IJNS
International Journal of Network Security

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

Vol. 19, No. 6 (Nov. 2017)

INTERNATIONAL JOURNAL OF NETWORK SECURITY

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

International Journal of Network Security

Vol. 19, No. 6 (Nov. 1, 2017)

The Capacity Analysis in the Secure Cooperative Communication System

Jong-Shin Chen[1], Cheng-Ying Yang[2], and Min-Shiang Hwang[3,4],
(Corresponding author: Min-Shiang Hwang)

Department of Information and Communication Engineering, Chaoyang University of Technology[1]
Taichung 41349, Taiwan, R.O.C.
Department of Computer Science, University of Taipei[2]
Taipei 10048, Taiwan, R.O.C.
Department of Computer Science and Information Engineering, Asia University[3]
No. 500, Lioufeng Raod, Wufeng Shiang, Taichung 41354, Taiwan
Department of Medical Research, China Medical University Hospital, China Medical University[4]
No.91, Hsueh-Shih Road, Taichung 40402, Taiwan
(Email: mshwang@asia.edu.tw)
(Invited Jan. 11, 2017)

Abstract

With the characteristic of spatial diversity and low cost, cooperative system is a tendency for the future communications. In the wireless communication system, there exist degradation factors such as signal fading, multipath transmission, signal inferences, bandwidth limitation and so on. In addition to these degradation factors, the wireless transmission is not a secure environment. The information might be leaked during the transmission. Currently, the issues of privacy and security have become increasingly important for the mobile users. Traditionally, the security scheme is applied to the higher network layer. Encryption can be complex and difficult without infrastructure. It is not suitable to apply to the equipment with low computing resources, such as Internet of Things (IoT) application. Within information theoretic security characterizes the fundamental ability of the physical layer to provide a secure transmission. Hence, this work concentrates on the secure cooperative communication system. Based on the Shannon third theorem on channel capacity, this work analyzes the secrecy capacity between the source station and the destination station. For a practical situation in the system, the scenario includes multiple source stations, multiple relay stations, multiple destination stations, and eavesdroppers. For the positive secrecy rate consideration, the maximum mutual information between the source station and the destination station and the minimum mutual information between the source station and the eavesdropper should be held. To ensure a secure communication, the derived theoretical solution could be applied to find the optimal relay assignment. Beyond the relay selection, some issues related to the secure cooperative communication are suggested for the future researches in the final.

Keywords: Internet of Things (IoT); Multiple Input Multiple Output (MIMO); Physical Layer Security; Secrecy Capacity; Secure Cooperative Communications; Shannon Third Theorem

1 Introduction

The wireless communications provide a number of multimedia services for the mobile users. However, there exist degradation factors, such as signal fading, multipath transmission, signal inferences, bandwidth limitation and so on because of the radio transmission. Under the condition of imitated transmission bandwidth, to improve system performance in the wireless systems could be a significant work. Especially, the spatial diversity techniques could be employed to improve the system performance [4, 7, 8, 11, 13]. For example, in the Multiple Input Multiple Output (MIMO) system, a spatial diversity gain is employed to improve the system performance. However, MIMO is with the high cost of hardware implementation because there are multiple antennas at both the transmitter and receiver [3, 7, 11]. Instead of MIMO technique, the cooperative communications with a relay channel increase the system capacity without extra antennas [8, 13].

Cooperative communication is an idea to employ the wireless channel to make communication nodes help each other to implement the communication process [11]. It benefits the wireless communication with the gain similar to that of MIMO. It improves the system capacity,

transmission speed, and system performance. On the other hand, it could reduce the power consumption at the communication ends to extend the lifetime of the system. It is suitable to provide the multimedia services for the mobile devices. In the cooperative communication systems, the relay station functions with a character of spatial diversity. Comparing with multiple carrier modulation schemes, the relay stations work as the receivers and the transmitters. The relay station not only forwards the transmitted information but also process the received signal. It provides a high throughput performance. The destination station could receive the information with a spatial diversity with employing the relay selection scheme. Even though the destination station has no multiple antennas, by employing the relay station as the virtual antenna, it increases the transmission data rate and provides a reliable channel capacity [7]. With a consideration of low cost, the cooperative communication system is a tendency in the future communications.

However, the wireless communication is not a secure environment for a highly private request. The issues of privacy and security have become increasingly important for the mobile users. Besides, security is the fundamental requirement for a personal communication. Secure communications enable the authenticated destination station could successfully receive the information from the source station. Also, it protects the transmitted information from the eavesdroppers to interpret. Traditionally, the secure communication depends on the cryptographic encryption at the application layer. The complex and difficult cryptography is the practical techniques without infrastructure for the secure communication in the presence of third parties [5, 10], i.e. eavesdroppers. The technique relates to construct and analyze the transmission protocols to overcome the influence of eavesdroppers to ensure the security constraints with confidentiality, integrity, and availability including authentication, and non-repudiation. Cryptographic encryption converts the meaningful information to be the apparent nonsense to avoid the eavesdroppers to release the desired and transmitted information. However, the encryption algorithms are developed based on the assumption of limited computational capability at the eavesdroppers [10]. Also, these encryptions assume there are a perfectly secret key management and the distribution scheme for the users. Hence, it is not practical for the wireless communication application. Especially, it is obvious for IoT application [19]. Besides, for the secure purpose, the social-aware networking has been proposed to the secure cooperative communication systems [6, 17]. The authentication protocol within the networking could be the preliminary limitation for access control scheme. Eventually, the secure communication could be hold based on the secrecy rate [9]. Hence, physical layer security has been proposed for this purpose [3, 5, 14, 20, 21].

In the cooperative communication system, the information is transmitted from the source station to the destination station with the help of relay stations [4]. Among the relay stations, the transmitted information is unwrapped in the presence of one or more eavesdroppers. The information could be eavesdropped from the source station or from the relay which the source station adopts in the cooperative communication. Hence, to provide a secure communication and service quality could become an important issue. In Section 2, the concept of the cooperative communication system is described and the quantity measurements of the information between the source station and the destination station are provided. Section 3 illustrates the analytical model for the secure communication and the theoretical requirement for the cooperative system is derived. Under the secure cooperative communication requirement, the constraint of the relay selection strategy is shown in Section 4. The conclusion and the further work suggestion are given in the final.

2 The Cooperative Communication

Similarly to the Multiple Input Multiple Output (MIMO) technique with a character of spatial diversity, the cooperative communication system uses single-antenna mobiles in a multi-user environment to share their antennas to create a virtual MIMO system and to improve the system performance. Basically, the concept of the cooperative communication is illustrated in Figure 1.

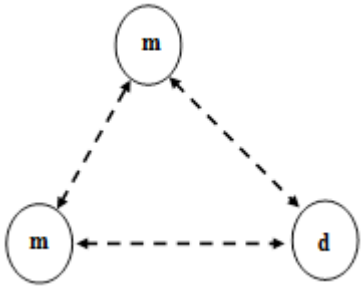


Figure 1: The concept of cooperative communication

In Figure 1, there are two mobile devices transmit the data to the same destination station simultaneously. Each device has its own antenna and cannot generate a spatial diversity. With the cooperation from the other device, it might be possible for one device to receive the other, the transmitted data can be forwarded with the same information to the destination station. One of these two mobile devices could be thought as the source station and the other is the corresponding relay station. With these three nodes, the source station, the relay station and the destination station, the capacity analysis of the cooperation communication system including these three nodes could be modeled as that in Figure 2.

In Figure 2, $h_{s,r}$ and $h_{r,d}$ denote as the channel response between source station to resource station and the

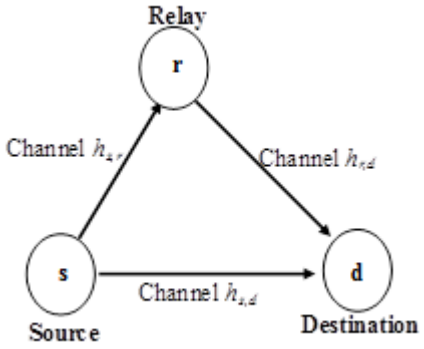


Figure 2: Analytical model for the cooperative communication

channel response between the resource station to destination station, respectively. The source station broadcasts the information to the destination station with both straight forward link and the assistant link with the relay station. This relay station might be another user in the system. The relay station functions as receiving the transmitted information from the source station and transmitting the information to the destination station. At the destination station, it multiple receives the information from the source station and the relay station. In the cooperative communication system, the destination station employs Maximal Ratio Combining (MRC) technique or Selective Combining (SC) technique to the received signals from the source station and the relay station [2]. It depends on the cooperative strategy used in the relay station. For example, in Amplify-and-Forward transmission mode, under AWGN channel, the maximize mutual information between the source and the destination becomes [18]:

$$I_{s,d} = \frac{1}{2}\log_2(1+\frac{P_s|h_{s,d}|^2}{N_0} + \frac{1}{N_0}\frac{P_sP_r|h_{s,r}|^2|h_{r,d}|^2}{P_s|h_{s,r}|^2+P_r|h_{r,d}|^2+N_0}) \quad (1)$$

where P_s is the signal power from the source station, P_r is the signal power from the relay station, and $n_{s,r}$ and $n_{s,d}$ are AWGN with the variance N_0. In Fixed Decode-and-Forward transmission mode, under AWGN channel, the mutual information between the source and the destination becomes [16]

$$I_{s,d} = \min\{I_{s,r}, I_{r,d}\} \quad (2)$$

where

$$I_{s,r} = \frac{1}{2}\log_2(1+SNR_{s,r}) = \frac{1}{2}\log_2(1+\frac{P_s|h_{s,r}|^2}{N_0})$$

and

$$I_{r,d} = \frac{1}{2}\log_2(1+SNR_{r,d}) = \frac{1}{2}\log_2(1+\frac{P_r|h_{r,d}|^2}{N_0}).$$

The system capacity depends on the maximum mutual information between the source station and the destination stations.

However, the wireless communication is not a secure environment. Within the theoretical information security characterizes [12], the fundamental ability of the physical layer provides a secure transmission. For example, channel coding and spread spectrum techniques provide secure communications. Hence, based on the Shannon third theorem on channel capacity, the secure communication could be hold based on the positive secrecy rate [1, 12]. The secrecy rate (i.e. secrecy capacity) of transmission is defined as the mutual information difference between the mutual information to the destination and that to the eavesdropper, i.e.

$$C_{s,d} = I_{s,d} - I_{s,e}. \quad (3)$$

3 The Secure Cooperative System

The secure cooperative system could be illustrated in Figure 3. There are a source station, a relay group, an eavesdropper group and a destination station in the system. In the system, the source station transmits the information. The information could be delivered directly to the destination station through the straightforward link between the source station and the destination. On the other hand, the information might be transmitted to the relay station and, then, delivered to the destination station with the help of the relay station. Similarly, the scenario of the information transmitted to the eavesdropper could be held in this wireless environment.

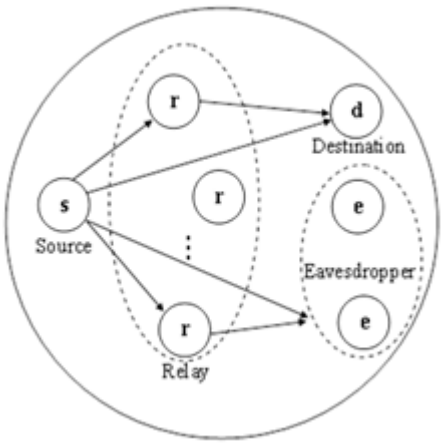
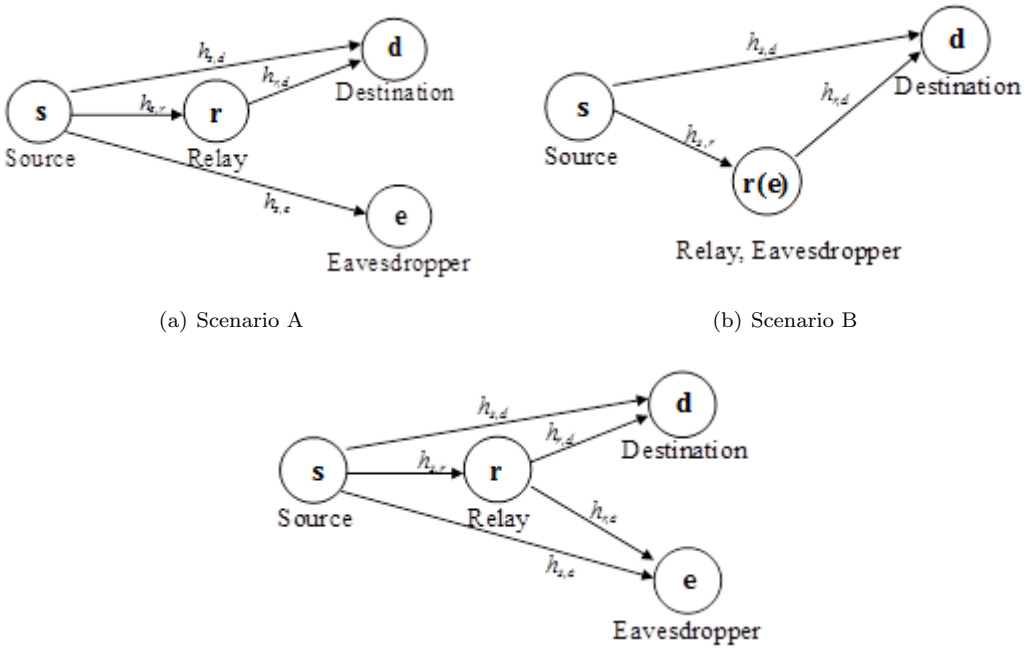


Figure 3: The cooperative communication environment

In order to consider the secure communication between the source station and the destination station, the location of the eavesdroppers could be considered with the following scenarios in Figure 4.

In Figure 4(a), the eavesdropper locates at the end communication link. The cooperative system employs



(a) Scenario A

(b) Scenario B

(c) Scenario C

Figure 4: The scenario for the location of eavesdroppers

the relay station to forward the information to the destination station. Hence, the mutual information between the source station and the destination could be obtained according to the previous theoretical derivation [16, 18]. Also, the mutual information between the source station and the eavesdropper could be obtained. The nodes in the cooperative communication system could work as the transmitter and receiver as that mentioned previously and each node could function as the relay station between the source station and destination station. Hence, in Figure 4(b), the relay could work as the eavesdropper to forward the information from the source station to the destination station. Similarly, the theoretical mutual information between the source station and the destination could be obtained according to the previous theoretical derivation. At the meantime, the analysis to mutual information between the source station and the eavesdropper could be considered as the case in Figure 4(a) with the same channel impulse response to the relay station, i.e. $h_{s,e} = h_{s,r}$. The case in Figure 4(b) could be considered as a special case of the scenario A. In Figure 4(c), the eavesdropper locates at the end communication link. With the different scenario to the scenario B, the relay station is not an eavesdropper and it forwards the transmitted information to the destination station. However, the eavesdropper receives the information from the source station and the relay station. For simplified analysis, the scenario C could be considered as the general case. For example, the situation in Figure 4(a) could be modified as that with the broken link between the relay station and eavesdropper in Figure 4(c). Hence, Figure 4(c) can be considered as the general situation for secure analysis and reshown in Figure 5.

