraction of the fuzzy based parameters module; (2) Fuzzy inference module; (3) Fuzzy decision module; and (4) Response module. To optimize and automate the fuzzy interface module an optimization process is done which includes three stages: Data preparations stage, ANFS stage, and GA stage.

4.1 Extraction of Fuzzy Based Parameters

In this module, a set of parameters chosen to be extracted from the network (this parameter should be the most affected parameters when attack nodes are presented in the network); in this system, the forward packet ratio (FPR) and the average destination sequence number (ADSN) are chosen [23] as an input to FIS. In addition, the fidelity level is chosen as an output from the FIS. To do that a neighbor table is presented to every node in the network to be able to store the number of forwarded data packets, the no. of the packets that the neighbor has been sending, and the destination sequence numbers that the node receives from the neighbor each time it sends an RREP message to it. Equation (1) can calculate FPR and ADSN for each neighbor as follows:

$$FPR = \frac{\text{no. of the packets that the neighbor has been send}}{\text{no. of forwarded data packets to the neighbor}} \quad (1)$$

ADSN for each node calculated by averaging the destination sequence numbers stored in the neighbor table in a predefined time slot.

4.2 Fuzzy Inference Module

An automation process is done to find the number of the membership function for each fuzzy set, the shape, and the position of each one to minimize the error that can be done by setting these parameters manually. To optimize and automate the fuzzy interface module an optimization process is done which includes three stages: Data preparations stage, ANFS stage, and GA stage.

4.3 Data Preparations Stage

A database is extracted from the network by recording all the activity of all the nodes in the network. Then a mapping process is done by mapping the normal activity with high FL (10 is chosen in the system) and the malicious activity with low FL (0 is chosen in the system). In the learning process, the normal activity and the malicious activity is known by the IP address of each node. After that, the input parameters "forward packet ratio" and "average destination sequence number" must be calculated from the database. The entire sets are divided into two groups training group and testing group the first is two-third of the data set and the second is the remaining third. See Figure 2 for the Data Preparations Stage process.

4.4 ANFS Stage

A generation of the initial individuals of the FIS is done in this stage which will be optimized in the GA stage. A Sugeno FIS with Gaussian MFs is chosen in this stage Figure 3 is the MFs for initial FIS. In addition, see Figure 4 for ANFS stage process.

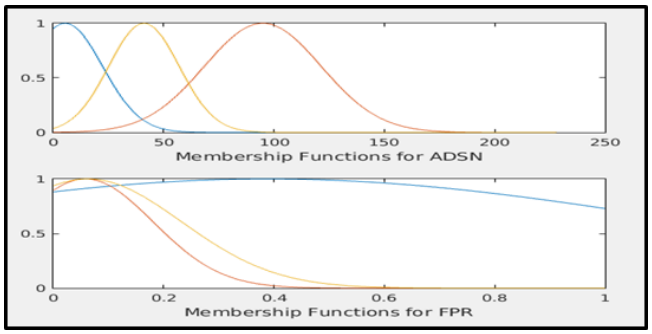

Figure 3: The membership function for initial FIS

4.5 GA Stage

This module used as an optimization tool. Since the GA deal with chromosomes, the variables presented to GA encoded by chromosome. Since each Gaussian MFs has two variables (mean "M" and standard deviation "SD") and each rule has three variables (p_i q_i r_i)) the chromosome should look like, see Equation (1) [22, 27]:

$$M_1SD_1M_2SD_2 \ldots\ldots\ldots p_1q_1r_1 \ldots\ldots\ldots p_3q_3r_3 \quad (2)$$

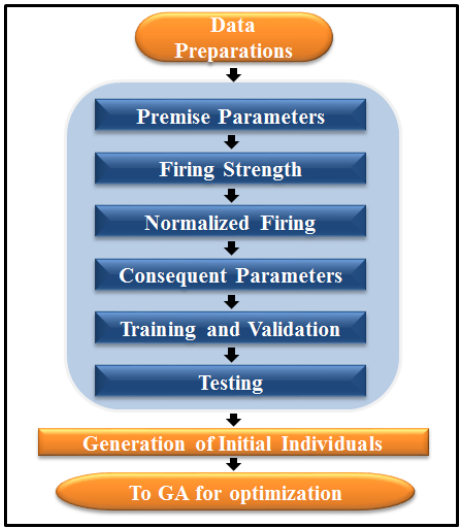

Figure 4: The ANFS stage

The initial population of individuals, which called the parent population is evaluated by the fitness function, which is the Mean Square Error (MSE). See Equation (3):

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(P_i - T_i)^2 \quad (3)$$

Where P_i is the value of from the GA system, T_i is the target value and n is the number of data in the training dataset. GA started with 25 randomly generated chromosomes, and their parameters were crossover percentage, mutation rate and population size with the values of 0.4, 0.15, and 25, respectively. Figure 9 shows GA optimization process and Figure 9 shows the optimized membership functions.

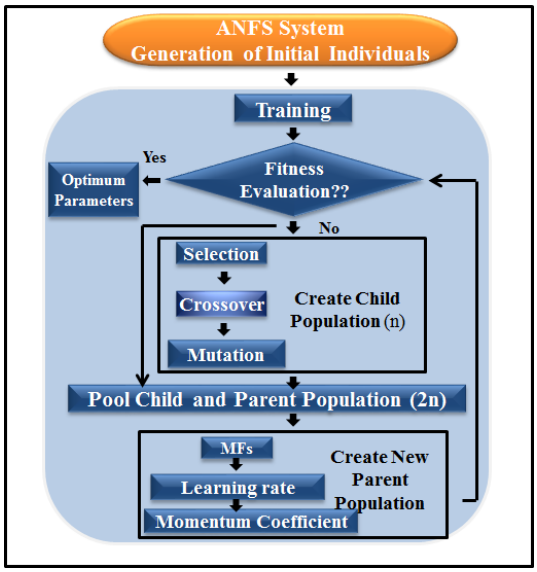

Figure 5: The genetic algorithm stage

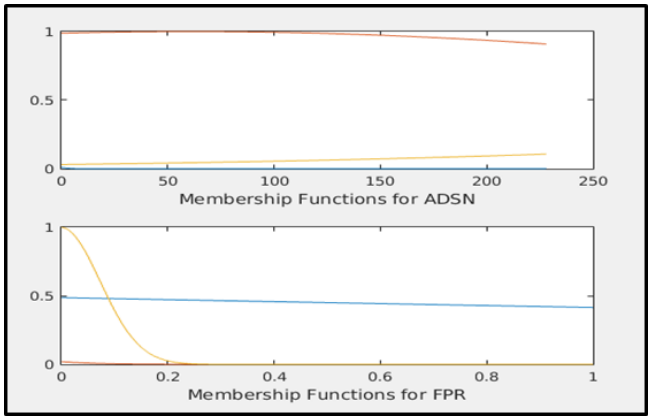

Figure 6: The membership functions for optimized FIS

4.6 Fuzzy Decision Module

A threshold value with three as a value is used in this module to distinguish the normal from the abnormal activity. If the resulted FL from the Fuzzy Inference Module is above the threshold value the node considers legitimate otherwise the node is malicious.

4.7 Response Module

Four actions are done when a malicious node detected:

- Delete the malicious from the routing table.
- Add the malicious node to a blacklist.
- Any route reply message comes from any node in the blacklist considers a fake route reply message and the source nod will not consider it.
- Send a message to the other nodes in the network to inform them about the malicious node.

5 Performance Evaluation

This section describes simulation methodology, network simulation configurations, and performance metrics.