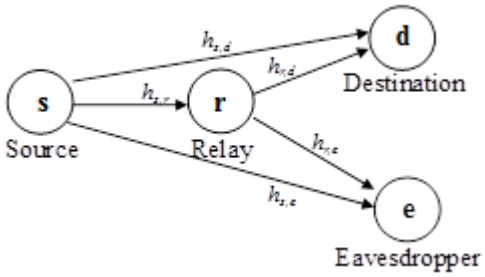


Figure 5: Analytical model for the secure cooperative communication

As the mentioned previously, for example, the maximize mutual information with AF mode between the source and the eavesdropper

$$I_{s,e} = \frac{1}{2}\log_2(1 + \frac{P_s|h_{s,e}|^2}{N_0} + \frac{P_r|h_{r,e}|^2}{N_0}) \quad (4)$$

Under the condition that the relay station could not decode the received signal correctly, the mutual information between the source station and the eavesdropper is

$$I_{s,e} = \frac{1}{2}\log_2(1 + \frac{P_s}{N_0}|h_{s,e}|^2). \quad (5)$$

The secrecy capacity of transmission is defined in Equation (3).

When the secrecy capacity is negative, the intercept event will be held and the eavesdropper could intercept the transmitted information successfully. Hence, the condition for a secure communication, the secrecy capacity $C_{s,d}$ should be positive. The maximum of secrecy capacity $C_{s,d}$ could be reached with maximizing the mutual information between the source station and the destination station and minimizing the mutual information between the source station and the eavesdropper. Hence, the relay selection strategy in the secure cooperative system could be employed with the concern of maximum the secrecy capacity in the system.

4 Relay Selection Strategy

For the relay selection, almost researchers concentrated on the situation that the single source station and discussed the relay assignment. However, in practical, there exist many source stations in the system. There are a lot of users requiring the relay stations to transfer the information. Based on this situation, relay selection should consider the multiple source stations, multiple relay stations and multiple destination stations in the system, as shown in Figure 6

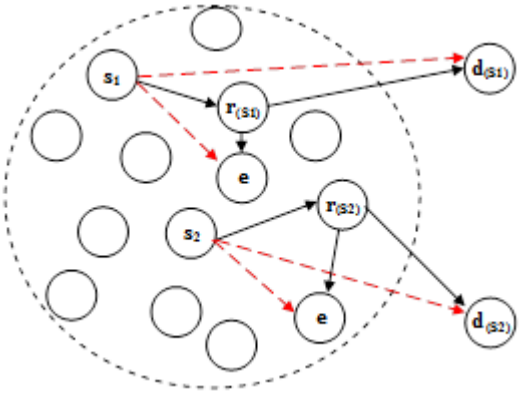


Figure 6: Fixed mode relay selection

The analysis to the relay selection is based on fixed mode in the cooperative communication system [16, 18]. It supposes that there are v nodes in the system and those nodes are denoted as set V. In the set V, there are k nodes as the source stations there are m nodes that could function as the source station and the relay station. These m nodes are denoted as set M. All the source stations are denoted as set S, i.e. $S \subseteq M$. $r(s)$ is defined as the set of the relay stations with forwarding the transmitted signal for the source station s. In this system, all source stations have their own destination stations. $d(s_i)$ represents the destination station for source station s_i. The destination station does not belong to set M. To analyze the secrecy capacity in the cooperative communications, initially, consider for the source station i transmits the information to the destination station d(si) with the relay station r_i. Under AWGN channel, for the example in AF mode, the mutual information between the source station i and the destination is described in Equation (1):

$$I_{s,d(s_i)} = \frac{1}{2}\log_2(1 + \frac{P_{s_i}|h_{s_i,d(s_i)}|^2}{N_0} + \frac{P_{r_i}||h_{r_i,d(s_i)}|^2}{N_0}).$$

Similarly, Equation (3) could be applied to DF mode if the relay station could correctly decode the transmitted signal and maximal ratio combining (MRC) strategy the equal gain for each forward link applied. However, if the relay station could not decode the transmitted signal correctly, Selective Combining (SC) strategy applied, the mutual information between the source station i and the destination is described in Equation (3) and could be rewritten as

$$I_{s_i,d(s_i)} = \frac{1}{2}\log_2(1 + \frac{P_{s_i}}{N_0}|h_{s_i,d(s_i)}|^2).$$

In the both modes, the mutual information between the source station and the eavesdropper is

$$I_{s_i,e} = \frac{1}{2}\log_2(1 + \frac{P_{s_i}|h_{s_i,e}|^2}{N_0} + \frac{P_{r_i}|h_{r_i,e}|^2}{N_0}).$$

The secrecy capacity in the cooperative system becomes

$$C_{s_i,d(s_i)} = I_{s_i,d(s_i)} - I_{s_i,e}.$$

To approach the maximal mutual information achieved in the system at the destination stations should consider the channel condition, under the situation of multiple source station, multiple relay stations, and multiple destination station environments. Hence, the relay selection strategy for secure cooperative communication could be developed based on the maximum mutual information between the source station i and the destination station, the minimum mutual information between the source station i and the eavesdropper and the positive secrecy capacity, i.e.

$$I_{s,d(s)} = \max_{r=(r_1,r_2,\cdots,r_k)\in R(s_1)\times R(s_2)\times\cdots\times R(s_k)} \sum_{i=1}^{k} I_{s_i,d(s_i)}$$

and

$$I_{s,e} = \min_{r=(r_1,r_2,\cdots,r_k)\in R(s_1)\times R(s_2)\times\cdots\times R(s_k)} \sum_{i=1}^{k} I_{s_i,e}$$

and, the positive secrecy capacity $C_{s_i,d(s_i)}$. Hence, the limitation to this problem could become

$$\begin{aligned} C &= \max \sum_{i=1}^{k}\sum_{j=1}^{m} \rho_{i,j} C_{s_i,d(s_i)} \\ &= \sum_{i=1}^{k}\sum_{j=1}^{m} \rho_{i,j} \cdot \{\max(I_{s,d(s)} - I_{s,e})\} \end{aligned}$$

under the conditions,

$$\begin{aligned} \sum_{i=1}^{k} \rho_{i,j} &\leq 1, \forall i = 1,2,\cdots,k, \quad \text{and} \\ \sum_{j=1}^{m} \rho_{i,j} &= 1, \forall j = 1,2,\cdots,m \end{aligned}$$

where $\rho_{i,j}$ is defined as the connection between the relay station i to the destination station j. Hence, how to choose the appropriate relay station i to approach the maximum mutual information becomes an important issue. The limitation for the relay selection strategy is with the above derivate equations.

5 Conclusion and Further Work

With the character of low cost, the cooperative system is a tendency for the future communications. For a practical situation in the cooperative system, the scenario includes multiple source stations, multiple relay stations, multiple destination stations, and eavesdroppers. This paper concentrates on the physical layer secure in the cooperative systems and develops the theoretical limitation for the relay assignment scheme. For the secrecy capacity in the system, it begins to analyze the theoretical mutual information between the source station and the destination station. The maximum mutual information could be achieved by the power management in the system. Also, it could be obtained with the appropriate relay selection strategy. On the other hand, in order to obtain the maximum the secrecy capacity, one possible solution is to achieve the minimum mutual information between the source station and the eavesdropper. To ensure the secure communication, based on the information theory, the secrecy capacity should be kept a positive value. By deriving the theoretical solution to the system performance in the secure cooperative system, this work applies the derived results to the considered environment to construct the optimal relay assignment scheme. By the way, the better relay selection strategy could be developed with maximizing the secrecy capacity in the system. Also, the effective relay selection algorithm could be developed in the future.

Other important issues to the secure cooperative communications including the power distribution, the coding schemes, the multiple access technique, and the transmission protocol and so on could be made further researches. Power control management is to find the appropriate power distribution among the relay stations. Obviously, it could be found in the theoretical mutual information analysis. Within the mathematical derivations, the transmitted power from the source station and the relay stations effects the system capacity. This power control issue for the relay stations could be included in the design to achieve the optimal throughput for the cooperative system. The coding schemes and multiple access techniques convert the desired information to be the nonsense data. It increases the secrecy capacity between the source station and the destination station to make sure the positive secrecy rate. These practical considerations and requirements on the system design could contribute to constructing a cooperative system as well as extensions to the fundamental idea of secure communication.

Acknowledgment

This research was partially supported by the Ministry Of Science and Technology, Taiwan (ROC), under contract no.: MOST 103-2632-E-324-001-MY3.

References

[1] J. Barros and M. R. Rodrigues, "Secrecy capacity of wireless channels," in *Proceedings of 2006 IEEE International Symposium on Information Theory*, pp. 356–360, 2016.

[2] E. Beres and R. S. Adve, "Selection cooperation in multi-source cooperative networks," *IEEE Transactions on Wireless Communications*, vol. 7, no. 1, pp. 118–127, 2008.

[3] X. Chen, L. Lei, H. Zhang and C. Yuen, "Large-scale MIMO relaying techniques for physical layer security: AF or DF?," *IEEE Transactions on Wireless Communications*, vol. 14, no. 9, pp. 5135–5146, 2015.

[4] Y. Chou, J. Zhu, X. Wang and V.C. Leung, "Improving physical-layer security in wireless communications using diversity techniques," *IEEE Network*, vol. 29, no. 1, pp. 42–48, 2015.

[5] L. Dong, Z. Han, A. P. Petropulu and H. V. Poor, "Improving wireless physical layer security via cooperating relays," *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1875–1888, 2010.

[6] X. Gu, L. Tang, and J. Han, "A social-aware routing protocol based on fuzzy logic in vehicular ad hoc networks," *Proceedings of 2014 International Workshop on High Mobility Wireless Communications (HMWC'14)*, pp. 12–16, 2014.

[7] L. Li, X. Zhou, H. Xu, G. Y. Li, D. Wang, and A. Soong, "Simplified relay selection and power allocation in cooperative cognitive radio systems," *IEEE Transactions on Wireless Communications*, vol. 10, no. 1, pp. 33–36, 2011.

[8] H. C. Lu and W. Liao, "Cooperative strategies in wireless relay networks," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 2, pp. 323–330, 2012.

[9] A. Mukherjee, S. A. A. Fakoorian, J. Huang and A. L. Swindlehurst, "Principles of physical layer security in multiuser wireless networks: A survey," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 3, pp. 1550–1573, 2014.

[10] D. W. K. Ng, E. S. Lo and R. Schober, "Robust beamforming for secure communication in systems with wireless information and power transfer," *IEEE Transactions on Wireless Communications*, vol. 13, no. 8, pp.4599–4615, 2014.

[11] A. Nosratinia, T. E. Hunter and A. Hedayat, "Cooperative communication in wireless networks," *IEEE Communications Magazine*, vol. 42, no. 10, pp. 74–80, 2004.

[12] C. E. Shannon, "Communication theory of secrecy systems," *Bell System Technology Journal*, vol. 29, pp. 656–715, 1949.

[13] K. Vardhe, D. Reynolds and B. D. Woerner, "Joint power allocation and relay selection for multiuser cooperative communication," *IEEE Transactions on Wireless Communications*, vol. 9, no. 4, pp. 1255–1260, 2010.

[14] L. Wang, N. Yang, M. Elkashlan, P. L. Yeoh and J. Yuan, "Physical layer security of maximal ratio combining in two-wave with diffuse power fading channels," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 2, pp. 247–258, 2014.

[15] Y. Wang and G. Noubir, "Distributed cooperation and diversity for hybrid wireless networks," *IEEE Transactions on Mobile Computing*, vol. 12, no. 3, pp. 596–608, 2013.

[16] J. H. Wen, C. H. Chiang, Y. S. Lin, C. Y. Yang, "Performance evaluation for the cooperative communication systems in decode-and-forward mode with a maximal ratio combining scheme," *WSEAS Transactions on Communications*, vol. 13, pp. 424–429, 2014.

[17] F. Xia, L. Liu, J. Li, A. M. Ahmed, L. T. Yang and J. Ma, "BEEINFO: Interest-based forwarding using artificial bee colony for socially-aware networking," *IEEE Transactions on Vehicle Technology*, vol. 64, no. 3, pp. 1–11, 2014.

[18] C.Y. Yang, Y.S. Lin and M.S. Hwang, "Downlink relay selection algorithm for amplify-and-forward cooperative communication systems," in *Proceedings of 2013 Seventh International Conference on Intelligent, and Software Intensive Systems (CISIS'13)*, pp. 331–334, 2013.

[19] Z. K. Zhang, M. C. Y. Cho and S. Shieh, "Emerging security threats and countermeasures in IoT," in *Proceedings of the 10th ACM Symposium on Information, Computer and Communications Security*, pp. 1–6, 2015.

[20] T. Zou, X. Wang and W. Shen, "Optimal relay selection for physical-layer security in cooperative wireless networks," *IEEE Journal on Selected areas in Communications*, vol. 31, no. 10, pp. 2099–2111, 2013.

[21] Y. Zou, J. Zhu, X. Wang and V. Leung, "Improving physical-layer security in wireless communications using diversity techniques," *IEEE Network*, vol. 29, no. 1, pp. 42–48, 2015.

Biography

Jong-Shin Chen was born in 1972. He received the B.S. and Ph.D. degrees in computer science from Feng Chia University, Taiwan, in 1996 and 2003, respectively. Currently, he is an associate professor in the Department of Information and Communication Engineering, ChaoYang University of Technology, Taiwan. His research interests include big-data mining, capacity planning, and wireless networking.

Cheng-Ying Yang received the M.S. degree in Electronic Engineering from Monmouth University, New Jersey, in 1991, and Ph.D. degree from the University of Toledo, Ohio, in 1999. He is a member of IEEE Satellite & Space Communication Society. Currently, he is employed as an Associate Professor at Department of Computer Science, University of Taipei, Taiwan. His research interests are performance analysis of digital communication systems, error control coding, signal processing and computer security.

Min-Shiang Hwang received the Ph.D. degree in computer and information science from the National Chiao Tung University, Taiwan in 1995. Dr. Hwang was the Chairman of the Department of Information Management, Chaoyang University of Technology (CYUT), Taiwan, during 1999-2003. He was a distinguished professor and Chairman of the Department of Management Information Systems, NCHU, during 2003-2011. He obtained the 1997, 1998, 1999, 2000, and 2001 Excellent Research Award of National Science Council (Taiwan). His current research interests include information security, cryptography, image compression, and mobile computing. Dr. Hwang has published over 200+ articles on the above research fields in international journals.

A New Secure Authentication and Key Exchange Protocol for Session Initiation Protocol Using Smart Card

Mourade Azrour, Yousef Farhaoui, Mohammed Ouanan

(Corresponding author: Mourade Azrour)

Department of Computer Science, M2I Laboratory, ASIA Team, Moulay Ismail University
BP 509, Boutalamine, 52000 Errachidia, Morocco
(Email: azrour.mourade@gmail.com)

(Received July. 26, 2016; revised and accepted Nov. 15 & Dec. 25, 2016)

Abstract

In today′s communications over Internet Protocol (IP), Session Initiation Protocol (SIP) is using establish, modify and terminate the sessions multimedia among participants. Authentication is the most security service required for SIP. Authentication HTTP Digest is the original authentication protocol proposed for SIP. However, this protocol is demonstrated insecure against different attacks. To improve the authentication, a different authentication protocols have being proposed. Very recently, Jiang et al. demonstrate that Zhang et al.'s scheme cannot resist to impersonation attack. Then, Jiang et al. proposed their protocol. However, in this paper we show that Jiang et al.'s protocol suffer from server spoofing attack. In order to overcome this problem we propose an improved SIP authentication protocol. The security analysis shows that the proposed protocol is more secure and can deal with several attacks.

Keywords: Authentication Protocol; Elliptic Curve Cryptography; Session Initiation Protocol; Smart Card

1 Introduction

Telephony over IP (ToIP) is a service that allows transferring voice communications flow on IP (Internet Protocol). This is the application that will require the IP infrastructure as the standard for all types of information or media. IP telephony is based on open standards. To establish ToIP communication two types of protocols are required which are signaling protocol and transport protocol. In recently decade, Session Initiation Protocol (SIP) is the most signaling protocol used for establishing, altering and terminating session multimedia between different users.