5.1 Simulation Methodology

The network simulated in four situations.Situation (1): The network without the presence of malicious nodes, Situation (2): The network with the presence of only black-Hole node, Situation (3): The network with the presence black-Hole node and gray-Hole nodes, Situation (4): The network with the presence of only black-Hole node and the IDS. Situation (5): The network with the presence of black-Hole node and gray-Hole nodes and the IDS. Each situation simulated with 1 m/s speed mobility and 20 m/s speed mobility

5.2 Network Simulation Configurations

The Network Simulation Configurations presented in Table 1.

Table 1: The network simulation configurations

Parameter Network Parameter	Value
Number of nodes	75 nodes
Coverage area	800×800 m
Transport layer	UDP protocol
Packet length	512 bytes
Send interval	0.025s
Mobility type	Random WP
Application layer for source nodes	UDP Basic Burst
No. of sources	2 to 12
Application layer for the other nodes	UDP Sink
Mac type	IEEE 802.11
Routing protocol	AODV
No. of black-hole attack node	1
No. of Gray-hole attack node	10
Initial position of black-Hole node	(400,400)

5.3 Performance Metrics

A number of two performance metrics used to evaluate the performance if the proposed system in the five situations, which are:

- Packet Delivery Ratio (PDR) [12] , which shows the ability to successfully deliver packets to the destination, which can be calculated by Equation (4):

$$PDR = \frac{\sum \text{No.of packets received by the destination node}}{\sum \text{no. of packets sent by source nodes}} \quad (4)$$

- Routing Overhead (ROH) [10], which shows the over heading in the routing related packets resulted by the use of the proposed IDS which can be calculated by Equation (5):

$$ROH = \frac{\sum \text{routing related packets in bytes}}{\sum \text{total routing/data transmissions in byte}} \quad (5)$$

6 Results and Discussions

In this section, the simulated result presented along with result discussion. Two different scenarios presented the first with low-speed mobility (1 m/s) and the second with high-speed mobility (20 m /s). Each scenario simulated in five situations (without attack, with a black-Hole attack only, with both black-Hole and gray-Hole attack, with black-Hole attack and IDS, with both black-Hole and gray-Hole attack and IDS).

The result of the PDR in low-speed mobility presented in Figure 7. The PDR in situation (1) is 99.74% in case of four or fewer source nodes but with an average of 84% in case of twelve or fewer source nodes, which means that the

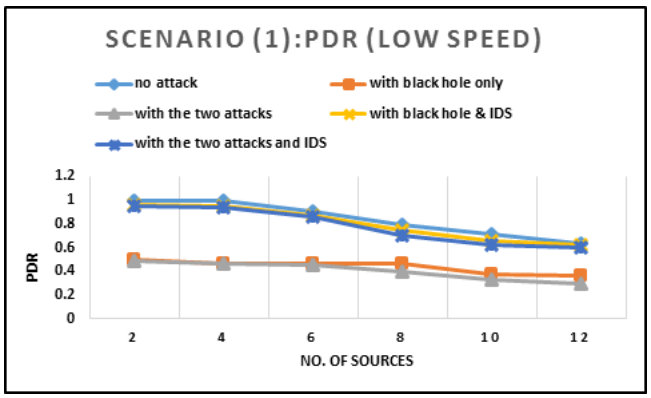

Figure 7: Packet delivery ration in Scenario (1)

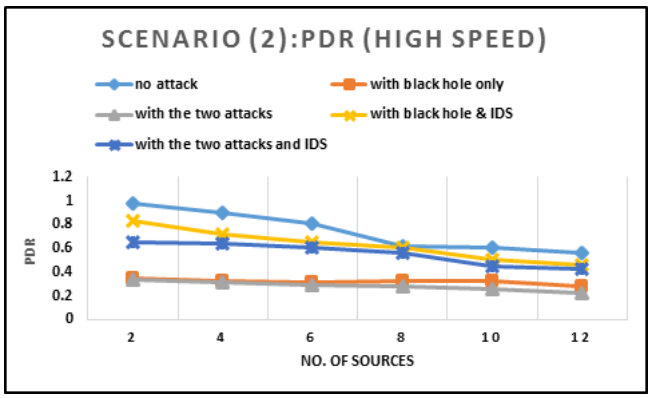

Figure 8: Packet delivery Ration in Scenario (2)

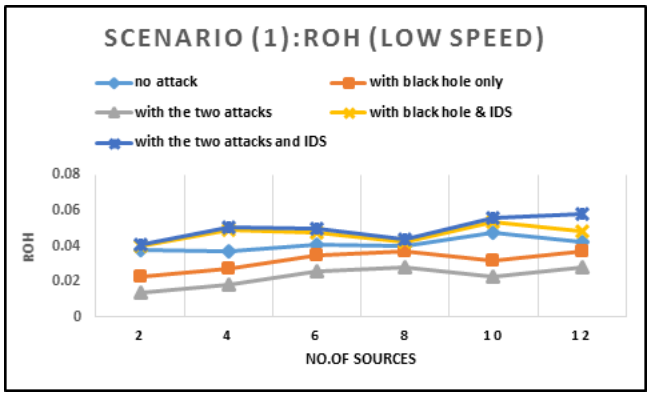

Figure 9: Routing overhead in Scenario (1)

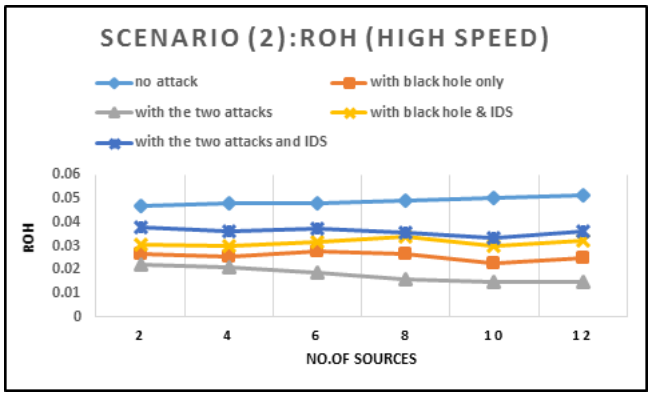

Figure 10: Routing overhead in Scenario (2)

ability to deliver packet decreases with the increase of the number of source nodes. In case situation (2) the average of PDR is 44% but situation (3) the PDR decreases to only 40% which means that the black-hole attack has the big influence on the network. However, with the use of the proposed IDS and the presence of black-Hole attack node, only the PDR increases to an average of 80% with only 4% decrease from the average percentage in case of no attacks. However, this percentage decreases to 78.1% in case of the presence of both attacks, which means that the presence of gray-Hole attack decreases the ability of the system with only 2%.

However, the increase in speed mobility changes the results completely, see Figure 8. The PDR in the case of no attacks is 74.36% in case of four or fewer source nodes but with an average of 93.63% in case of twelve or fewer source nodes. In the case of the presence of a black-Hole node in the network the average of PDR is 31.8% but in the presence of both black-Hole and gray-hole attack node the PDR decreases to only 28.3%. However, with the use of the proposed IDS and the presence of black-Hole attack node, only the PDR increases to an average of 63% with 10% decrease from the average percentage in case of no attacks. In addition, this percentage decreases to 56% in case of the presence of both attacks, which means that in the case of the two attacks and high-speed mobility the system has poor performance.

RoH is another performance evaluation, which presented in Figure 9 with low-speed mobility. The average percentage of RoH in the case of no attack is 4%, which decreases to 3.1% in case of the presences of black-Hole node and decreases again to 3.3 in case of the presence of both attacks this is because of the decrease of the route maintenance packets in the network in the absence of IDS. However, the percentage does up in case of the presences of black-Hole Node and the proposed IDS to 4.7% and to 5% in the presence of both attacks and the proposed IDS because of the use of the IDSRERR message, which is used by the detecting node to inform the other nodes about the attack.