Authentication is the most security service required by SIP. The original SIP authentication protocol is HTTP Digest Authentication. This protocol was found vulnerable to deferent attacks. In order to reinforce SIP authentication, a large community has been participated by proposing the different protocols based on various mechanisms. Generally, authentication protocol can be categorized as Password Authentication protocol [17, 21, 22], ID-based protocol [7] and Elliptic Curve Cryptography based protocol [2].

In 2005, Yang et al. [25] demonstrated that the original SIP authentication protocol is vulnerable to Off-line password guessing attack and stolen verifier attack. So, they based on Diffie-Hellman Key Exchange [3] to propose their protocol which is secure against Off-line password guessing attack, server spoofing attack and replay attack. However, the protocol of Yang et al. requires maintenance and configuration of the passwords table. In addition, it based on the discrete logarithm problem which requires an important computation time cost. Therefore, it is not suitable for applications with low memory and limited computing capability. In 2006, Huang et al. [9] proposed a new protocol based on one-way hash functions. After comparing the computational complexity of their protocol with the Yang et al.'s protocol they concluded that their protocol is the fastest. Moreover, Jo et al. [12] demonstrated that the protocol of Yang et al. and the protocol of Huang et al. are both vulnerable to Off-line password guessing attack.

To overcome this weakness, Durlanik and Sogukpinar [4] based on the Yang et al.'s protocol to propose another SIP authentication protocol using the Elliptic Curve Cryptography Diffie-Hellman (ECCDH) [13]. They demonstrated that their protocol reduces the computation time cost. Because it uses a small size key but offer the same security offered by the Diffie-Hellman large key size. However, Yoon et al. [27, 29] introduced that the protocol of Durlanik and Sogukpinar cannot resist the stolen verifier attack and Denning-Sacco attack.In 2008, Wu et al. [23] proposed a new SIP authentication and key exchange protocol based

on elliptic curve cryptography (ECC). Wu et al. prove that their protocol is secure against man-in-the-middle attack, replay attack, Off-line password guessing attack and server spoofing attack. Unfortunately, this protocol is vulnerable to Off-line password guessing attack, Denning-Sacco attack and stolen verifier attack [26]. In the same year, Tsai [19] proposed an authentication protocol for SIP based on random nonce. The protocol uses one-way hash functions, and a bit-wise exclusive-or(XOR) operation to encrypt and decrypt messages. As result, the calculation time cost is reduced when it compared with the existing protocols. For this, it is desirable for applications with low computing capability. However, Yoon et al. [28], then Arshad and Ikram [1] found that Tsai′s protocol is vulnerable to Off-line password guessing attack, server spoofing attack and stolen verifier attack. One year later, Yoon and Yoo [28] proposed a new secure SIP authentication protocol. They demonstrated that their protocol is secure against the man-in-the-middle attack, Off-line password guessing attack, replay attack, modification attack, Denning-Sacco attack and stolen verifier attack. In addition, it provides mutual authentication, known key secrecy, session key secrecy and perfect forward secrecy. However, Liu and Koenig [14] demonstrated that this protocol is vulnerable to Off-line password guessing attack and partition attack. In 2011, Arshad and Ikram [1] demonstrated that Tsai et al.'s protocol is vulnerable to Off-line password guessing attack and stolen verifier attack, and it does not provide key known secrecy and perfect forward secrecy. As result, Arshad and Ikram presented an authentication protocol for SIP based on ECC. In 2012, Xie [24] showed that the protocol of Yoon and Yoo is insecure against stolen verifier attack and Off-line password guessing attack. Based on these attacks Xie proposes a new SIP authentication protocol. Then, he demonstrated that his protocol is more secure, and it is faster when it compared with existing protocols. However, Xie′s protocol is shown vulnerable to Off-line password guessing attack. In the same year, Tang et al. [18] noted that the protocol introduced by Arshad and Ikram is not secure against Off-line password guessing attack. In order to deal with this problem, they suggested another secure and efficient SIP authentication protocol based on Elliptic Curve Discrete Logarithm Problem (ECDLP).

In 2013, Zhang et al. [30] introduced for the first time smart-card-based protocol and key exchange for SIP. Then, they demonstrated that their protocol is secured against different attacks. However, Tu et al. [20], Irshad et al. [10], Zhang et al. [31], and Jiang et al. [11] demonstrated that Zhang et al.'s scheme is insecure against impersonation attacks. To solve the problem Jiang et al. [11] proposed a new SIP authentication protocol. Then, they proved that their scheme resist to various attacks. However, in this paper we demonstrate that Jiang et al.'s protocol is vulnerable to server spoofing attack. In order to overcome this weakness, we propose a secure and efficient SIP authentication protocol using smart card and based on elliptic curve cryptography.

The remainder of this paper is organized as follows. Section 2 delivers general information on the architecture and the original SIP authentication protocol. In Section 3, we review briefly Jiang et al.'s scheme. A cryptanalysis of Jiang et al.'s scheme is given in Section 4. In Section 5, we present our secure and efficient SIP authentication protocol. The security analysis and performance comparison are presented in Sections 6 and 7, respectively. Finally, section 8 concludes the paper.

2 Preliminaries

Session Initialization Protocol was initiated by the Multiparty Multimedia Session Control Group (MUSICG) in RFC 2543 [5]; then it was taken over and maintained by the SIP Group of the Internet Engineering Task Force (IETF). The first works are started from 1995, which resulted in a first version of SIP with the publication of RFC 2543 [6] in 1999; then a second version of SIP was published in 2002 to correct certain defects of the previous version.

SIP is a text-based protocol built on the basis of protocols such as HTTP or SMTP. The exchanges are in the form of dialogues (peer-to-peer relationships between agents) that include transactions (request/response). It is a widely used protocol, mainly for telephony applications on IP.

2.1 SIP Architecture

The architecture of SIP consists of a proxy server, redirect server, register server, location server, and User agents. The role of each component is described as follows.

User Agent Client (UAC): generates SIP requests before they were sent;

User Agent Server (UAS): generates answers to SIP requests (accepting, refusing, or redirecting);

User Agent (UA): it can be a SoftPhone (software) or HardPhone (IP phone). It is able to generate, send and receive SIP requests. It can act at the same time as a UAC and UAS;

Registrar Server: handles the registration of SIP terminals. This is a server that accepts SIP REGISTER requests;

Proxy Server: is a server which is connected to fixed or mobile terminals (UA). It plays the role of a server and client;

Redirect Server: is a server that accepts SIP requests, translates the SIP address of a destination network IP address and returns them to the client;

Location Server: The responsibility of the location server is to maintain information on the current location of the user agent. It provides the proxy server, redirect server, and register server, it allows for them to look up or register the location of the user agent.

2.2 HTTP Digest Authentication Protocol

The authentication of SIP is the most security service recommended by the IETF in RFC 2617 [16]. If a user wants to get access into the SIP services, he/she must be authenticated by server. In addition to needing to know if a user's identity is legitimate or not. The user also needs to know if the server with which it communicates is the legal server or not.

HTTP Digest Authentication for SIP is based on the mechanism challenge/response. Before the protocol execution, the client and the server share the password, the latter is used to verify the client′s identity. The messages exchanged between the server and the clients during authentication procedure are illustrated in Figure 1. and they are described as follows:

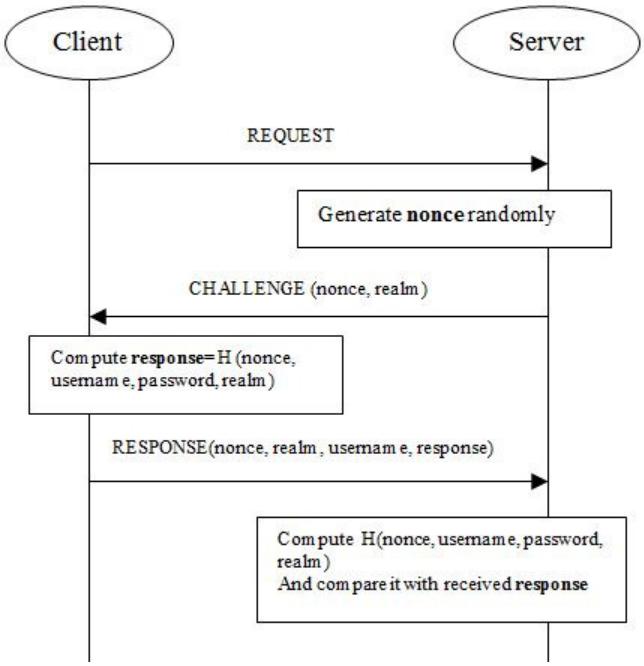


Figure 1: HTTP digest authentication

Step 1. Client → Server: REQUEST

The client sends a REQUEST to the server;

Step 2. Server → Client: CHALLENGE (nonce, realm)

After receiving REQUEST; the server generates CHALLENGE that includes a nonce and the client′s realm. Note that realm is used to verify username and password. Then the server sends back CHALLENGE to the client;

Step 3. Client → Server: RESPONSE (nonce, realm, username, response)

After receiving CHALLENGE from the server, the client computes the response by using received nonce, username, secret password, and realm. response=F(nonce, username, password, realm). Note that $F(\cdot)$ is a one-way hash function. Next, the client sends back the original REQUEST with the computed response, username, nonce and realm;

Step 4. According to username the server extracts the client's password. Then, the server verifies wither nonce is correct or not. If it is correct, the server computes F(nonce, username, password, realm) and uses it to compare it with the response. If they match, the server authenticates the identity of the client.

2.3 Elliptic Curve Cryptography

Elliptic curve cryptography (ECC) was introduced by Neal Koblitz in 1985 [25]. ECC proposed as an alternative to established public-key systems such as DSA and RSA. ECC have lately received a lot attention in information security. The main reason for the attractiveness of ECC is the fact that there is no sub-exponential algorithm known to solve the discrete logarithm problem on a properly chosen elliptic curve. This means that ECC uses the keys of small size but offer the same levels of security offered by the Diffie-Hellman key large size. Some benefits of having smaller key size include faster computations, and reductions in processing power, storage space and bandwidth. This makes ECC ideal for constrained environments such as cellular phones and smart cards [15].

The elliptic curve is a cubic equation of the form in Equation (1):

$$E : y^2 + axy + by = x^3 + cx^2 + dx + e \tag{1}$$

where a, b, c and e are real numbers.

In cryptosystem, the elliptic curve equation is defined as the form in Equation (2) over a prim finite field $(F)_p$, where $(a, b) \in F_p$ and $4a^3 + 27b^2 \neq 0(mod\ p)$. Given an integer $k \in (F_p)^*$ and a point $P \in E_p(a, b)$, the scalar multiplication kP over $E_p(a, b)$ can be computed as in Equation (3).

$$E_p(a, b) : y^2 = x^3 + ax + b(mod\ p) \tag{2}$$

$$kP = (P + P + ... + P)_{(k\ times)} \tag{3}$$

Definition 1. *Given two points P and Q over $(E)_p(a, b)$, the elliptic curve discrete logarithm problem (ECDLP) is to find an integer $k \in (F_p)^*$ such as $Q = kP$.*

Definition 2. *Given three points P,sP and kP over $E_p(a, b)$for s,$k \in (F_p)^*$, the computational Diffie-Hellman problem (DHP) is to find the point skP over $E_p(a, b)$.*

Definition 3. *Given two points P and $Q = sP+kP$ over $E_p(a, b)$ for $s, k \in (F_p)^*$, the elliptic curve factorization*

problem (ECFP) is to find two points sP and kP over $E_p(a,b)$.

3 Review of Zhang et al.'s Scheme

In this section we briefly review the Jiang et al.'s [30] authentication scheme for SIP. The Jiang et al.'s scheme consists of four phases: the setup phase, the registration phase, and the authentication phase. The notations used in this paper are shown in Table 1.

Table 1: Notions and their explanations

Notations	Explanations
U	The remote user
S	The remote server
X → Y:M	X sends a message M to Y
username	The identity of user U
PW	The password of user U
$E_p(a,b)$	An elliptic curve equation with order n
s	The long-live secret key of server S
$P_pub = sP$	The long-live public key of server S
SK	A session key
$h(\cdot)$,$h_1(\cdot)$,$h_2(\cdot)$	Three secure one-way hash functions
Z_q^*	Multiplication group of Z_q
$\|$	The string concatenation operator
$E_s(\bullet)$	Symmetric key encryption under the key s

3.1 System Setup Phase

Step 1. The server selects an elliptic curve equation $E_p(a,b)$ with the order n, and chooses a base point P over $(E)_p(a,b)$, where n is a large number for the security consideration. Then, it chooses a random number $s \in_R (Z_p)^*$ as the secret key and computes the public key $(P)_{pub} = sP$;

Step 2. The server selects three one-way hash functions, $h(\cdot) : \{0,1\}^* \to \{0,1\}^k$, $h_1(\cdot) : G \times \{0,1\}^* \times \{0,1\}^* \to \{0,1\}^k$, $h_2(\cdot) : G \times G \times \{0,1\}^* \to \{0,1\}^k$,where G is a cyclic addition group generated by P over $(E)_p(a,b)$;

Step 3. The server publishes $\{E_p(a,b)$, P, P_{pub}, $h(\cdot)$, $h_1(\cdot)$, $h_2(\cdot)\}$ and keeps s in secret.

3.2 Registration Phase

In this phase, the user registers on the SIP server through a secure channel. When a user wants to login into the remote server, he/she firstly should register to the remote server..The details of this phase are as follows.

R1: The user U selects his or her username, password PW and a random number $a \in_R Z_p^*$. After that, U computes $h(PW\|a)$ and sends $\{h(PW\|a), username\}$ to the server through a secure channel.

R2: after receiving the registration information, the server computes $R = h(h(PW\|a)\|username)s(-1)P$ and $X = h(username\|s)P$. Then the server stores R and X into the smart card and issues it to U.

R3: Upon receiving the card, U stores a in the card. Then the card contains (R, X, a).

3.3 Authentication Phase

Whenever the user wants to login into the remote server, he/she performs the following.

A1: $U \to S : REQUEST(username, V, W)$ U inserts his/her smart card into a card reader and inputs his/her username and password PW. Then, U's smart card picks a random number $b \in_R (Z_p)^*$, and computes V= $bR + X$ and $W = h(h(PW\|a)\|username)P_{pub}$. Next, the card sends a request message $REQUEST(username, V, W)$ to the server.

A2: $S \to U : CHALLENGE(realm, Auth_s, S, r)$ After receiving the request message, the server S computes $(X)' = h(username\|s)P$ and $W' = s^2(V - X')$. Then, it checks $W \stackrel{?}{=} W'$, if true it chooses two random integers $c \in_R Z_p^*$ and $r \in_R Z_p^*$. Then computes $S = cP$, $K = cs(V - X')P$, $SK = h_1(K\|r\|username)$ and $Auth_s = h_2(K\|W\|r\|SK)$. Next, it sends message $CHALLENGE(realm, Auth_s, S, r)$ to U over a public channel.

A3: $U \to S : RESPONSE(realm, Auth_u)$ Upon receiving message $CHALLENGE(realm, Auth_s, S, r)$, U computes $K = bh(h(PW\|a)\|username)S$ and $SK = h_1(K\|r\|username)$ and verifies if $Auth_s \stackrel{?}{=} h_2(K\|h(h(PW\|a)\|username)bP_{pub}\|r\|SK)$.
If so, U computes $Auth_u = h_2(K \| h(h(PW \| a) \| username) bP_{pub} \| r + 1\|SK)$ and sends $RESPONSE(realm, Auth_u)$ back to the server over public channel. Otherwise, it deletes received information and the protocol stops.

A4: After receiving the RESPONSE message, the server verifies $Auth_u \stackrel{?}{=} h_2(K\|W'\|r + 1\|SK)$. If the message is authenticated, the server sets SK a shared session key with user U. Otherwise, it deletes received information and the protocol stops.

3.4 Password Changing Phase

This phase is similar to the Zhang et al.'s password changing phase. When the user U wants to update its password, it needs to agree on a session key with the server via the authentication phase in advance.

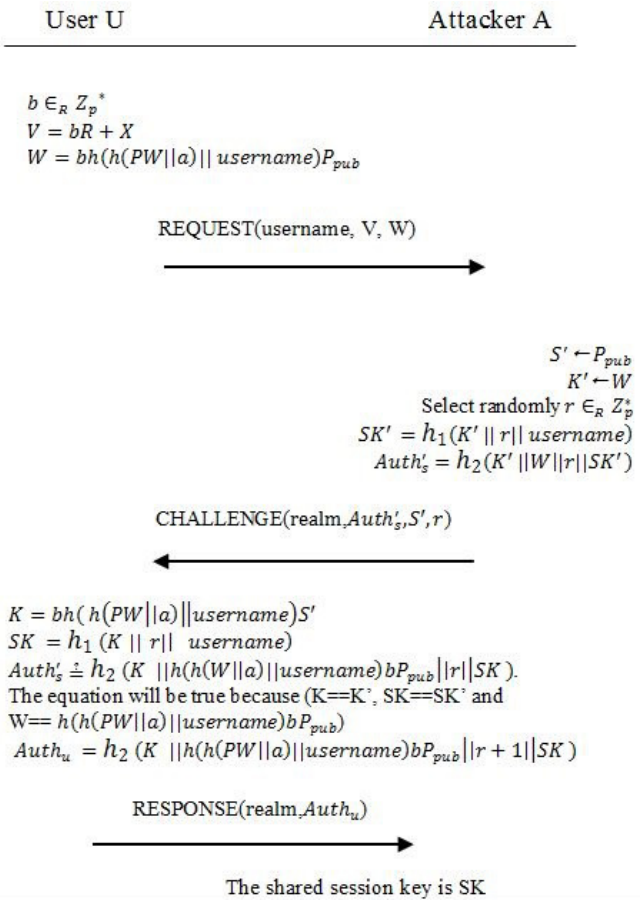


Figure 2: Server spoofing attack on Jiang et al.'s scheme

4 Cryptanalysis of Jiang et al.'s Scheme

Jiang et al. claimed that their protocols can resist various attacks. However, in this section, we will show that the Server spoofing attack, not as they claimed, is still effective in Jiang et al.'s protocol.

Let A be an attacker. A can eavesdrops the message REQUESTusername,V,W transmitted between server S and user U. A can get server′s public key, because S has published it with other parameters. Then, A can execute server spoofing attack. The detail of attack is illustrated in Figure 2 and is presented as follows.

Step 1. U inputs his username and password PW after inserting his smart card in card reader. The card generates randomly a number $b \in_R Z_p^*$ and computes $V = bR + X$ and $W = bh(h(PW\|a), username)P_{pub}$. Then, the card sends a request message REQUEST($username$, V, W) to S.

Step 2. A eavesdrops message REQUEST($username$, V, W) and get $username$, V, W. he/she generates a random number $r \in_R Z_p^*$. Next he/she get server public key, put its value in $S'(S' \leftarrow P_{pub})$ and put value of W in $K'(K' \leftarrow W)$. Then he/she computes $SK' = h_1(K'\|rl|username)$ and $Auth'_s = h_2(K'\|W\|r\|SK')$. Next, A sends message CHALLENGE(realm, $Auth'_s, S', r$) to U.

Step 3. Upon receiving message CHALLENGE (realm, $Auth'_s, S', r$), U computes $K = bh(h(PW \| a) \| username)S'$ and $SK = h_1(K \| r \| username)$ and verifies if $Auth'_s \stackrel{?}{=} h_2(K\|h(h(PW\|a)\|username)bP_{pub}\|r\|SK)$. The user will find true because:

$$\begin{aligned} K &= bh(h(PW\|a)\|username)S' \\ &= bh(h(PW\|a)\|username)P_{pub} \\ &= W \\ &= K' \\ SK &= h_1(K\|r\|username) \\ &= h_1(K'\|r\|username) \\ &= SK' \\ W &= bh(h(PW\|a)\|username)P_{pub} \\ &= h(h(PW\|a)\|username)bP_{pub}. \end{aligned}$$

Then user U authenticates attacker A and sends to him RESPONSE thinking that he/she communicate with a legal server S.

According to previous analysis, the adversary can easily impersonate identity of server at any time. The user U does not know whether the one he contacts is that the valid server or not. So the adversary can impersonate the server successfully. Therefore, Jiang et al.'s protocol is vulnerable to the server spoofing attack.

5 Our Proposed Protocol

In this section, in order to overcome weakness in Jiang et al.s protocol, we propose an improved and efficient authentication and key agreement protocol for SIP. Our protocol consists of four phases, which are system setup phase, registration phase, authentication and key agreement phase, and password changing phase. These phases are described as follows.

5.1 System Setup Phase

In this section, the server selects an elliptic curve equation $E_p(a, b)$, over a finite field F_q, an additive group G of order p and P a base point generator with order n over equation $E_p(a, b)$, n is a large prime of height entropy. Then, the server picks a random integer $s \in_R Z_p^*$ as its secrete key, and computes its public key $P_{pub} = sP$. Next, the server chooses three one-way hash functions $h(\cdot)$, $h_1(\cdot)$ and $h_2(\cdot)$. Finally, the server publishes all parameters except its private key, which it is saved secretly.

5.2 Registration Phase

When user wants to register in server and become a legal user, he has to perform the following steps.

R1: The user U selects his or her username, password PW and a random number $a \in_R Z_p^*$. After that, U computes $h(PW\|a)$ and sends $\{h(PW\|a), username\}$ to the server over a secure channel.

R2: after receiving the registration information, the server computes $R = h(h(PW\|a)\|username)s^{-1}P$ and $X = h(username\|s)P$. Then, the server stores R and X into the smart card and issues it to U.

R3: Upon receiving the card, U stores a in the card. Therefore, user card contains (R,X,a).

5.3 Authentication and Key Agreement Phase

As illustrated in Figure 3, whenever a legal user U wishes to log into the server, he/she have to inserts his/her smart card in card reader and inputs his/her username and password PW. Next, the following steps will be executed between server S and user U.

Auth 1: $U \rightarrow S$:$REQUEST(username, V, W)$

After inserting the smart card in card reader and inputting the username and password; the smart card of user U chooses a random $b \in_R Z_p^*$, and computes $V = bR + X$, $Y = bh(h(PW\|a), username)$ and $W = YP_{pub}$, then, he/she sends a request message $REQUEST(username, V, W)$ to the server over a public channel.

Auth 2: $S \rightarrow U$: $CHALLENGE(realm, Auth_s, S, r)$

When server S gets the request message, it computes $X' = h(username\|s)P$ and $W' = s^2(V - X')$. Then, it verifies $W \stackrel{?}{=} W'$. If true, U is authenticated and the server S picks randomly two integers $c, r \in_R Z_p^*$. Then, it computes $S = cP$, $K = cs(V - X')$, $SK = h_1(K\|r\|username\|X')$ and $Auth_s = h_2(K\|W\|r\|SK\|X')$. Next, it sends message $CHALLENGE(realm, Auth_s, S, r)$ to U over a public channel.

Auth 3: $U \rightarrow S$:$RESPONSE(realm, Auth_u$

Once the user U receives the CHALLENGE message, it calculates $K = YS$ and $SK = h_1(K\|r\|username\|X)$. Then, checks $h_2(K\|W\|r\|SK\|X)$ if is true, the server is authenticated. Then, user U computes $Auth_u$ as following $Auth_u = h_2(K\|W\|r + 1\|SK\|X)$ and sends $RESPONSE(realm, Auth_u)$ back to the server over public channel. Otherwise, it stops the protocol and deletes received and calculated parameters.

Auth 4: After receiving the RESPONSE message, the server computes $h_2(K\|W'\|r + 1\|SK\|X')$ and verifies that it equal to received $Auth_u$. If successful, the server sets SK a shared session key with user U. Otherwise, it stops the protocol and deletes received and calculated parameters

5.4 Password Changing Phase

This phase is similar to Zhang et al.s password changing phase. When the user U wants to update its password, it needs to agree on a session key with the server via the authentication phase in advance. The details of this phase are described as following.

Pass 1. $U \rightarrow S$: $(username, e, New_u)$

The user U chooses its new password PW^* and two random integers a^*, $e \in_R Z_p^*$ and computes $h(PW^*\|a^*)$ and $tag_u = h(username\|e\|h(PW^*\|a^*))$, it then uses SK to encrypt the new parameters: $New_u = E_{KS}(username\|e\|h(PW^*\|a^*)\|tag_u)$. Next, it sends message$(username, e, New_u)$ to server.

Pass 2. $S \rightarrow U$: (New_s)

Upon receiving the information, the server decrypts the message and then checks the validity of the authentication $tag_u \stackrel{?}{=} h(username\|e\|h(PW^*\|a^*))$. If it is valid, the server computes the new secret information $R^* = h(h(PW^*\|a^*)\|username)s^{-1}P$ and $tag_s = h(username\|e + 1\|R^*)$. Then, it sends encryption information $New_s = E_{KS}(R^*\|tag_s)$ to the user U.

Pass 3. The user U decrypts received message and verifies the validity of $tag_s \stackrel{?}{=} h(username\|e + 1\|R^*)$. If it is valid, the user U stores R^* and a^* in its smart card.

6 Security Analysis

In this section we will prove that our protocol provide mutual authentication and session key secrecy. Moreover, we will show that its secure against several attacks especially server spoofing attack, user impersonation attack, Denning-Sacco attack, replay attack, stolen verifier attack, offline password guessing attack, and man-in-the-middle attack.

6.1 Mutual Authentication

Mutual authentication means that both the user and server are authenticated to each other within the same protocol. In the proposed scheme the server can authenticate user after receiving REQUEST by checking W, and after receiving RESPONSE by checking $Auth_u$. Upon receiving message CHALLENGE user can authenticate the server by testing validity of $Auth_s$. Consequently, the proposed protocol provides mutual authentication.

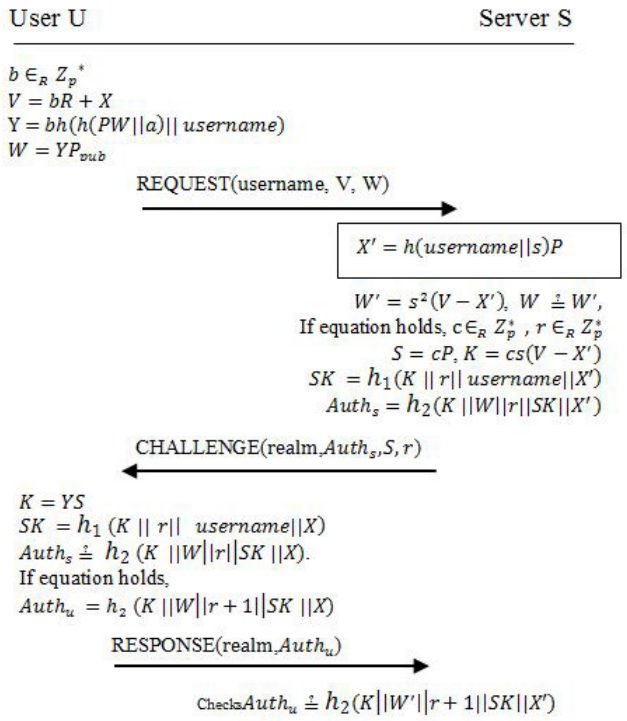


Figure 3: Authentication phase of our proposed scheme

6.2 Session Key Secrecy

Session key security means that at the end of the key exchange anyone cannot know the session key excepting the legal communication parties (the user and the server). In the proposed scheme the session key is computed in this way $SK = h_1(K\|r\|username\|X)$ where $X' = h(username\|s)P$ and $K = bh(h(PW\|a)\|username)S$. Since, PW, a and s are secret the session key cannot calculate by anyone except the server and the client. Therefore, our proposed protocol provides session key secrecy.

6.3 Server Spoofing Attack

The proposed scheme can resist against server spoofing attack. Assume that attacker Alice wants to impersonate the server and spoof user U, Alice has to computes $Auth_s = h_2(K\|W'\|r\|S\|K\|X')$. However, Alice does not have any information about a server secret key s. Then, she cannot compute K,SK and X'. Therefore, Alice cannot forge a valid CHALLENGE message.

6.4 User Impersonation Attack

Assume that attacker Alice wishes to connect to the server as legitimate user U. Alice has to prove its validity by forging two messages $REQUEST(username, V, W)$ and $RESPONSE(realm, Auth_u)$. While Alice need to know some secret information PW,a and X. Therefore, Alice is not capable to send the two validate messages. As result, our scheme can resist user impersonation attack.

6.5 Denning-Sacco Attack

The Denning-Sacco attack is when User or Server compromises an old session key and an attacker tries to find a long-term private key (e.g. user password or server private key) or other session keys.

In our scheme, the session key is calculated in this way $SK = h_1(K\|r\|username\|X)$ or $SK = h_1(K\|r\|username\|X')$. If an attacker obtains a session key, he will have to break the one-way hash function to get K,r and X or X. Then he have to know a secrete a and face the ECDLP if he want to guess password (PW') and verify the validity of $Auth'_s \stackrel{?}{=} h_2(K\|h(h(PW'\|a)\|username)bP_{pub}\|r\|SK\|X)$. So, the proposed scheme is secure against Denning Sacco attack.

6.6 Replay Attack

A replay attack is applied when an adversary reuse the information obtained in a protocol, trying to impersonate or deceive another legitimate participant. The following explain why the proposed protocol can resist to this attack. The adversary Alice may intercept the messages $REQUEST(username, V, W)$ and $RESPONSE(realm, Auth_u)$ from the User U and try to impersonate a legitimate user. However, she cannot calculate V, W and $Auth_u$ since she don't know server secret key. Alice has to face the ECDLP, if she wants get the correct one by guessing the secret key s from V or W. after replaying REQUEST or RESPONSE the server will detect the attack via comparing if $W \stackrel{?}{=} s^2(V - X')$ or $Auth_u \stackrel{?}{=} h_2(K\|W'\|r + 1\|SK\|X')$. Now, Suppose that Alive intercepts the message $CHALLENGE(realm, Auth_s, S, r)$ and try to replays it to impersonate the legal server. In order, to be authenticated by the user, Alice have to computes the value of $h_2(K\|W\|r\|SK\|X)$ using secret PW,X,a, $K = YS$ and $SK = h_1(K\|r\|username\|X)$. Since Alice don't have information about secret parameters she cannot computes a valid $Auth_s$. As result the proposed protocol withstand replay attack.

6.7 Stolen Verifier Attack

The stolen verifier attack means that an adversary steals the secret information from the server, like user's password. Then, the adversary uses it directly to masquerade as a legitimate user in a user authentication connection.

In the proposed scheme, any user's secret is stored in server database, so the attackers cannot obtain the user's secret information from server. Therefore, our proposed protocol is secure against stolen verifier attack.

6.8 Offline Password Guessing Attack

Password guessing attacks means that when an attacker interposes the communication between user and server

then he can guess the correct secret password by repeatedly guessing possible passwords and verifying the correctness of the guesses.