With high speed in Scenario (2) in Figure 10, the RoH changes completely where the highest percentage value of 5% happens in the case of no attacks due to the route maintenance process. Noticing that the value decreases to 2.5% in case of the presence of black-Hole attack only and decreases again to 1.8% in case of the presence of both attacks. In addition, the value goes up again in Situation (4) to 3.1% and increases again to 3.6% when the effect of attacks increases the route maintenance packets.

7 Conclusions

In this paper, a proposed fuzzy IDS presented against both black-Hole and gray-Hole attack. This system de-

pends on an automated process to set the values of the MFs parameter in the fuzzy inference module to prevent errors from setting the parameters values manually. From the simulated results it appears that the black-Hole attack has more influence on the network than the gray-hole attack. It is proven that the network improved with an average of 36% in the presence of black-Hole attack only, and with an average of 37.8% with the presence of both attacks in case of low-speed mobility by using the proposed IDS in the PDR with an increase of 2.5% in the RoH. But in the case of high-speed mobility the network improved with an average of 31% in the presence of black-Hole attack only, and with an average of 27% with the presence of both attacks in case of high-speed mobility by using the proposed IDS in the PDR with an increase of 1.8% in the RoH.

References

[1] I. Abasikeleş-Turgut, M. N. Aydin, and K. Tohma, "A realistic modelling of the sinkhole and the black hole attacks in cluster-based WSNs," *International Journal of Electronics and Electrical Engineering*, vol. 4, no. 1, pp. 74–78, Feb. 2016.

[2] Y. Altman and A. Y. Keren, *System and Method for Automated Configuration of Intrusion Detection Systems*, US Patent 9,479,523, Oct. 25 2016.

[3] K. Balakrishnan, J. Deng, and V. K. Varshney, "Twoack: Preventing selfishness in mobile ad hoc networks," in *IEEE Wireless Communications and Networking Conference*, vol. 4, pp. 2137–2142, 2005.

[4] E. V. Balan, M. K. Priyan, C. Gokulnath, and G. U. Devi, "Fuzzy based intrusion detection systems in MANET," *Procedia Computer Science*, vol. 50, pp. 109–114, 2015.

[5] G. Banerjee, A. Kumari, A. Thakur, K. Kumari, and J. Parit, "An analysis on characteristics, challenging issues and comparisons of routing protocols of MANET," *International Journal of Scientific Research in Science, Engineering and Technology*, vol. 2, no. 2, pp. 876–879, 2016.

[6] A. O. Bang and P. L. Ramteke, "MANET: History, challenges and applications," *International Journal of Application or Innovation in Engineering & Management*, vol. 2, no. 9, pp. 249–251, 2013.

[7] S. Brar and M. Angurala, "Review on grey-hole attack detection and prevention," *International Journal of Advance research , Ideas and Innovations in Technology*, vol. 2, no. 5, pp. 1–4, 2016.

[8] A. Chaudhary, V. Tiwari, and A. Kumar, "Analysis of fuzzy logic based intrusion detection systems in mobile ad hoc networks," *International Journal of Information Technology*, vol. 6, no. 1, pp. 690–696, 2014.

[9] D. Dave and P. Dave, "An effective black hole attack detection mechanism using permutation based acknowledgement in MANET," in *International Conference on Advances in Computing, Communications and Informatics (ICACCI,14)*, pp. 1690–1696, 2014.

[10] S. Goswami, S. Joardar, C. B. Das, S. Kar, and D. K. Pal, *Performance Analysis of Three Routing Protocols in MANET Using the NS-2 and ANOVA Test with Varying Speed of Nodes*, Ad Hoc Networks, 2017.

[11] M. Gupta and K. K. Joshi, "A review on detection and prevention of gray-hole attack in MANETs," *International Journal of Scientific & Engineering*, vol. 4, no. 11, 2013.

[12] N. Kaur, "Implementing MANET security using CBDS for combating sleep deprivation & DOS attack," *International Journal for Science and Emerging*, vol. 16, no. 1, pp. 6–12, 2014.

[13] V. Khandelwal and D. Goyal, "Blackhole attack and detection method for AODV routing protocol in MANETs," *International Journal of Advanced Research in Computer Engineering & Technology*, vol. 2, no. 4, pp. pp. 1555–1559, 2013.

[14] T. Link and B. D. Gadong, "Performance analysis of MANET under black hole attack using AODV, olsr and tora," in *Computational Intelligence in Information Systems: Proceedings of the Computational Intelligence in Information Systems Conference (CIIS'16)*, vol. 532, Springer, pp. 198, 2016.

[15] S. Marti, T. J. Giuli, K. Lai, and M. Baker, "Mitigating routing misbehavior in mobile ad hoc networks," in *Proceedings of ACM the 6th Annual International Conference on Mobile Computing and Networking*, pp. 255–265, 2000.

[16] D. J. Norris, "Fuzzy logic system," in *Beginning Artificial Intelligence with the Raspberry Pi*, pp. 111–143, 2017.

[17] J. Ramkumar and R. Murugeswari, "Fuzzy logic approach for detecting black hole attack in hybrid wireless mesh network," in *2014 IEEE International Conference on Innovations in Engineering and Technology (ICIET'14)*, vol. 3, pp. 877–882, 2014.

[18] D. Sabarish and C. Ranjani, "Enhanced DSR protocol for detection and exclusion of selective black hole attack in MANET," *International Journal of Computer Applications*, vol. 112, no. 14, 2015.

[19] S. T. P. Saurabh, "A PDRR based detection technique for blackhole attack in MANET," *International Journal of Computer Science and Information Technologies*, vol. 2, no. 4, pp. 1513–1516, 2011.

[20] M. Sengar, P. P. Singh, and S. Shiwani, "Detection of black hole attack in MANET using fbc technique," *International Journal of Emerging Trends & Technology in Computer Science*, vol. 2, no. 2, pp. 269–272, 2013.

[21] G. Sharma and M. Gupta, "Black hole detection in MANET using AODV routing protocol," *International Journal of Soft Computing and Engineering*, vol. 2, 2012.

[22] K. Shimojima, T. Fukuda, and Y. Hasegawa, "Self-tuning fuzzy modeling with adaptive membership

function, rules, and hierarchical structure based on genetic algorithm," *Fuzzy Sets and Systems*, vol. 71, no. 3, pp. 295–309, 1995.

[23] J. Singh, "Fuzzy logic based intrusion detection system against blackhole attack on AODV in MANET," *IJCA Special Issue on Network Security and Cryptography*, pp. 28–35, 2011.

[24] M. M. Singh, A. Singh, and J. K. Mandal, "A snapshot of black hole attack detection in MANET," *International Journal of Computer Applications*, vol. 116, no. 14, 2015.

[25] R. Singh and D. Kumar, "MANET: Security issues and behavior analysis of routing protocol using NS-2," *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 5, no. 4, 2015.

[26] P. Tahmasebi and A. Hezarkhani, "A hybrid neural networks-fuzzy logic-genetic algorithm for grade estimation," *Computers & Geosciences*, vol. 42, pp. 18–27, 2012.

[27] K.-S. Tang, K.-F. Man, Z.-F. Liu, and S. Kwong, "Minimal fuzzy memberships and rules using hierarchical genetic algorithms," *IEEE Transactions on Industrial Electronics*, vol. 45, no. 1, pp. 162–169, 1998.

[28] D. Vydeki and R. S. Bhuvaneswaran, "Effect of clustering in designing a fuzzy based hybrid intrusion detection system for mobile ad hoc networks," *Journal of Computer Science*, vol. 9, no. 4, pp. 521–525, 2013.

[29] M. Wahengbam and N. Marchang, "Intrusion detection in MANET using fuzzy logic," in *3rd national conference on Emerging trends and applications in computer science (NCETACS'12)*, pp. 189–192, 2012.

Biography

Mohamed Abdel-Azim received the Ph.D. degree in Electronics and Communications Engineering from the Faculty of Engineering-Mansoura University-Egypt by 2006. After that he worked as an assistant professor at the electronics & communications engineering department, then awarded the associate professor degree in 2012 until now. He has 130 publications in various international journals and conferences. His current research interests are in multimedia processing, wireless communication systems, and field programmable gate array (FPGA) applications. He had awarded the best Ph.D. thesis at Mansoura University at 2007. He is the executive director of scientific computing center and the consultant for IT in Mansoura University.

Hossam El-Din Salah, associated professor at electronics and communications department -Faculty of Engineering - Mansoura University, BSc of electronics from Faculty of Engineering - Mansoura University 1993, MSc of electrical communications engineering Faculty of Engineering - Mansoura University - 2000, PhD of electrical communications engineering Faculty of Engineering - Mansoura University - 2008.

Menas Ebrahim received the Master degree in Electronics and Communications Engineering from the Faculty of Engineering-Mansoura University-Egypt by 2017. She worked as a Teaching assistant at the electronics& communications engineering department in MISR Engineering & Technology higher institute since 2012.

Network Security Situation Awareness Based on the Optimized Dynamic Wavelet Neural Network

Huang Cong[1] and Wang Chao[2]
(Corresponding author: Wang Chao)

China Tobacco Guangxi Industry Co., Ltd, China[1]
No. 28 Research Institute, China Electronics Technology Group Corporation, China[2]
No. 1, Alfalfa East Street, Qinhuai District, Nanjing 210007, China
(Email: diguatongxue@163.com)
(Received Mar. 28, 2017; revised and accepted June 26, 2017)

Abstract

In order to analyze the evolvement trend of the network threat and to explore the self-perception and control problem of the security situation, the dynamic wavelet neural network model is integrated into the model design, and a kind of network security situation awareness based on the optimized dynamic wavelet neural network is put forward, so as to enhance the interaction and cognitive ability between the layers of the network security system. On the basis of the analysis of the model components and their functions, the dynamic wavelet neural network algorithm is applied to obtain the accurate decision of the heterogeneous sensors for the network security events. Combined with the deduction of the relationship between the threat grade and threat genes, the shortcoming of the necessity to handle the complicated relationships among network components during the process to obtain the threat genes is overcome, and the hierarchical situation awareness method including the service level and network level is proposed to improve the expressiveness for the network threats. The simulation results show that: The network security situation awareness and method based on the optimized dynamic wavelet neural network can integrate the heterogeneous security data with dynamic perception on the evolution trend of the threat, and have the ability of self-regulation and control to certain extent, which has achieved the goal of situation awareness, and provided new methods and means for the supervision and management of network.

Keywords: Awareness; Multi-source Fusion; Network Security Situation; Wavelet Neural Network

1 Introduction

Nowadays, the role of the network and information technology is becoming more and more important in the field of economy, society and national defense. As a result, it has risen to the height of the interests of the state and all the people. It has gradually become an important factor in the economic development and national strategic deployment. However, the heterogeneous heterogeneity and complexity of the current network system, the continuous deterioration of the utilization environment, the continuous expansion of the scale and the emergence of a variety of emerging applications, the traditional "lugging holes, building high walls and anti attack" security model is lack of self-adaptability, unified scheduling and effective coordination, resulting in the widespread network intrusion and sabotage, which has led to major economic losses, adverse social impact, and even major fatal crash and casualty accident due to serious dereliction of duty.

Network security situation awareness (NSSA) is considered as a new approach to solve the problems in the field of network security [15, 21]. It can fuse the security events detected by network components and real-time perceive the network security situation and the risks faced, and has become a hot research field with the cutting-edge international and cross-disciplinary nature. In 1999, Bass *et al.* [1] proposed a multi-sensor fusion based on the situation awareness model, and the fusion model became representative at this stage of research, including the three level model composed of element extraction, state perception and situation prediction proposed by Tadda *et al.* [20], and the network situation fusion perception and risk awareness model proposed by Shen *et al.* [17]. At the same time, the visualization technology is also an significant branch of the initial phase of NSSA research.

Lawrence Berkeley National Laboratory, the National Center for Advanced Security Systems Research of the United States and other foreign military departments and research centers have developed "Spinning Cube Potential Doom" [18], NVisionIP [25] and other visual situation awareness software, and even until now, the trend of visualization for situation is still a major research direction [19]. In the period of NSSA research, the framework model and the visualization tools are developed. How-

ever, the methods and mechanisms involved in the NSSA are still validated. In the large-scale network, the visualization of network traffic and connectivity is also important. It is difficult to accurately obtain network security situation. In 2006, Chen *et al.* [3] proposed a hierarchical awareness method of network security threat situation quantification. Although the threat weight in this work still depended on the expert experience and fusion perception, the proposed hierarchical threat awareness concept had important impetus influence on the research of situation awareness.

At this stage, there was vigorous development in the research direction of the analytic hierarchy process (AHP) according to Hu *et al.* [13]. However, some problems, including the incomplete knowledge of situation knowledge, large subjective dependence of situation threat genes, and difficulties in obtaining such methods are still existed. Since 2008, fusion perception has become a hot topic in NSSA research field, and fusion algorithm is one of the core contents of the research process. Due to the randomness and suddenness of network security events, it is difficult to obtain the prior and conditional probabilities, moreover, it is hard to deal with the uncertainties in the fusion process.

D-S evidence theory meets the demand of multi-source fusion, and the demand of data traffic is small. The reasoning process has low requirement to prior probabilities and good adaptabilities in dealing with uncertainty. The fusion-perception method based on D-S linear weighting is introduced into D-S evidence by Wei *et al.* [23], while Zhang *et al.* [26] adopted the average method to improve D-S merge rule to deal with NSSA fusion perception problem. Research on NSSA fusion perception confirms the feasibility of fusion perception. But, there are still some problems such as the non-normalization and fusion conflict in the process of D-S evidence fusion.

Since 2010, the cognitive ability and feedback control of situation awareness have received a great attraction from network security researchers, and there have been many representative research results. For example, according to [5], author thought cognitive perception is an important challenge in the field of information fusion, and has discussed the formal theory basis of cognitive situation awareness (like dependency theory and extended constructive function). Gong *et al.* [11] emphasized the feedback control structure of NSSA research in the proposed framework of network situation, and argued that extended control cycle model (OODA) provided a data fusion mechanism to deal with multiple concurrency and latent interaction. Zhang *et al.* [24] constructed an NSSA model based on Markov game. Although the core of the research was to build a tripartite game model by using risk communication networks, the concept of security system reinforcement emphasized not only the realization of threat situation, but also the control of system state.

Neural network is considered to be a new mechanism to solve the problem in NSSA. In the literature, there are several researches on cognitivering, cross-layer structure and self-adaptability taking into account. Thomas *et al.* [8] agreed that neural networks should employ designs similar to those of the OODA ring, enhancing the cognitive ability of the system. Cross-layer design is another research hotspot of neural network, and also a widely accepted structural form in academia. Clark *et al.* [4] and Shakkottai *et al.* [16] considered that the cross-layer structure was the basis of neural networks and could overcome the shortcomings of the traditional network level information interaction difficulties, which had been initially applied and tested in the wireless network channel management [7], self-interference [6], path selection [2] and other studies. unfortunately there were still the shortcoming only for the specific network level, overlap optimization and other issues.