Suppose an attacker records all messages (REQUEST, CHALLENGE and RESPONSE) transmitted between user and server, then extract $username$, V, W, $realm$, $Auth_s$, S, r and $Auth_u$, and tries to guess the password PW^* and verifies its correctness. Since the attacker does not know any information about values of s, a and b he cant compute K, X, SK and $h(h(PW\|a)\|username)bP_{pub}$. Then, he cant verify the calculated V, W, $Auth_s$ or $Auth_u$.

If attacker steals user card he can get R, a and X, he must to know s to checks $h(h(PW^*\|a)\|username)s^{-1}P$. However, he will face ECDLP to extract s from $X = h(username\|s)P$. Therefore, our proposed scheme is safe against password guessing attack.

6.9 Man-in-the-Middle Attack

Man-in-the-middle attacks means that the attacker makes independent connections with the victims and relays messages between them, making them believe that they are talking directly to each other over a private connection. However, the entire conversation is controlled by the attacker.

In our protocol all messages are authenticated by server or user, to know their origin. In addition, at the end of authentication, the session key is shared between user and server, so the following messages will be encrypt using session key. To replay these messages, an attacker needs to know a session key. But, he cannot calculate it since he does not know s, a, X, PW and b. As result, our protocol is secure against Man-in-the-middle attack.

7 Performance Comparison

In this section, the performance of our proposed authentication and key agreement schemes is compared with other related authentication protocols. In this comparison a very lightweight operations like string concatenation operation, Exclusive-OR operation are not examined, because there computation cost is negligible. The notations used are illustrated as follows.

- T_h: Computational cost of one-way hash operation.
- T_{pm}: Computational cost of elliptic curve point multiplication.
- T_{pa}: Computational cost of elliptic curve point addition.
- T_{inv}: Computational cost of modular inversion.
- T_{EKs}: Computational cost of symmetric encryption.
- T_{DKs}: Computational cost of symmetric decryption.

In the registration phase of our protocol the user uses one hash function and the server computes $2T_h + 2T_{pm} + 1T_{inv}$. When the user need to be authenticated by server, it calculates $5T_h + 3T_{pm} + 1T_{pa}$ and the server computes $4T_h + 4T_{pm} + 1T_{pa}$. In the password changing phase the user computes $3T_h + 1T_{EKs} + 1T_{DKs}$ and the server computes $3T_h + 1T_{pm} + 1T_{inv} + 1T_{EKs} + 1T_{DKs}$.

In Table 2, we have illustrated the security performance of related schemes, as we can show ours protocol is secure against stolen verifier attack, Denning-Sacco attack, off-line password guessing attack, replay attack, man in the middle attack, server spoofing attack, insider attack, impersonation attack and denial of service attack. But, the Jiang et al. protocol is not secured against Server Spoofing attack and suffer from impersonation attack, and don′t provide security against man-in-the middle attack and Denning Sacco attack. So, we can say that our protocol is more secured if it is compared with Jiang et al.'s scheme.

According to Table 3, we can observe that our protocol reduce the number of T_{pm} from 4 to 3 in the authentication phase, if it's comparede with the same phase of Jiang et al. protocol. Hence, we can say that authentication phase of our protocol is faster than the same phase of Jiang et al.'s protocol. So, our protocol is more efficient than Jiang et al.'s protocol.

Table 2: Security comparison

Attacks	Zhang et al.	Tu et al.	Jiang et al.	Lin et al. [8]	Ours
Stolen Verifier	Yes	Yes	Yes	Yes	Yes
Denning Sacco	Yes	-	-	-	Yes
Password Guessing	Yes	Yes	Yes	Yes	Yes
replay	Yes	Yes	Yes	Yes	Yes
Man in the Middle	Yes	No	-	-	Yes
Server Spoofing	-	No	No	Yes	Yes
Impersonation	No	No	No	Yes	Yes
Mutual Authentication	Yes	Yes	Yes	Yes	Yes
Session Key Secrecy	Yes	-	Yes	Yes	Yes

8 Conclusion

In this article, we demonstrated that the protocol proposed by Jiang et al. cannot withstand server spoofing attacks. In order to overcome this weakness we proposed an efficient and secure SIP authentication scheme. By

Table 3: Computational comparisons between our protocol and related protocols

Phases	Entities	Zhang et al.	Tu et al.	Jiang et al.	Lin et al.	Ours
Registration Phase	User	$1T_h$	$1T_h$	$1T_h$	$1T_h$	$1T_h$
	Server	$1T_h + 1T_{pm} + 1T_inv$	$1T_h + 1T_{pm}$	$2T_h + 2T_{pm} + 1T_inv$	$2T_h + 1T_{Eks}$	$2T_h + 2T_{pm} + 1T_{inv}$
Authentication Phase	User	$6T_h + 4T_{pm} + 1T_{pa}$	$5T_h + 4T_{pm} + 1T_{pa}$	$4T_{pm}$	$6T_h + 3T_{pm} + 1T_{EKs} + 1T_{DKs}$	$5T_h + 3T_{pm} + 1T_{pa}$
	Server	$4T_h + 4T_{pm} + 1T_{pa}$	$5T_h + 3T_{pm}$	$4T_h + 4T_{pm} + 1T_{pa}$	$5T_h + 3T_{pm} + 2T_{EKs} + 2T_{DKs}$	$4T_h + 4T_{pm} + 1T_{pa}$
Password Changing Phase	User	$3T_h + 1T_{EKs} + 1T_{DKs}$	$3T_h + 1T_{EKs} + 1T_{DKs}$	–	$6T_h + 3T_{pm} + 1T_{EKs} + 1T_{DKs}$	$3T_h + 1T_{EKs} + 1T_{DKs}$
	Server	$3T_h + 1T_{pm} + 1T_{EKs} + 1T_{DKs} + 1T_{inv}$	$3T_h + 1T_{pm} + 1T_{EKs} + 1T_{DKs}$	–	$5T_h + 3T_{pm} + 2T_{EKs} + 2T_{DKs}$	$3T_h + 1T_{pm} + 1T_{EKs} + 1T_{DKs} + 1T_{inv}$
Total		$18T_h + 10T_{pm} + 2T_{pa} + 2T_{inv} + 2T_{EKs} + 2T_{DKs}$	$18T_h + 9T_{pm} + 1T_pa + 2T_{EKs} + 2T_{DKs}$	$12T_h + 10T_{pm} + 2T_{pa} + 1T_{inv}$	$25T_h + 12T_{pm} + 7T_{EKs} + 6T_{DKs}$	$18T_h + 10T_{pm} + 2T_{pa} + 2T_{inv} + 2T_{EKs} + 2T_{DKs}$

analyzing our scheme, we show that it is secure against various attacks and can provide many security services. Then, we conclude that our proposed protocol is suitable for Telephony over IP applications

References

[1] R. Arshad and N. Ikram, "Elliptic curve cryptography based mutual authentication scheme for session imitation protocol," *Multimedia Tools Appl*, vol. 66, no. 2, pp. 165–178, 2013.

[2] M. Azrour, M. Ouanan, and Y. Farhaoui, "SIP authentication protocols based on elliptic curve cryptography: Survey and comparison," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 4, no. 1, pp. 231–239, 2016.

[3] W. Diffie and M. Hellman, "New directions in cryptology," *IEEE Transaction on Information Theory*, vol. 22, no. 6, pp. 644–654, 1976.

[4] A. Durlanik and I. Sogukpinar, "Sip authentication scheme using ecdh," *World Enformatika Socity Transations on Engineering Computing and Technology*, vol. 8, pp. 350–353, 2005.

[5] J. Franks, P. Hallam-Baker, J. Hostetler, S. Lawrence, P. Leac, A. Luotonen, and L. Stewart, "Http authentication: Basic and digest access authentication," Tech. Rep. RFC 2617, June 1999.

[6] M. Handley, H. Schulzrinne, E. Schooler, and J. Rosenberg, "SIP: Session initiation protocol," Tech. Rep. RFC 2543, Mar. 1999.

[7] H.H. Hilinc, Y. Allaberdiyev, and T. Yanik, "Efficient ID-based authentication and key argeement protocols for the session initiation protocol," *Turkish Journal of Electrical Engineering and Computer Sciences*, vol. 23, pp. 560–579, 2015.

[8] H.Lin, F. Wen, and C.Du, "An anonymous and secure authentication and key agreement scheme for session initiation protocol," *Multimed Tools Appl*, vol. DOI 10.1007/s11042-015-3220-2, 2016.

[9] H. Huang, W.Wei, and G. E. Brown, "A new efficient authentication scheme for session initiation protocol," in *The 9th Joint Conference on Information Sciences*, 2006.

[10] A. Irshad, M. Sher, E. Rehman, ChS. Ashraf, MU. Hassan, and A. Ghani, "A single round-trip sip authentication scheme for voice over internet protocol using smart card," *Multimed Tools Applications*, 2013.

[11] Q. Jiang, J. Ma, and Y. Tian, "Cryptanalysis of smart-card-based password authenticated key agreement protocol for session initiation protocol of zhang et al," *International Journal of Communication Systems*, vol. 28, no. 7, 2014.

[12] H. Jo, Y. Lee, M. Kim, S. Kim, and D. Won, "Offline password guessing attack to yangs and huangs authentication schemes for session initiation protocol," in *The 5th International Joint Conference on INC, IMS and IDC (NCM'09)*, pp. 618–621, Aug. 2009.

[13] N. Koblitz, "Elliptic curve cryptosystems," *Mathematics of Computation*, vol. 48, no. 177, pp. 203–209, 1987.

[14] F. W. Liu and H. Koenig, "Cryptanalysis of a sip authentication scheme," in *12th IFIP TC6/TC11 International Conference (CMS'11)*, pp. 134–143, 2011.

[15] J. Lopez and R. Dahab, "An overview of elliptic curve cryptography," Tech. Rep., June 2000.

[16] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, and E. Schooler, "Sip: Session initiation protocol," Tech. Rep. RFC 3261, June 2002.

[17] M. Stanek, "Weaknesses of password authentication scheme based on geometric hashing," *International Journal of Network Security*, vol. 18, no. 4, pp. 798–801, 2016.

[18] H. Tang and X. Liu, "Cryptanalysis of arshad et al.'s ecc-based mutual authentication scheme for session initiation protocol," *Multimedia Tools and Applications*, vol. 65, no. 3, pp. 165–178, 2013.

[19] J. L. Tsai, "Efficient nonce-based authentication scheme for session initiation protocol," *International Journal of Network Security*, vol. 8, no. 3, pp. 312–316, May 2009.

[20] H. Tu, N. Kumar, N. Chilamkurti, and S. Rho, "An improved authentication protocol for session initiation protocol using smart card," *Peer-to-Peer Networking and Applications*, vol. 8, no. 5, pp. 903–910, 2014.

[21] Y. Wang and X. Peng, "Cryptanalysis of two efficient password-based authentication schemes using smart cards," *International Journal of Network Security*, vol. 17, no. 6, pp. 728–735, 2015.

[22] J. Wei, W. Liu, and X. Hu, "Secure and efficient smart card based remote user password authentication scheme," *International Journal of Network Security*, vol. 18, no. 4, pp. 782–792, 2016.

[23] L. Wu, Y. Zhang, and F. Wang, "A new provably secure authentication and key agreement protocol for sip using ecc," *Computer Standards and Interfaces*, vol. 31, no. 2, pp. 286–291, 2009.

[24] Q. Xie, "A new authenticated key agreement for session initiation protocol," *International Journal of Communication Systems*, vol. 25, no. 1, pp. 47–54, 2012.

[25] C. C. Yang, R. C. Wang, and W. T. Liu, "Secure authentication scheme for session initiation protocol," *Computers and Security*, vol. 24, pp. 381–386, 2005.

[26] E. J. Yoon and K. Y. Yoo, "Cryptanalysis of nake protocol based on ecc for sip and its improvement," in *Second International Conference on Future Generation Communication and Networking Symposia*, 2008.

[27] E. J. Yoon and K. Y. Yoo, "Cryptanalysis of ds-sip authentication scheme using ecdh," in *International Conference on New Trends in Information and Service Science*, pp. 642–647, Aug. 2009.

[28] E. J. Yoon and K. Y. Yoo, "A new authentication scheme for session initiation protocol," in *International Conference on Complex, Intelligent and Software Intensive Systems (CISIS'09)*, pp. 549–554, 2009.

[29] E. J. Yoon, K. Y. Yoo, C. Kim, Y. S. Hong, M. Jo, and H. H. Chen, "A secure and efficient sip authentication scheme for converged voip networks," *Computer Communications*, vol. 33, no. 14, pp. 1674–1681, 2010.

[30] L. Zhang, S. Tang, and Z. Cai, "Efficient and flexible passwordauthenticated key agreement for voice over internet protocolsession initiation protocol using smart card," *International Journal of Communication Systems*, vol. 27, no. 11, p. 2691–2702, 2013.

[31] L. Zhang, S. Tang, and Z. Cai, "Cryptanalysis and improvement of password-authenticated key agreement for session initiation protocol using smart cards," *Secur Communication Networks*, 2014.

Biography

Mourade Azrour received his Master's degree in Computer Science and Distributed Systems in 2014 from Departments of Mathematics and Computer Science, Faculty of Science, University Ibn Zohr, Agadir, Morocco. He is currently a Ph.D. candidate of the Moulay Ismail University, Faculty of sciences and Techniques, Errachidia, Morocco.His research interests include Authentication Protocols, Computer and Network Security and Cryptography.

Yousef Farhaoui is an professor, Department of Computer Science in Faculty of sciences and Techniques, Moulay Ismail University, Morocco. Received his PhD degree in computer security from the University Ibn Zohr. His research interest includes computer security, Data Mining, Data Warehousing, Data Fusion etc..

Mohammed Ouanan was born in Morocco on 1971. He received the B.S. degree in applied Mathmatics and the PhD degree in applied Mathmatics, all from the Faculty of science, University of Fez, Morroco, in Morocco, in 1997, 1999, and 2002 respectively. Since JULAY 2007, he has been an Assistant Professor with the Department of Computer Sciences, Faculty of Sciences and Techniques Errachidia, Moulay Ismail University, Meknes, Morocco. His research interest includes security, indexed image, statistical signal processing, multidimensional signal modelling, and Mathmatics.

The Integrated Artificial Immune Intrusion Detection Model Based on Decision-theoretic Rough Set

Rui-Hong Dong, Dong-Fang Wu, Qiu-Yu Zhang
(Corresponding author: Qiu-Yu Zhang)

School of Computer and Communication, Lanzhou University of Technology
No. 287, Lan-Gong-Ping Road, Lanzhou 730050, China
(Email: zhangqylz@163.com)
(Received Aug. 16, 2016; Revised and accepted Dec. 7 & 25, 2016)

Abstract

The intrusion detection methods used in the industrial control network generally have a higher false positive rate. Considering this issue and improving the detection performance of intrusion behaviors, an integrated artificial immune intrusion detection model based on decision-theoretic rough set was proposed in this paper. Firstly, by the approach of decision-theoretic rough set attributes reduction algorithm (DTRSA), attributes reduction was finished. And the rule set was obtained from the training data which has the binary string form. Secondly, taking into consideration of the negative selection algorithm (NSA), the rule set produced the corresponding detector sets. Vaccine mechanism was added into the model. Finally, real time dendritic cell algorithm (rtDCA) analyzed the environment and antigen information. The antigen matching threshold was obtained. Considering the intrusion behaviors and antigen matching threshold, the dynamic increases of rule set was achieved. Experimental results show that the proposed model obtained the lower false positive rate (FP) and the true positive rate (TP) reached to 95.5%. And both known and unknown intrusion detections had the high performance.