The autonomic and dynamic configuration ability of neural networks are also considered as a feasible way to realize adaptive system. Gomez *et al.* [10] proposed a neural network framework abstraction layer and operating environment (ALOE), which can provide dynamic configuration, resource awareness and operation control for the real-time system platform. Gupta *et al.* [12] and Ogiela *et al.* [14] thought that neural networks have the ability to reason and perceive autonomously, and can simulate cognitive functions,including learning, memory, reasoning and perception.

Furthermore, they can be applied to security, information system decision-making and many other fields. The above mentioned studies have provided feasible theoretical basis for the realization of intelligent systems with autonomous characteristics, perception and learning ability via using neural networks, while foreign military and research institutions start their own research program from the dynamic configuration of neural network capacity, for example, National Natural Science Foundation of the United States, DARPA and NASA funded the BNA (bio-networking architecture project) [9], Tbatou [22] and other research, also, the European Union launched the seventh Framework Program (FP7).

Such research programs have validated that neural networks can build adaptive systems in a systematic and empirical manner and can provide a new approach to self-management. According to the development of NSSA, NSSA has transitioned from perception network to perception control network. In this paper, based on the existing research results, the wavelet neural network is integrated into the research of the security situation awareness, and a model of network security situation awareness is proposed. The research is applicable to the dynamic wavelet neural network algorithm in the heterogeneous sensor environment, and the dynamic awareness is achieved for the perception of external environmental information, controlling of internal operating state, and the establishment of the bridge between the discrete control and continuous control, so as to achieve the purpose of situation awareness.

2 Dynamic Wavelet Neural Network Security Situation Awareness

Through the cross-layer optimization and integration of CPSO-DS, the data accuracy, consistency and other issues have been solved. And the situation awareness is the dynamic evolution view of the network system generated on the basis of the fusion results, with the situation factor extraction and hierarchical threat awareness as the key issue.

2.1 Situation Elements Extraction

For a successful NSSA system, effective perception depends on accurate factor extraction. The composition of the situation elements should include all the key factors in the network that can cause changes in the situation, such as the attack threat gene (threat level), attack intensity (event frequency), asset importance, and so on. Among them, the importance of assets and attack strength are attributes that are easy to be obtained, but the threat gene generation is the current difficult content in the study. From the current research status, empirical recursive and AHP analytic methods are the most successful research methods in threat gene research. However, the experience recursive method is more dependent on expert experience, subjectivity is strong and difficult to obtain. AHP method needs to make a complex deduction to the influence of the problem component and subordinate relationship.

In this work, the wavelet neural network is introduced into NSSA and its fitting with NSSA is studied. Assuming that there are n targets in the network environment, it is required assign the threat genes. The decision-making target can obtain m threat genes for n different types of events. Each threat gene is treated as a random variable x_i with value taken as 1 and -1, respectively. The purpose is to make the random variable satisfy a distribution with the mean 0. Let $X_n = \sum_{i=1}^{n} x_i$, $Y = \frac{X_n}{\sqrt{n}}$ (X, n), when $n \to \infty$, Y obeys a normal distribution, then X_n gradually obeys the normal distribution $N(0, n)$, and then the horizontal ordinate of the normal distribution curve is the importance of the decision-making target (the size of the threat gene), and the vertical ordinate is the number of decision targets (sorted by the threat level from highest to lowest). According to the characteristics of the normal distribution, the threat gene pattern can be described as follows: The greater the impact on the network security situation, the threat gene of the event closer to the first quadrant of the first position; the other hand, the threat gene in the second quadrant more left Position, as shown in Figure 1. The following will simplify the acquisition of the threat gene by reasoning, so that the threat gene can be easily calculated only at the known threat level.

Perform equidistant partition on the normal distribution curve vertical axis with as the scale factor, as presented in Figure 2, making transformation on the curve in the second quadrant with line $f(x) = A$ as the axis of symmetry, and moving the vertical axis to the left by 3σ.

$$y = \begin{cases} 3\sigma + \sqrt{-2\sigma^2 \ln[\sigma\sqrt{2\pi}x]}, 0 < x < \frac{1}{\sqrt{2x}} \\ 3\sigma, x = \frac{1}{\sqrt{2\pi}} \\ 3\sigma - \sqrt{-2\sigma^2 \ln[\sigma\sqrt{2\pi}(\frac{2}{\sigma\sqrt{2\pi}} - x)]}, \frac{1}{\sigma\sqrt{2\pi}} < x < \frac{2}{\sigma\sqrt{2\pi}} \end{cases} \quad (1)$$

From Equation (1), it can see that the target range is $(0, \frac{2}{\sigma\sqrt{2\pi}})$. Divide it into n equal parts $(x_i = \frac{i}{n} \times \frac{2}{\sigma\sqrt{2\pi}}$, $1 \le i \le n)$, and introduce into Equation (1), then y_i corresponding to x is the threat gene of the queue level i. The maximum threat gene is $ny_{max} \approx 6\sigma$, then the i-th threat gene (G_i) can be quantified as

$$G_i = \frac{y_i}{y_{max}} \begin{cases} \frac{1}{2} + \frac{\sqrt{-2\ln\frac{2i}{n}}}{6}, 1 \le i < \frac{n}{2} \\ \frac{1}{2}, i = \frac{n}{2} \\ \frac{1}{2} - \frac{\sqrt{-2\ln[2-\frac{2i}{n}]}}{6}, \frac{n}{2} < i < n. \end{cases} \quad (2)$$

At this point, the threat gene can be obtained only by knowing the different threat types (n), and ranking the threat degree of each type of threat to the network (i), to obtain the threat gene of the i-th level event. When NSSA is applied to networks with different attack susceptibilities, it is only necessary to rearrange the rank of the threat type. This method can greatly reduce the complexity of the acquisition of threat gene and improve the current status of the acquisition of threat genes with strong subjectivity, high complexity, and dependence on the expert experience.

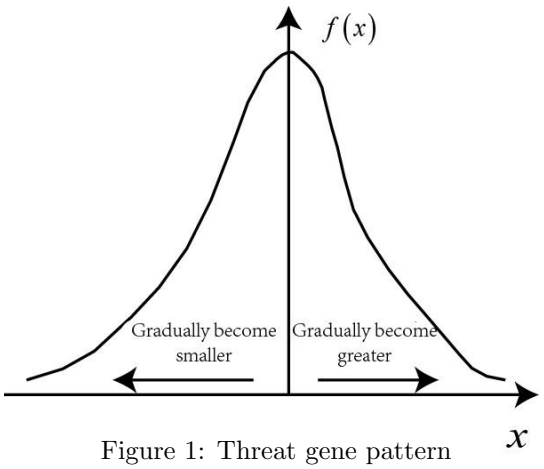

Figure 1: Threat gene pattern

2.2 Threat Quantitative Awareness

In this paper, the network security situation is divided into two different levels, including service-level and network-level. The core idea is: At two different levels, the security situation values are both based on the sensitivity of the network system for the type of attack, with the situation elements as the central point of view to achieve the hierarchical awareness.

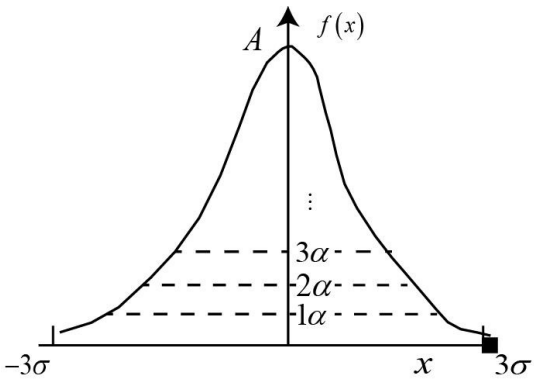

Figure 2: Threat equal interval separation

2.2.1 Service Security Situation

Definition 1. *(Threat Gene). In the time window tw, the service $s_i(0 \le i \le u)$ is subjected to n different types of attacks $a_{ij}(0 \le j \le n)$. According to the degree of threat to the service s_i, n attacks can be divided into $g(1 \le g \le n)$ different threat levels (a variety of different types of attacks can fall under the same threat level), then the threat gene of grade k is*

$$l_k = \begin{cases} \frac{1}{2} + \frac{\sqrt{-2\ln\frac{2k}{n}}}{6}, 1 \le k < \frac{n}{2} \\ \frac{1}{2}, i = \frac{n}{2} \\ \frac{1}{2} - \frac{\sqrt{-2\ln[2-\frac{2k}{n}]}}{6}, \frac{n}{2} < k < n. \end{cases} \quad (3)$$

According to the different types of attacks, from the Definition 1 the degree of threat to attack the service quantization weight can be determined. Additionally, the service security situation is also related to the attack strength, and Definition 2 can be obtained accordingly.

Definition 2. *(Service Security Situation). Under the premise of Definition 1, the threat gene quantization weight of g different attacks is $l_k(1 \le k \le g)$; service s is subject to a total N_i number of various types of attacks, in which, the number of type $j(0 \le j \le g)$ attack is denoted as N_{ij}, known as the attack strength, and $N_i = \sum_{j=0}^{s} N_{ij}$ is satisfied. Then the security situation of service $s_i(0 \le i \le u)$ is $N_i = \sum_{j=0}^{g} N_{ij}$.*

In which, u is the number of the services. The purpose of Equation (4) using 10^{l_k} is to emphasize the importance of the threat genes, and weaken the impact of attack strength on the service security situation.

$$V_{S_i} = \sum_{k=1}^{g} N_{ik} 10^{l_k} \quad (4)$$

2.2.2 Network Security Situation

Network security situation is composed of the host security situation within the time window *tw* and the number of host and so on.

Definition 3. *(Network Security Situation). In the time window tw, there are v hosts in the network system NS, where in the importance degree of the host $H_i(1 \le i \le v)$ is $g_{H_i}(1 \le i \le v)$, then the network security situation is*

$$V_{NS} = \sum_{i=1}^{v} (V_{H_i} g_{H_i}) \quad (5)$$

In which, the host weight is determined by the number of the key services running on the host, the asset value, and whether there is the presence or confidential data as $(t_{H1}, t_{H2}, \cdots, t_{Hv})$, on conduct normalization to obtain the host weighting:

$$g_{H_i} = \frac{t_{H_i}}{\sum_{j=1}^{v} t_{H_j}} \quad (6)$$

In the network system, the greater the value of VNS is, it indicates that the more serious threat that network system is facing; on the other hand, the network system is relatively safe. In contrast to intrusion detection, security situation awareness techniques map discrete alarm events into continuous security situation evolution curves, and visually express the threats and evolving trends of current network systems. Combined with the visualization technology, it can generate hierarchical, multi-dimensional dynamic evolution view, intuitive perceptual service, host and network and so on, which have provided new methods and means for the monitoring and management of the network, and can also be applied to construct the multi-dimensional dynamic evolution view, as well as provide the reference for safety control.

3 Simulation Experiment and Analysis

3.1 Fusion Ability

According to the research demand, the network topologies are designed.The three kinds of sensors, including Netflow, Snort and Snmp, are deployed to detecting data at different levels. The cross-layer heterogeneous sensor data transmission and formatting are realized by XML technology. Furthermore, the relationship between the three sensors and cross-layer perception ring components, as well as the structure of the Snmp sensor design are shown in Figure 3.

The training set and test set select 20% and 9% of the 10% data set of DARPA 99 intrusion detection data (See Table 1). The data selection process is based on the proportion of traffic in the real network and Netpoke, so as to achieve the greatest degree of simulation Internet. Adopting the results of three kinds of sensors to conduct several rounds of training on the CPSO-DS fusion engine with the population size of 55, and search the optimization weights in [0, 1]. The influence of noise on the weight optimization is reduced by the combination of the offline

optimization and on-line adjustment. And the expected variance and Netflow port change rate and flow access ratio is adopted to obtain the heterogeneous sensor BPA, as listed in Table 1.

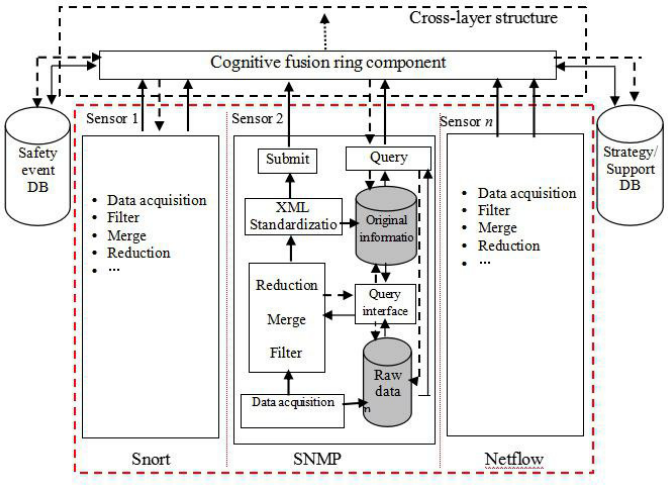

Figure 3: Relations between sensors and cognitive circle components

According to the network topology, the CPSO-DS is applied to fuse the alarms generated by the replay test set and fuse with the un-weighted traditional DS and the empirical weight DS. The PSO-DS of the two data sources fusion is compared at the detection rate (DR) and false discovery rate (FDR) aspects, as shown in Table 2.

The experimental results show that CPSO-DS multi-source fusion is superior to the other methods in detection rate and false alarm rate. In addition, compared with the two sensor research in literature [9], the increase of sensor number can improve the detection rate and reduce false alarm rate. During the experiment, from the point of view of U2R and R2L, it is more difficult to improve the detection efficiency simply by increasing the number of sensors. More host-based sensors should be designed to improve their detection capability. Besides, from the two data sources and three data sources on the performance comparison, an increasing of the data source can improve the accuracy, but sometimes can not significantly enhance the accuracy of fusion, in other words, for multi-source integration, it is not necessary better if there are greater number of sources, and the accuracy of the fusion is closely related to the detection performance and characteristics of the added sensor itself.

3.2 Hierarchical Awareness

3.2.1 Service Security Situation

On the basis of the output of the CPSO-DS fusion engine, the threat awareness can be performed according to the steps of factor extraction, feature quantization and layer awareness. The attack method needs to quantify the attack strength, attack type and threat gene, etc. The attack strength can be determined statistically in the time window through the output of the CPSO-DS fusion engine. Attack types and their threat levels are listed in Table 1, and the threat genes (n = 5, g = 4) are calculated according to Equation (3). The simulation network runs for a total of one week (From June 10, 2013 to June 16, 2013), the attacker consists of two terminals, which autonomously arrange the attack time, the defense side has no knowledge of the hacking simulation behavior, the test set attack data is selectively replayed to the local area network and the dynamic evolution curve of the security situation of a certain service is obtained according to Equation (4). Figure 4 shows the evolution of security situation for the running three services of Email, Http and Snmp on host H1. The abscissa is time, the length is 7 days, time window is 2 hours; and vertical ordinate is the service security situation value, which expresses the threat level of the attack to the service.