Keywords: Decision-theoretic Rough Set; Detector; Integrated Artificial Immune; Intrusion Detection System; Rule Set

1 Introduction

Industrial control system (ICS) is widely used in many national critical infrastructures. According to the statistics, more than 80% national critical infrastructures use ICS to achieve the automation of the industrial production. Therefore, the security of ICS can affect the national security and economy development directly. At the beginning of ICS development, the design of industrial control network didn't consider the security requirements. Following the disappearances of the network physical closure, the "Stuxnet" and "Flame" security events happen on the industrial control area. The security situation of industrial control network becomes worse. The main information security problem of ICS is closely related to the security of industrial control network. As the vital technique in the network security, intrusion detection approaches in industrial control network attach more focus from the researchers.

For the improvement of intrusion detection system (IDS), many approaches included probability statistics, neural network, support vector machine, genetic algorithm and artificial immune system [7, 10, 11], were proposed to research the network intrusion detection. According to the different intrusion detection objects, intrusion detection methods are divided into two types. One is named as anomaly detection which is used in the detection of the unknown intrusion. And, the other is called misuse detection that is used to detect the known intrusion. Both of them have advantages and disadvantages [14]. In [6], the simulation experiments and analysis of the integrated scheme for the anomaly and misuse detection were achieved. By the combination of structural features and behavioral characteristics two detection measures, the integrated system obtained the good detection performance. But, the experimental data has the high dimension of the condition attributes, which leads to the high time complexity. In this condition, for the real-time of system, the attributes reduction approach was recommended. According to [19], the above mentioned integrated scheme for two methods was also adopted. By the rough set algorithm (RSA), the rule set was obtained from the training data. And the corresponding detectors which actively participated in the intrusion detection were produced. Using vaccine mechanism, the producing process of detectors was optimized. But, as Figure 4 shown, comparing with the rough set, the decision-theoretic rough

set has a better characterization degree of domain. The characterization speed of domain is also faster.

In the artificial immune intrusion detection researches, Forrest, Castro and Greensmith [3, 8, 9] make more contributions. In [18], an improved artificial immune intrusion detection method was introduced. To obtain the superior antibodies, by the using of rough set and fuzzy set, a scheme of antibody based on the rough set was proposed. And the speed of intrusion detection was kept. According to [17], for the low detection performance of clone selection algorithm (CSA), a features selection method was used in the improvement of CSA. With the clone and mutation of some objects which have excellent characteristics, the classification ability of classifiers and algorithm performance were improved. In [12], the negative selection algorithm (NSA) included the fixed and variable r-contiguous bits two kinds matching methods, was introduced. The performance of NSA was improved via the real-time adjustment of binary string matching length. According to [15], in the process of features selection, using the filter approach, the result of attributes reduction was input into the dendritic cell algorithm (DCA). As the experiment shown, the performance of DCA was improved and the requirement of calculation ability was low.

Considering the shortcomings of the above researches, an integrated artificial immune intrusion detection model based on decision-theoretic rough set (DTRSIAI-IDM) was proposed in this study. By the DTRSA, the attributes reduction of experimental data was achieved. And the complexity of condition attributes was efficiently decreased. Meanwhile, self and nonself rule set was obtained from the training data. The corresponding nonself detectors which were used in the misuse detection were produced. To overcome the disadvantages of randomly producing detectors, vaccine mechanism was added into the NSA. Then, the real time dendritic cell algorithm (rtDCA) real-timely captured the antigen and environment information. The abnormal judgment was made via the computing of antigen abnormal index and matching threshold. To improve the speed of intrusion detection, the abnormal behaviors real-timely gave feedbacks to the rule set and detector sets. Finally, as the simulation experiments shown, comparing with the traditional rough set algorithm, the DTRSA got a better description of the positive domain and a lower complexity of experiment data. And the quality of detectors could be guaranteed. By the above proposed integrated scheme, the artificial immune model got a high true positive rate and a lower false positive rate.

The rest of the paper is organized as follows: Section 2 introduces the model and workflow of the integrated artificial immune intrusion detection approach. Section 3 describes all the algorithms included in the model and analyzes the complexity. In Section 4, firstly, it introduces the preprocessing of the data and analysis of the detectors length. Secondly, it discusses the merits and demerits of rough set (RS) and decision-theoretic rough set (DTRS). Thirdly, it analyzes and compares the detection performance of the proposed model. Finally, we conclude our paper in Section 5.

2 DTRSIAI-IDM

According to the research of Common Intrusion Detection Frame (CIDF) [19], as Figure 1 shown, an integrated intrusion detection model is designed. The model includes three function modules: rule set module (thymus), rule matching module (tissue) and analysis center module (lymph gland). By the preprocessing of KDD99 data, the experimental data translates into binary string form. Then, the data set, as the information of antigen, is input into the rule matching module and participates in the detection.

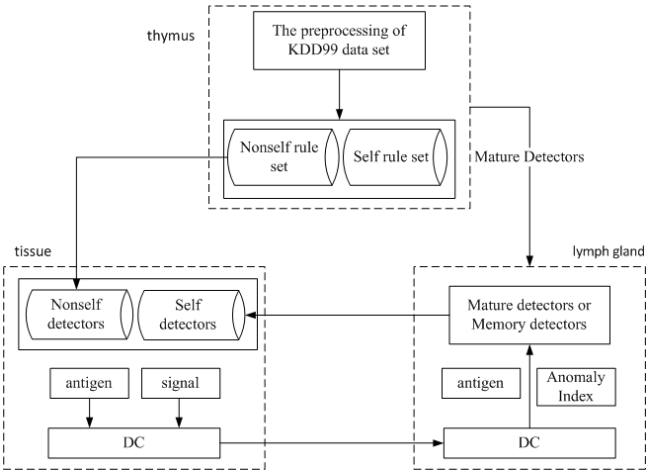


Figure 1: DTRSIAI-IDM integrated model

DTRSIAI-IDM includes four algorithms: integrated artificial immune algorithm based on decision-theoretic rough set (DTRSIAIA), DTRSA, NSA and rtDCA. By the using of the DTRSA approach, the data is processed and the rule set is obtained. According to the rule set, NSA produces the detector sets which meet the demands of detection system. For the undetected antigens, dendritic cell model captures the antigen and environment information. According to the metastasis threshold of dendritic cells, the states of dendritic cells are real-timely updated. And the match threshold is also computed. Then, the rule matching module receives the feedbacks from the anomaly detections. There is the general workflow of the model, as shown in Figure 2.

3 The Proposed Model

3.1 DTRSIAIA

In [6], the conception of matching threshold was mentioned. In the integrated artificial immune system, combining the dynamic anomaly index of antigen, the ap-

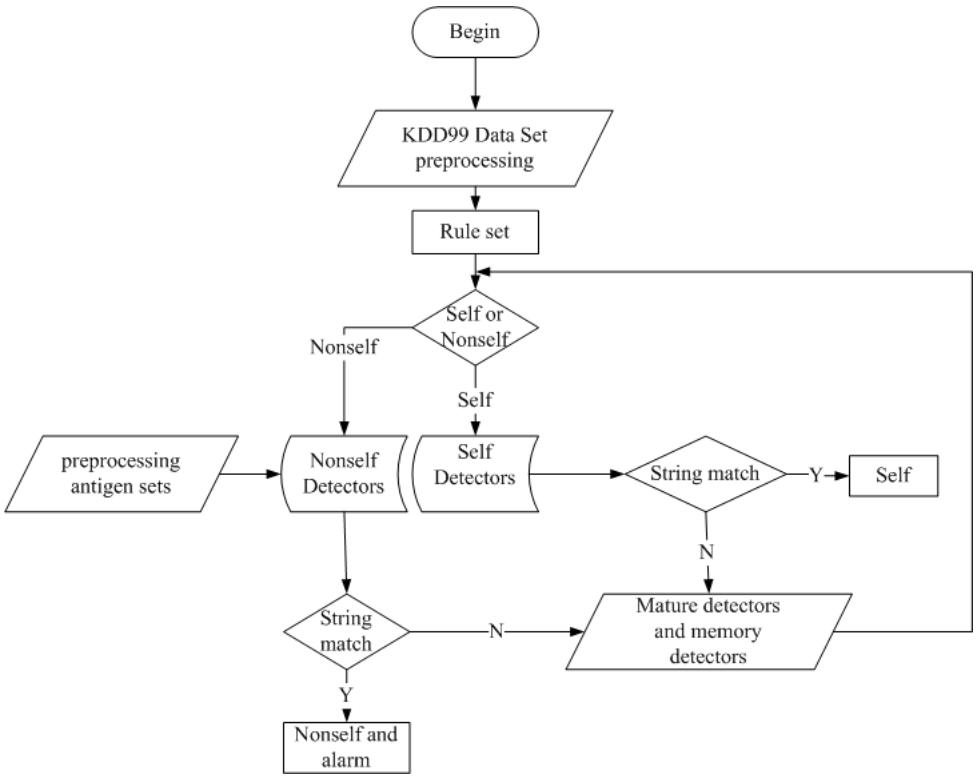


Figure 2: The workflow of the DTRSIAI-IDM

proach of computing matching threshold between detectors and antigen is proposed.

$$Y_\alpha = (L - \varepsilon)e^{-a(X_\alpha - \delta)}. \quad (1)$$

where the Y_α is the matching threshold of α type antigen. And, a is a constant. For the memory detectors, $a = -1$. In rtDCA, δ is the anomaly threshold. L is the length of detectors. And, $L = 24$. In the binary space, ε is self-radius which decides the coverage area of detectors. According to the Equation (2), when the anomaly degree is higher, the matching threshold is lower. And, when the anomaly degree of antigen is decreasing, the matching threshold is increasing. X_α is the dynamic anomaly index of α type antigen.

$$X_\alpha = \frac{m_a}{\sum_{i=1}^{A} A_i}. \quad (2)$$

where α is the antigen set which has the same value. And, the number of α type mature antigens is m_a. A_i is the total transformation number of i type antigens. The total number of all the types is A.

3.2 DTRSA

As the Ref. [16] describes, the study introduces the follow definitions: decision-theoretic rough set, α-positive domain, positive domain reduction and α-positive domain global significance.

Algorithm 1 DTRSIAIA

```
Input: Antigen and preprocessing information flow
Output: The detected anomaly antigens
while input data do
    if Semi ≤ Mat then
        Add the input antigens to the self-set
    end if
    Check and produce the detectors (NSA)
    Using rtDCA to deal with the antigens and the environment information
    According to Equation (2), real-timely analyze the input data and obtain the dynamic anomaly index
    According to Equation (1), when we know the anomaly index, computing the matching threshold
    Using detectors to detect the antigens
    Produce the alarm signals
end while
```

Definition 1. *(Decision-theoretic Rough Set) A decision-theoretic table is following tuple:*$DT = \{U, A_t = \{C \cup D\}, \{V_a | a \in A_t\}, \{I_a | a \in A_t\}\}$, *where* $U = \{x_1, x_2, .., x_m\}$ *is a finite nonempty set of objects.* A_t *is a finite nonempty set of all condition attributes.* $A_t = C \bigcup D$ *, where* C *is the condition attributes set and* D *is the decision-theoretic set. And,* V_a *is a nonempty set of values of* $a \in A_t$*.* $I_a : U \to V_a$ *is an information function that maps an object in* U *to exactly one value in* V_a*.*

Definition 2. *(α-positive domain) Given a decision-*

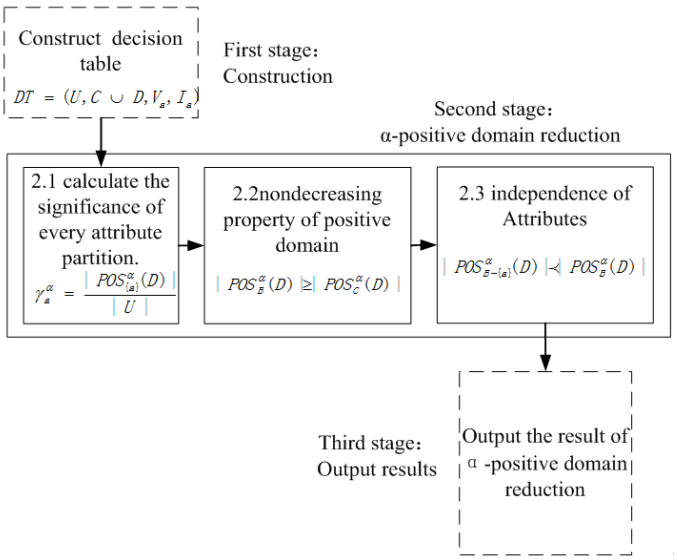


Figure 3: The attributes reduction model for DTRSA

theoretic table is following tuple: $DT = \{U, A_t = \{C \cup D\}, \{V_a|a \in A_t\}, \{I_a|a \in A_t\}\}$. *The decision-theoretic rough set is related with decision attributes and condition attributes. The* α*-positive domain is defined as:*

$$POS_B^\alpha = \bigcup_{X \in \frac{U}{D}} \underline{apr_C^\beta}(X). \tag{3}$$

where $\underline{apr_C^\alpha}(X)$ *is the* α *approximations of* X, $\underline{apr_C^\alpha}(X) = \{x \in U|P(X|[x]_c) \geq \alpha\}$.

Definition 3. *(positive domain reduction) Given a decision-theoretic table is following tuple:* $DT = \{U, A_t = \{C \cup D\}, \{V_a|a \in A_t\}, \{I_a|a \in A_t\}\}$. $\alpha \in [0, 1]$, *when the attributes subset* $B \subseteq C$ *has the following characteristics:*

1) *Positive domain no decreasing property:* $|POS_B^\alpha(D)| \geq |POS_C^\alpha(D)|$

2) *Independence of attributes:* $\forall \alpha \in B, |POS_{B-\{a\}}^\alpha(D)| < |POS_B^\alpha(D)|$

Definition 4. *(*α*-positive domain global significance) Given a decision-theoretic table is following tuple:* $DT = \{U, A_t = \{C \cup D\}, \{V_a|a \in A_t\}, \{I_a|a \in A_t\}\}$.*Let the condition probability threshold* $\alpha \in [0, 1]$, $a \in C$, *a is one attribute. The* α*-positive domain global significance is defined as follow:*

$$\gamma_a^\alpha = \frac{|POS_B^\alpha(D)|}{|U|}. \tag{4}$$

Attention, in reduction algorithm, there is one low time complexity belongs to positive domain global significance. In DTRSA, a heuristic reduction algorithm is adopted. The α-positive domain global significance, as the heuristic function, is selected to compute positive domain reduction. As shown in Figure 3, that is the basic thought of the DTRSA.

Algorithm 2 DTRSA

1: Input: The decision-theoretic table $DT = \{U, A_t = \{C \cup D\}, \{V_a|a \in A_t\}, \{I_a|a \in A_t\}\}$ and the *condition probability threshold* $\alpha \in [0, 1]$
2: Output: α attributes reduction
3: Preprocessing the *positive domain* attributes set $r = \emptyset$ $M \times N$
4: According to Equation (3), computing the *positive domain* of attributes. According to Equation (4), computing the *significance* of every condition attributes.
5: Obtain the global significance list γ, and show in descending order.
6: **while** $|POS_B^\alpha(D)| < |POS_C^\alpha(D)|$ **do**
7: Let $a = P(i)$, add a to α-positive reduction attributes set,$R = R \bigcup \{a\}; i = i + 1$
8: **end while**
9: **while** R does not meet this property **do**
10: $\forall a \in R$,
11: **if** $|POS_{B-a}^\alpha(D)| \geq |POS_B^\alpha(D)|$ **then**
12: $R = R - \{a\}$
13: **end if**
14: Using detectors to detect the *antigens*
15: **end while**

3.3 NSA

NSA includes two parts: producing the detectors and anomaly detection. By the DTRSA approach, the rule set is obtained. According to the affinity, NSA randomly produces some detectors which meet the self-radius. In [5], by the Equation (5), the affinity between detectors and self-samples is computed.

$$affinity = \sqrt{\sum_{i=1}^{L}(x_i - y_i)^2}. \tag{5}$$

where x_i is the i feature bit and y_i is the i feature bit.