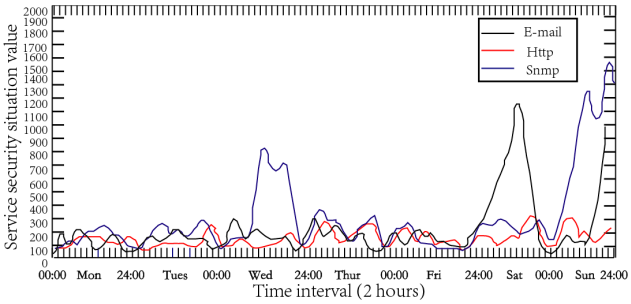

Figure 4: Service security situation after self-adapting

As can be seen in Figure 4, both Http and Snmp services are heavily attacked on Sunday, on Wednesday afternoon and at night, Http's security situation should also be concerned, and Snmp is threatened late Saturday night; email service is stable during the whole monitoring period. According to the service security wavelet network curve, the network analyst should strengthen the supervision of the managed service in the time of severe threat and make further inspection on the vulnerability and configuration, etc. From the service security situation view can also be seen in a certain period of time, security threats there is a gradual and serious regularity, the administrator should be based on service security trends and trends in advance to take appropriate measures. Owing to the space limitation, this work only shows the E-mail, Http and Snmp security situation on host H1, and will not elaborate the host H2 and H3 service security situation evolution view.

3.2.2 Network Security Situation

The perception of network security situation requires to determine the weight of the importance of the host, but the determination of the host weight is more complicated than the service weight, which is related to the host asset value (V_h), service criticality (C_s), access frequency level (A_f) and confidentiality (D_c) and other factors, the im-

Table 1: Basic experiment data

Type	Train. Set	Experi. Set	$BPA_{Netflow}$	BPA_{Snort}	BPA_{Snmp}	$w_{Netflow}$	w_{Snort}	w_{Snmp}	Threat Grade	Threat Gene
R2L	226	98	0.098	0.194	0.347	0.26	0.71	0.67	1	0.726
U2R	31	11	0.146	0.203	0.261	0.23	0.91	0.69	1	0.726
DoS	78291	33665	0.283	0.189	0.167	0.93	0.34	0.22	2	0.611
Probe	822	354	0.367	0.288	0.188	0.88	0.60	0.41	3	0.389
New	*	*	0.106	0.126	0.037	0.58	0.71	0.43	0	1

Table 2: Fusion ability

Parameter	Traditional D-S(%)	Empirical Weighted D-S(%)	PSO-DS (Two Data Sources) (%)	CPSO-DS (Three Data Sources) (%)
DR	73.33	82.60	86.67	88.11
FDR	9.86	5.80	5.63	5.06

portance level is shown in Table 3, in which, the host complex importance $t_{H1} = k_V V_{hi} + k_C C_{si} + k_A A_{fi} + k_D D_{ci}$, $k_V = 0.2$, $k_C = 0.3$, $k_A = k_D = 0.25$. Based on the host security situation, we can obtain the composite importance weight of the host using Table 3, and use Equation (5) to generate the security situation evolution view of the whole network system, as depicted in Figure 5.

Table 3: Host weight grade

Host	V_{hi}	C_{si}	A_{fi}	D_{ci}
H_1	High	Medium	Medium	High
H_2	Medium	Low	Medium	Medium
H_3	High	Low	Low	Low

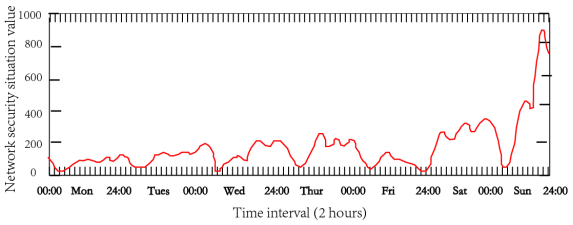

Figure 5: Network security situation

The evolution of the entire network's security situation over the course of a week can be seen from Figure 5 that On Wednesday, Thursday, and Saturday, there was a concern that have been caused by a hacker attempting to attack the network, although the need for individual hosts and services for security maintenance, but still did not have a huge impact on the entire network. Network system in June 16, 2013 appeared very serious attack situation, requiring administrators to focus on monitoring and taking appropriate action, so as to avoid the entire network system inefficient and even the collapse of the situation.

In addition to dynamically assessing the threat situation of services, hosts and networks, the hierarchical awareness method proposed in this paper has good environmental adaptability and shows certain cognitive ability. In this article, based on the weight coefficient of the threat gene acquisition method, the situation awareness system is applied to the new network, and the administrator only needs to re-sort the grades of the attacks that the network is sensitive to. Assuming a new network environment, the sensitivity of the attack on the services in order of unknown attacks, DoS, U2R and R2L and Probe, the threat genes in Table 1 in accordance with the Equation (3) can be adjusted, as shown in Table 4, and according to Table 1, Table 4, Equation (4) and (5), the security wavelet network curve at service and network levels can be generated, and it is not required to make complicated association analysis on components and elements of composition situation. Under the same attack, as plotted in Figure 4, the security situation evolution of E-mail, Http, and Snmp on the host H1 after the adaptation of the threat genes is shown in Figure 6. Similarly the security situation of the monitored service can also be perceived and shows a trend similar to the situation evolvement trend as Figure 4, but the same attack data on different networks will usually show different degree of threat.

Table 4: New threat gene

Attack Type	Threat Grade	Threat Gene
R2L	2	0.611
U2R	2	0.611
DoS	1	0.726
Probe	3	0.389
New	0	1

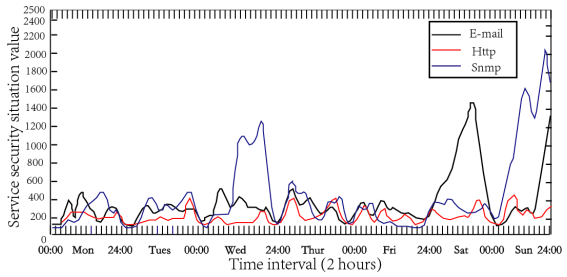

Figure 6: Service security situation after self-adapting

4 Conclusions

In this paper, existing models, fusion algorithms and perceptual methods of the network security awareness have been analyzed. Also, a network security situational awareness and control fusion model is put forward and discussed. Under the guidance of the model, CPSO-DS dynamic wavelet neural network algorithm that is applicable to the heterogeneous network is discussed, on the basis of the acquisition of the threat genes, comprehensive analysis on the two different levels of services and network security situation awareness methods are conducted, and the situation gradient is applied to achieve the self-regulation of the network security situation. Simulation results reveal that the network security situation awareness and control model based on the optimized dynamic wavelet neural network and its method can accurately identify the network security events and dynamically perceive the threat evolution trend at different levels.

References

[1] T. Bass, "Multi-Sensor data fusion for next generation distributed intrusion detection systems," in *Proceeding of the IRIS National Symposium on Sensor and Data Fusion*, vol. 4, pp. 24-27, 1999.

[2] J. S. Chen, C. Y. Yang and M. S. Hwang, "The Capacity Analysis in the Secure Cooperative Communication System," *International Journal of Network Security*, vol. 19, No. 6, pp. 863-869, 2017.

[3] Y. Chen and J. Chou, "On the privacy of user efficient recoverable off-line e-cash scheme with fast anonymity revoking," *International Journal of Network Security*, vol. 17, no. 6, pp. 708-711, 2015.

[4] D. D. Clark, C. Partridge and J. C. Ramming, "A knowledge plane for the internet," in *Proceeding of the 2003 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications (SIGCOMM'03)*, pp. 3-10, 2003.

[5] R. Dapoigny and P. Barlatier, "Formal foundations for situation awareness based on dependent type theory," *Information Fusion*, vol. 14, no. 1, pp. 87-107, 2013.

[6] A. R. Emilio, W. Stefan, L. V. Roberto, R. Taneli and W. Risto, "Wideband full-duplex MIMO relays with blind adaptive self-interference cancellation," *Signal Processing*, vol. 130, pp. 74-85, 2016.

[7] M. Erdelj, M. Krl and E. Natalizio, "Wireless Sensor Networks and Multi-UAV systems for natural disaster management," *Computer Networks*, vol. 124, pp. 72-86, 2017.

[8] C. Fortuna and M. Mohorcic, "Trends in the development of communication networks: Cognitive networks," *Computer Networks*, vol. 53, no. 9, pp. 1354-1376, 2009.

[9] P. Gandotra, R. K. Jha and S. Jain, "A survey on device-to-device (D2D) communication: Architecture and security issues," *Journal of Network and Computer Applications*, vol. 78, pp. 9-29, 2017.

[10] I. Gomez, V. Marojevic and A. Gelonch, "ALOE: An open-source SDR execution environment with cognitive computing resource management capabilities," *IEEE Communications Magazine*, vol. 49, no. 9, pp. 76-83, 2011.

[11] Z. H. Gong and Y. Zhuo, "Research on cyberspace situational awareness," *Journal of Software*, vol. 21, no. 7, pp. 1605-1619, 2010.

[12] M. Gupta, "On fuzzy logic and cognitive computing: some perspectives," *Scientia Iranica*, vol. 18, no. 3, pp. 590-592, 2011.

[13] W. Hu, J. Li and X. Jiang, "A hierarchical algorithm for cyberspace situational awareness based on analytic hierarchy process," *High Technology Letters*, vol. 13, no. 3, pp. 291-296, 2007.

[14] M. R. Ogiela and I. You, "Cognitive and secure computing in information management," *International Journal of Information Management*, vol. 33, no. 2, pp. 243-244, 2013.

[15] E. U. Opara, O. A. Soluade, "Straddling the next cyber frontier: The empirical analysis on network security, exploits, and vulnerabilities," *International Journal of Electronics and Information Engineering*, vol. 3, no. 1, pp. 10–18, 2015.

[16] S. Shakkottai, T. Rappaport and P. Karlsson, "Cross-Layer design for wireless networks," *IEEE Communications Magazine*, vol. 41, no. 10, pp. 74-80, 2003.

[17] D. Shen, G. Chen, J. B. Cruz, J. L. Haynes, M. Kruger and E. Blasch, "A Markov game theoretic approach for cyber situational awareness," in *Proceeding of the Multi-Sensor, Multi-Source Information Fusion: Architectures, Algorithms, and Applications*, LNCS 6571, pp. 1-11, Springer, 2007.

[18] Y. Shi, R. Li, Y. Zhang and X. Peng, "An immunity-based time series prediction approach and its application for network security situation," *Intelligent Service Robotics*, vol. 8, no. 1, pp. 1-22, 2015.

[19] H. Shiravi, A. Shiravi and A. A. Ghorbani, "A survey of visualization systems for network security," *IEEE Transactions on Visualization andComputer Graphics*, vol. 18, no. 8, pp. 1313-1329, 2012.

[20] G. Tadda, J. J. Salerno, D. Boulware, M. Hinman and S. Gorton, "Realizing situation awareness within

a cyber environment," in *Proceeding of the Multi-Sensor, Multi-Source Information Fusion: Architecture, Algorithms, and Applications*, LNCS 6242, pp. 1-6, Springer, 2006.

[21] A. Tayal, N. Mishra and S. Sharma, "Active monitoring & postmortem forensic analysis of network threats: A survey," *International Journal of Electronics and Information Engineering*, vol. 6, no. 1, pp. 49–59, 2017.

[22] Z. Tbatou, A. Asimi, Y. Asimi, Y. Sadqi and A. Guezzaz, "A New Mutuel Kerberos Authentication Protocol for Distributed Systems," *International Journal of Network Security*, vol. 19, no. 6, pp. 889-898, 2017.

[23] Y. Wei, Y. F. Lian and D. G. Feng, "A network security situational awareness model based on information fusion," *Journal of Computer Research and Development*, vol. 46, no. 3, pp. 353-362, 2009.

[24] Y. Zhang, X. B. Tan, X. L. Cui and H. S. Xi, "Network security situation awareness approach based on Markov game model," *Journal of Software*, vol. 22, no. 3, pp. 495-508, 2011.

[25] Y. Zhao, P. C. Fan, H. T. Cai, Z. G. Qin and H. Xiong, "Attribute-based encryption with non-monotonic access structures supporting fine-grained attribute revocation in m-healthcare," *International Journal of Network Security*, vol. 19, no. 6, pp. 1044-1052, 2017.

[26] Y. Zhang, S. G. Huang, S. Z. Guo and J. M. Zhu, "Multi-Sensor data fusion for cyber security situation awareness," *Procedia Environmental Sciences*, vol. 10, pp. 1029-1034, 2011.

Biography

Huang Cong was born in the Jiangxi of Guangxi in 1974, the Master of Business School, Guangxi University. Now he is the Engineer of China Tobacco Guangxi Industry Co., Ltd. His main research direction is the computer network security.

Wang Chao was born in Nanjing in 1989, who is the Doctor of Nanjing University of Science and Technology, and his main research direction is the computer network security. Now, he is working in No. 28 Research Institute, China Electronics Technology Group Corporation, Nanjing, China.

Guide for Authors
International Journal of Network Security

IJNS will be committed to the timely publication of very high-quality, peer-reviewed, original papers that advance the state-of-the art and applications of network security. Topics will include, but not be limited to, the following: Biometric Security, Communications and Networks Security, Cryptography, Database Security, Electronic Commerce Security, Multimedia Security, System Security, etc.

1. Submission Procedure

Authors are strongly encouraged to submit their papers electronically by using online manuscript submission at http://ijns.jalaxy.com.tw/.

2. General

Articles must be written in good English. Submission of an article implies that the work described has not been published previously, that it is not under consideration for publication elsewhere. It will not be published elsewhere in the same form, in English or in any other language, without the written consent of the Publisher.

2.1 Length Limitation:

All papers should be concisely written and be no longer than 30 double-spaced pages (12-point font, approximately 26 lines/page) including figures.

2.2 Title page

The title page should contain the article title, author(s) names and affiliations, address, an abstract not exceeding 100 words, and a list of three to five keywords.

2.3 Corresponding author

Clearly indicate who is willing to handle correspondence at all stages of refereeing and publication. Ensure that telephone and fax numbers (with country and area code) are provided in addition to the e-mail address and the complete postal address.

2.4 References

References should be listed alphabetically, in the same way as follows:
For a paper in a journal: M. S. Hwang, C. C. Chang, and K. F. Hwang, ``An ElGamal-like cryptosystem for enciphering large messages," *IEEE Transactions on Knowledge and Data Engineering*, vol. 14, no. 2, pp. 445--446, 2002.

For a book: Dorothy E. R. Denning, *Cryptography and Data Security*. Massachusetts: Addison-Wesley, 1982.

For a paper in a proceeding: M. S. Hwang, C. C. Lee, and Y. L. Tang, ``Two simple batch verifying multiple digital signatures," in *The Third International Conference on Information and Communication Security (ICICS2001)*, pp. 13--16, Xian, China, 2001.

In text, references should be indicated by [number].

Subscription Information

Individual subscriptions to IJNS are available at the annual rate of US$ 200.00 or NT 7,000 (Taiwan). The rate is US$1000.00 or NT 30,000 (Taiwan) for institutional subscriptions. Price includes surface postage, packing and handling charges worldwide. Please make your payment payable to "Jalaxy Technique Co., LTD." For detailed information, please refer to http://ijns.jalaxy.com.tw or Email to ijns.publishing@gmail.com.