When the detector sets cannot meet the demands of model, NSA producing detectors randomly and computing the affinity. When the affinity is less than matching threshold, the antigen sample is abnormal.

Algorithm 3 DTRSA(Produce the detectors)

```
Input: Self set, Self-radius, Detector length and the number of required detectors
Output:Detectors set
while the number of current detectors ≤ the number of required detectors do
    Randomly produce some fixed length detectors
    for every self-sample do
        According to Equation (5), computing the affinity between detectors and self-sample
        if candidate detectors are not in the self-radius range then
            Add the candidate detectors into the detectors set
        end if
    end for
end while
```

Algorithm 4 DTRSA(Anomaly detection)

```
Input: Anomaly detection output the unknown antigens and self-detectors set
Output:Detection results
while input date do
    for every detector do
        Computing the affinity between antigen and detectors
        if affinity less than matching threshold then
            Abnormal antigen
        else
            Normal antigen
        end if
    end for
end while
```

3.4 rtDCA

DC can capture three kinds of information: PAMP signal, Danger Signal and Safe Signal. In [2], the preprocessing scheme of data was researched. Receiving the input information, by the Equation (6), dendritic cell produces the output information and the dynamic anomaly index. The output information includes: Co stimulation signal (Csm), Semi mature signal (Semi) and Mature signal (Mat). The equality about the information can be written as follows:

$$O_j = (1 + I_{in}) \sum_{i=0}^{I} \omega_{ij} S_i. \quad (6)$$

where S_i is input information and O_j is output information. I_{in} is inflammatory signal. Let ω_{ij} be the transformation weight. I is the number of signal types. S_0, S_1, S_2 are separately corresponding to PAMP signal, Danger Signal and Safe Signal. O_0, O_1, O_2 separately represent three signals: Csm, Semi and Mat. Table 3 describes the signal transformation rule.

Algorithm 5 rtDCA

```
Input: Antigen flow and preprocessing signal flow
Output: Dynamic anomaly index of antigens
(Tissue)
Initialization of dendritic cell population
while input the data do
    Update the antigen structure and signal matrix
    for dendritic cell do
        Capture and store antigen
        According to Equation (6), processing the signal
        if Csm ≥ Migration threshold then
            if Mat ≥ Semi then
                cell context = 1
            else
                cell context = 0
            end if
            Dendritic cell migrate into lymph gland
            Add a new dendritic cell into the tissue
        end if
    end for
end while
(T cell population)
for dendritic cell do
    Computing the anomaly index of every kinds antigen
end for
```

Dendritic cell not only capture the antigen and environment information, but produce the output information. Dependent on concentration of input information, the environment information can be obtained. Then, computing the antigen anomaly index and responding the abnormal feedbacks.

3.5 Algorithm Complexity Analysis

The analyses for the above discussed algorithms were made, as the Table 1 shown. In the DTRSA, let M be the number of condition attributes. The number of decision-theoretic attribute is 1. Let the training set be N_1. And, let the number of undetected antigens be N. After the misuse detection, N_2 is the number of antigens. And, let the number of self-detectors is K. Let K_1 be the number of nonself detectors. And L is the length of detectors. Let $R(R < M)$ to express the number of α-positive domain objects. The time complexity of DTRSA is $O((N_1+1)*2^{R+1}+2*(R+M)N_1)$. And the space complexity is $O((MN_1KL)$. Comparing with the RSA mentioned in [10], the time and space complexities of DTRSA are equivalent with RSA. In the NSA, the time complexity

is $O(LN_2K)$. And, the space complexity is $O((L+1)K)$. For the rtDCA, the time complexity is $O(MN_2^2)$ and the space complexity is $O(N_2(L+1))$. The time complexity of DTRSIAIA is $O(NL(N_2+NK))$ and the space complexity is $O(MN_1K_1L)$.

Table 1: Analyses of algorithm complexity

Algorithm	Time complexity	Space complexity
DTRSA	$O((N_1+1)*2^{R+1}$ $+2*(R+M)N_1)$	$O((MN_1KL)$
NSA	$O(LN_2K)$	$O((L+1)K)$
rtDCA	$O(MN_2^2)$	$O(N_2(L+1))$
DTRSIAIA	$O(NL(N_2+NK))$	$O(MN_1K_1L)$

4 Experiments and Analysis

In intrusion detection, considering the security problems of industrial control network, true positive rate (TP), false positive rate (FP) and the speed of detection are the main evaluating indicators for the intrusion detection system. By the above indicators, the real-time and effectiveness of the system can be analyzed. As the input data, KDD99 data participates in the experiments and verifies the performance of DTRSIAIA. The simulation experiment is finished in the windows 7 system.

4.1 Date Set

For intrusion detection systems, KDD 99 is a standard test data which includes connection and attack items. There are almost 38 kinds of attacks which include smurf, nmap and rootkit. The dimension of KDD99 data is 41. According to the demands of experiments, preprocessing of data can be finished. In [6], the preprocessing approach of the data was proposed. On the one hand, antigen expresses the structural features. On the other hand, signals represent the behavioral characteristics.

1) Antigen

In Table 2, transform the selected attributes into the corresponding binary strings and construct the new antigen.

2) Input Information

According to the information gain attributes selection approach [19], divide 10 selected attributes into three kinds of signals.

- PAMP Signal: attributes 25, 26, 29, 38 and 40
- Danger Signal: attributes 23 and 24
- Safe Signal: attributes 12, 31 and 32

Let x to express the value of attribute. When $x \in [m,n]$, this domain can be divided into PAMP signal or Danger signal. Otherwise, it is the Safe signal. The normalization of the data can be processed by the following equation.

$$f(x) = \left\{ \begin{array}{l} 0, x \in [0,m) \\ \frac{100x}{n-m}, x \in [m,n] \\ 100, x \in (n,+\infty) \end{array} \right\}. \quad (7)$$

For the attribute 12, the numerical range is $[0, 0.99]$. According to $[min, max]$, the numerical range is constructed. The mean value express the condition of this kind of signal.

3) Experimental Parameters

The sum of dendritic cells is 10. The range of migration threshold is $[50, 400]$. Let the abnormal threshold $\delta = 0.35$. For the approach of NSA, the length of detector is 24. For the memory detectors, let $a = -1$. And, for the general detectors, let $a = -1/2$. The following Table 3 is an advised threshold value.

Table 3: Threshold value

Weight	Csm Signal	Semi Signal	Mat Signal
PAMP Signal	2	0	2
Safe Signal	1	0	1
Danger Signal	2	3	-3

4.2 Length of Detector

In the experiment, comparing with the monotonicity of rough set, the decision-theoretic rough set has the non-monotonicity. Select 1,000 message records, increase one attribute every time and record the description of the positive domain. As Figure 4 shown, the sum of attributes is 10.

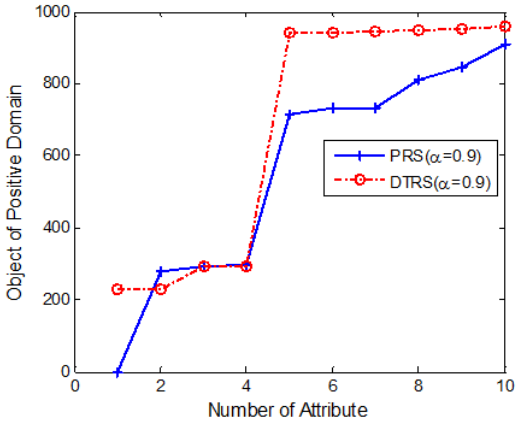


Figure 4: Attributes reduction description of the positive domain

In Figure 4, the dashed line describes the change situation of decision-theoretic rough set α positive domain.

Table 2: The construction of antigen

Attribute number	Transformation	Length
2	TCP,UDP and ICMP expressed by 00,01,10	2
3	Transform the number of the value into binary form	7
4	Transform the number of the value into binary form	4
5	Low, Middle, High and highest expressed by 00,01,10,11	2
6	Low, Middle, High and highest expressed by 00,01,10,11	2
12	Expressed by 0 or 1	1
28	Low, Middle and High expressed by 00,01,10	2
30	If the value equal to 1, expressed by 1, or 0	1
31	If the value equal to 1, expressed by 1, or 0	1
36	Low, Middle, High and highest expressed by 00,01,10,11	2

And, the dashed line does not meet the strict monotonicity. The solid line describes the change situation of the rough set. The dashed line is mostly on the upper of the solid line. And, account for the change situation of the positive domain, decision-theoretic rough set has a better description of the domain objects, as Figure 4 shown.

When $\alpha = 0.7, 0.8, 0.9$, Figure 5 describes the significance situation of attributes. When the value of α is decreases, the significance of attributes is increasing. And, in the decision-theoretic rough set, several attributes can mostly describe the whole positive domain which originally needs all condition attributes to represent. The message record includes 41 attributes. According to [4], the result of attributes reduction is $D = \{2\ 3\ 4\ 5\ 6\ 12\ 28\ 30\ 31\ 36\}$. Therefore, according to Table 2, the length of detector is 24. From the following Table 4, after the attributes reduction, the rule set is obviously decreasing. Add the rule sets into the self or nonself rule sets.

Table 4: The situation of DTRSA attributes reduction

Rule number	Before reduction	After reduction
PAMP Signal	8000	149
Safe Signal	2000	117

4.3 Self-radius

Researching the intrusion detection problems of industrial control network, the ROC (Receiver Operating Characteristic Curve) can be used to express the TP and FP of the detection. Adjust the self-radius continually, record the TP and FP [1]. Then, the ROC can be obtained, as Figure 6 shown.

Let $r = 0$, antigen completely matches detector sets. And, following the increasing of the self-radius, the TP and FP are also raising. When $r = 9$, the trend of the curve is balanced and the TP is higher than 0.9. For a higher TP and a lower FP, $r = 9$ can be selected to construct the detector sets.

4.4 Comparing the Performance

Compare the performance of DTRSIAIA with immune algorithm, rough set and support vector machine. The Table 5 lists the TP and FP of every approach.

Table 5: Comparing the performance

Name	True positive rate	False positive Rate
RSAI-IID [19]	0.9786	0.0268
SVM [20]	0.931	0.0081
RS-FSVM [13]	0.91	0.1424
rtDCA [6]	0.930	0.025
IAIS [6]	0.9691	0.0321
CSA [17]	0.996	0.1
NSA [12]	0.95	0.01
DTRSIAIA	0.955	0.020

As the Table 5 shown, comparing with other rough set algorithms, DTRSIAIA keeps the higher TP and lower FP. In [6], comparing with rtDCA, the FP is decreasing. And, for the IAIS approach, the proposed approach has a lower FP. In [19], the classifier method based on rough set has a higher TP and FP. The FP of proposed method is

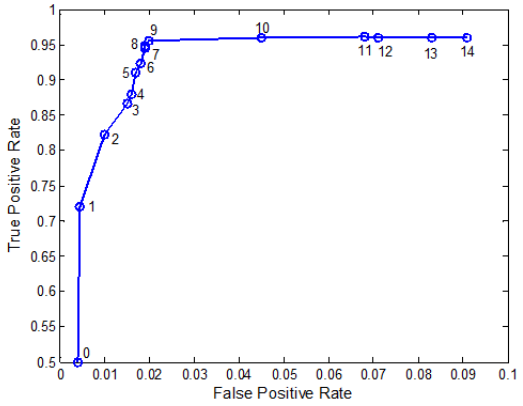


Figure 6: The ROC of the self-radius

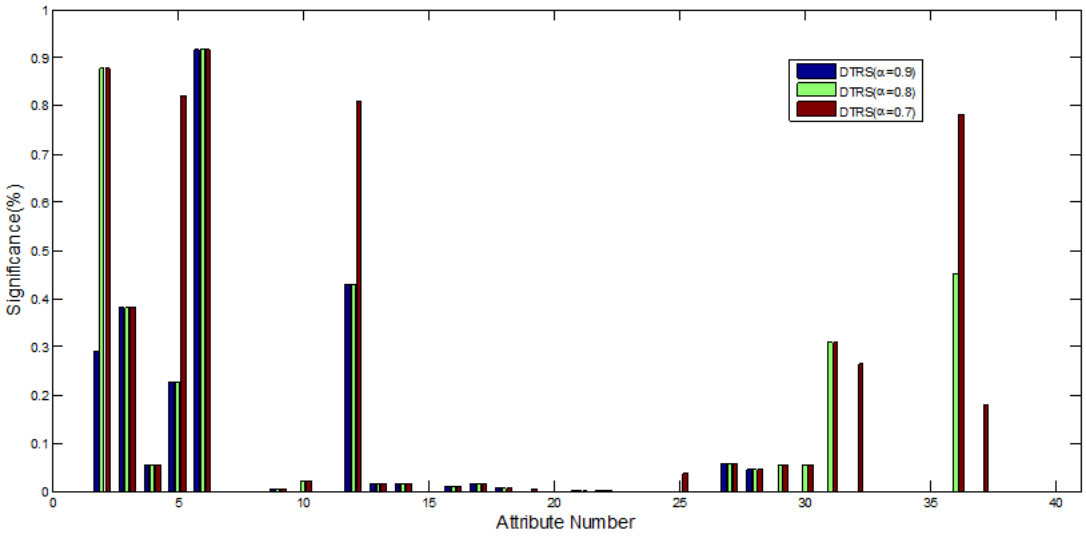


Figure 5: Significances of condition attributes

0.02 and TP is 0.955. In [17], an improved clone selection intrusion detection method was proposed. And the TP is 0.996, but the FP is 0.1. The FP is higher than the approach in this research. In [12], the FP is 0.01, but the TP is also lower. In [20], the FP is 0.0081 and TP is lower. In [13], the DTRSIAIA has a higher TP and lower FP.

According to Figure 6, $r \in [8, 10]$ is suitable for the simulation experiments. And the TP is higher than 0.9. Considering the analysis of algorithm in Section 3.5 , the DTRSIAIA has a lower complexity of algorithm. And, as Figure 4 shown, the DTRSA has a better description of the positive domain.

5 Conclusions

An integrated artificial immune intrusion detection algorithm based on decision-theoretic rough set was proposed. The effectiveness of the proposed intrusion detection method was proved, which deals with the intrusion detection problems in industrial control network effectively. Firstly, by the DTRSA approach, the attributes reduction was finished. Comparing with the RSA, the DTRSA is a better choice for the operation of attributes reduction. Secondly, overcoming the disadvantages of randomly producing detectors, the vaccine mechanism was added into the NSA. Then, the qualified detectors were produced. Finally, using rtDCA to analyze the antigen and environment information, rtDCA real-timely responded the feedbacks to the rule set. By this method, the real-time property was kept. In the experiment, the TP of DTRSIAIA is 0.955 and FP is 0.02. And, with the operation of DTRSA, the result of attributes reduction is obvious. The data complexity is decreasing. However, the time complexity of DTRSA is higher.

Next, the research will pay more attention and time to improve the performance of algorithm and prove the completeness.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (No. 61363078), the Natural Science Foundation of Gansu Province of China (No. 1310RJYA004). The authors would like to thank the anonymous reviewers for their helpful comments and suggestions.

References

[1] M. H. Aghdam and P. Kabiri, "Feature selection for intrusion detection systemusing ant colony optimization," *International Journal of Network Security*, vol. 18, no. 3, pp. 420–432, 2016.

[2] A. A. Al-Hasan and E. S. M. El-Alfy, "Dendritic cell algorithm for mobile phone spam filtering," *Procedia Computer Science*, vol. 52, pp. 244–251, 2015.

[3] L. N. De Castro and F. J. Zuben, "The clonal selection algorithm with engineering applications," in *Proceedings of the GECCO*, pp. 36–39, Las Vegas, USA, July 2000.

[4] A. Chaudhary, V. N. Tiwari, and A. Kumar, "A new intrusion detection system based on soft computing techniques using neuro-fuzzy classifier for packet dropping attack in manets," *International Journal of Network Security*, vol. 18, no. 3, pp. 514–522, 2016.

[5] W. Chen, X. Ding, T. Li, and T. Yang, "Negative selection algorithm based on grid file of the feature space," *Knowledge-Based Systems*, vol. 56, pp. 26–35, 2014.

[6] Y. Chen, Q. Zhang C. Feng, and C. Tang, "Integrated artificial immune system for intrusion detection," *Journal of China Institute of Communications*, vol. 33, no. 2, pp. 125–131, 2012.

[7] B. A. Fessi, S. Benabdallah, N. Boudriga, and M. Hamdi, "A multi-attribute decision model for intrusion response system," *Information Sciences*, vol. 270, pp. 237–254, 2014.

[8] S. Forrest, A. S. Perelson, L. Allen, and R. Cherukur, "Self-nonself discrimination in a computer," in *Proceedings of the IEEE Computer Society Symposium on Research in Security and Privacy*, pp. 202–212, Oakland, California, May 1994.

[9] J. Greensmith, U. Aickelin, and S. Cayzer, "Introducing dendritic cells as a novel immune-inspired algorithm for anomaly detection," in *Proceedings of the International Conference on Artificial Immune Systems*, pp. 153–167, Banff, Alberta, Canada, Aug. 2005.

[10] C. Guo and Y. Zhou, Y. Ping, Z. Zhang, G. Liu, and Y. Yang , "A distance sum-based hybrid method for intrusion detection," *Applied Intelligence*, vol. 40, no. 1, pp. 178–188, 2014.

[11] K. S. Anil Kumar and V. Nanda Mohan, "Adaptive fuzzy neural network model for intrusion detection," in *Proceedings of the Contemporary Computing and Informatics (IC3I'14), 2014 International Conference on. IEEE*, pp. 987–991, Mysore, India, Nov. 2014.

[12] A. Lasisi, R. Ghazali, and T. Herawan, "Negative selection algorithm: a survey on the epistemology of generating detectors," in *Proceedings of the First International Conference on Advanced Data and Information Engineering*, pp. 167–176, Kuala Lumpur, Malaysia, 2014.

[13] L. Li and K. Zhao, "A new intrusion detection system based on rough set theory and fuzzy support vector machine," in *Proceedings of the International Workshop on Intelligent Systems and Applications*, pp. 1–5, Wuhan, China, May 2011.

[14] S. Qing, J. Jiang, H. Ma, W. Wen and X. Liu, "Research on intrusion detection techniques: A survey," *Journal-China Institute of Communications*, vol. 25, no. 7, pp. 19–29, 2014.

[15] H. A. Le Thi, X. T. Vo A. V. Le, and A. Zidna, "A filter based feature selection approach in msvm using dca and its application in network intrusion detection," in *Proceedings of the the 6th Asian Conference on Intelligent Information and Database Systems*, pp. 403–413, Bangkok, Thailand, Apr. 2014.

[16] Y. Yao and Y. Zhao, "Attribute reduction in decision-theoretic rough set models," *Information Sciences*, vol. 178, no. 17, pp. 3356–3373, 2008.

[17] C. Yin, L. Ma, and L. Feng, "A feature selection method for improved clonal algorithm towards intrusion detection," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 30, no. 5, 2016.

[18] L. Zhang, Z. Bai, S. Luo, and G. Cui, "A dynamic artificial immune-based intrusion detection method using rough and fuzzy set," in *Proceedings of the Information and Network Security (ICINS'13)*, pp. 1–7, Beijing, China, Nov. 2013.

[19] L. Zhang, Z. Bai, S. Luo, K. Xie, G. Cui, and M. Sun , "Integrated intrusion detection model based on rough set and artificial immune," *Journal on Communications*, vol. 34, no. 9, pp. 166–176, 2013.

[20] H. Zhao, "Intrusion detection ensemble algorithm based on bagging and neighborhood rough set," *International Journal of Security and Its Applications*, vol. 7, no. 5, pp. 193–204, 2013.

Biography

Dong Rui-hong, vice researcher, worked at school of computer and communication in Lanzhou university of technology. His research interests include network and information security, information hiding and steganalysis analysis, computer network.

Wu Dong-fang received the BS degrees in Computer Science and Technology from Northwest University for Nationalities, Gansu, China, in 2015. Currently, he is studying for his masters degree at Lanzhou University of Technology. His research interests include industrial control network security.

Zhang Qiu-yu, researcher/PhD supervisor, graduated from Gansu University of Technology in 1986, and then worked at school of computer and communication in Lanzhou University of Technology. He is vice dean of Gansu manufacturing information engineering research center, a CCF senior member, a member of IEEE and ACM. His research interests include network and information security, information hiding and steganalysis, multimedia communication technology.

A New Mutuel Kerberos Authentication Protocol for Distributed Systems

Zakariae Tbatou, Ahmed Asimi, Younes Asimi, Yassine Sadqi, Azidine Guezzaz
(Corresponding author: Zakariae Tbatou)

Department of Mathematics, Faculty of Sciences, Ibn Zohr University, Agadir, Morocco
B.P 8106, City Dakhla, Agadir, Morocco
(Email: tbatou.zakariae@gmail.com)
(Received July 18, 2016; revised and accepted Sept. 25 & Oct. 25, 2016)

Abstract

In recent years, distributed systems, including cloud computing, are becoming increasingly popular. They are based on traditional security mechanisms that focus on access control policies and the use of cryptographic primitives. However, these mechanisms do not implement some more advanced security properties, including authentication policies. Kerberos V5, the most recent version, is a successful protocol that is designed to authenticate clients to multiple networked services. In this paper we propose a new mutuel Kerberos authentication protocol for distributed systems based upon Kerberos V5 and Diffie Hellman models. it is composed of three phases: 1) registration phase, based on the Diffie Hellman model, enabling the design and reliable exchange of client's authentication parameters to the authentication server side; 2) communication phase, based upon the two functions S2KexS () and DKexS (), which aims to the exchange of encryption keys and creates a secure the communication channel between client and server of services and 3) renewal phase for updating the client authentication parameters. Our security analysis and performance evaluation demonstrate that our scheme creates a secure channel to a more secure password exchange. Hence, it reduces the chance that a password will be guessed from the parameters stored or exchanged between client and authentication server, which make our proposed protocol efficient against dictionary and brute force attacks. The results proved by the behavior study show the success of our scheme and the easily of implementation. *Keywords: Authentication; Cloud Computing; Cryptographic Primitives; Diffie Hellman Model; Distributed Systems; Kerberos V5*

1 Introduction and Notations

Most authentication mechanisms are based only on password [8, 11, 20, 26]. In these regimes, the distant server maintains a table to record information of each user's password, and exploits them to verify the corresponding user privileges. However, although they are widely used in many applications in real life, authentication systems based on password suffer from several attacks [1, 3, 20, 27], such as dictionary attacks [23] brute force, steals data, Guessing Attacks [4, 14, 22, 27], etc. to respond to these issues, Kerberos V5 presents a strong protocol of network authentication for client/ server applications [13, 31]. It uses a (KDC), and tickets distribution center (TDC) [12, 24, 25]; in the sense that it never transmits passwords [21, 25, 31]. It exchanges encrypted messages with limited life by adding entities called tickets [12, 25]. All authentication requests are routed through the centralized KDC server [24, 25]. The latter defines a unique namespace for different clients [12]. In our approach, we assume that the communication between realms and service servers is based on Single Sign On.

In this paper, we begin by presenting the authentication dialog and the different cryptographic primitives for keys generation. In the third section, we propose a new communication scheme with a description of the three phases: 1)the registration phase based on Diffie Hellman model [7] and a dynamic salt generator (RGSCS) [2]); 2)the communication phase based upon the two functions S2KexS () and DKexS ().

The first function aims to generate a basic key from the footprint of the password and a dynamic salt per session. Based on this basic key, the second function is designed to generate encryption keys and 3)the renewal phase for updating the clients authentication parameters. The fourth section describes a behavioral study of the three phases with the use of regenerator of salts (RGSCS), dynamic and cryptographically secure. The five section studies the security analysis of our approach by evaluating the impact of its three phases on the robustness of the Kerberos V5 protocol. We end this paper with a conclusion. In all that follows, we denote by Table 1.

Table 1: Notations

Symbols	Meaning
C:	Client.
S:	Server of services.
KDC:	Key Distribution Center.
ID:	Identity of client.
ID_R:	Identity of releam.
Pwd:	User's password.
$\mathbb{N}$:	The set of natural numbers.
$salt_i$:	Dynamic pseudorandom sequence.
$\|\|$:	Concatenation.
==:	Comparaison.
mod:	Modulo operation.
$K_{x,y}$:	Session key shared between x and y.
$\{m\}K_x$:	m encrypted by the secret key.
$T_{x,y}$:	Ticket of x to use y.
$A_{x,y}$:	Authenticator of x for y.
$\mathbb{F}_p$:	Finite field of order a prime number p.
$\mathbb{F}_p^*$:	Cyclic multiplicative group of all non zero elements in $\mathbb{F}_p$ of order $p-1$.
$S(i)$:	$(i+1)^{th}$ binary string position of S.

2 Related Work

The modern Kerberos has undergone several major revisions. In each review, significant improvements have been made like scalability and security. The version 1 through 3 were used internally and as to version 4 was the first version distributed to the public was Kerberos V4, which has been limited in some nations due to the limitations of used encryption algorithms. These limitations made norms to evolve a new protocol that contains all the features presented in the Kerberos V4, with the addition of features such as extensible encryption types and more transparent authentication to create the version 5 of Keberos [13, 25]. After all these changes and with the development of computer system, Kerberos V5 still vulnerable against attacks such as attacks by brute force and dictionary. They still represent a real challenge for this protocol. These conclusions made thinking several researchers to propose solutions such as the use of asymmetric cryptographic primitives [17], in order to make the keys generation more reliable, or the introducing of new technologies such as smart card [16]. In this section, we present the communication phase based on two strong points: cryptographic primitives and tickets, and the various requests exchanged between a client and the KDC server to access a service.

2.1 Communication Dialogue

The communication dialogue in Kerberos V5 introduces three entities: a client, a centralized KDC server and a server of services. Authentication requests are routed through the centralized server KDC [12, 25] as described in Figure 1.

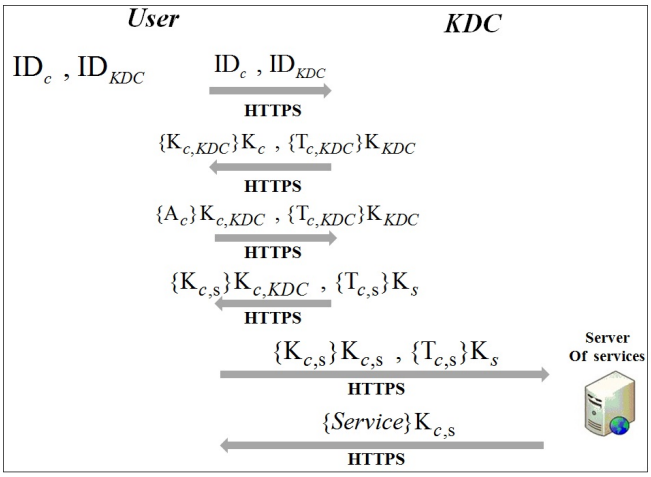


Figure 1: Description of Kerberos V5 queries

In Kerberos V5, the ticket distribution center acts as an intermediary of various requests exchanged between client and server of services to authenticate the client before access to the wanted service, based on two entities: tickets, which are used to authenticate client to the ticket distribution center and an authenticator to validate the client's identity to the server of services.

2.2 Cryptographic Primitives and Diffie Hellman Problem

Kerberos V5, in its communication phase, uses three encryption keys. Referring to [12, 25], the steps to generate these three keys are as follows:

- Regeneration of the basic key either by the random-to-key () function from a random bit string, or by the String-to-key() function from a password and a salt.
- Regeneration of these three keys associated to this based key by the key derivation function called Derived-key().

The Diffie-Hellman protocol is a method for two computer users to generate a shared private key with which they can then exchange information across an insecure channel. We refer to [15, 19] and we deduce the following results.

Definition 1. *A primitive element of $\mathbb{F}_p$ is a generator of a cyclic units group $\mathbb{F}_p^*$.*

Definition 2. *The Diffie Helman problem is the following : given a prime number, a primitive element g of $\mathbb{F}_p$, and $g^a \ mod \ p$ and $g^b \ mod \ p$, find $g^{ab} \ mod \ p$.*

Definition 3. *The generated Diffie Helman problem is the following : given a finite cyclic group G, a primitive element g of G, and group elements g^a and g^b, find g^{ab}.*

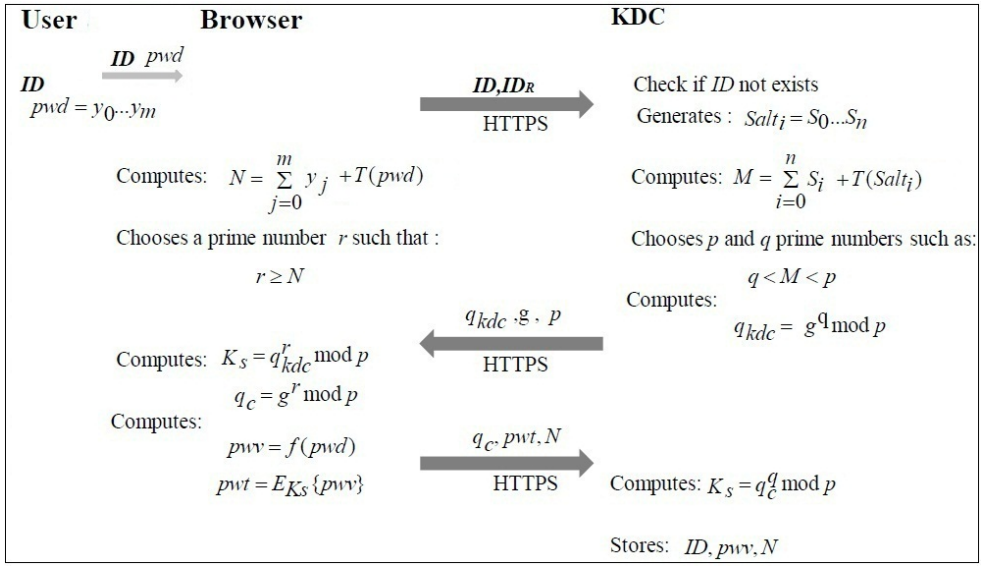


Figure 2: Description of the new registration phase

According to Definitionss [1, 2, 3] the use of Diffie-Hellman causes some problems at the implementation level: 1) the problem to determine with effective way the primitive elements of a finite field [5, 6]; 2) the difficulty of implementation specifically the complexity of the computation time and performance especially in systems require the notion of time [10], and 3) the synchronization problem relatively to the time system. In our approach, we have took into consideration these problems with using the Diffie-hellman principle by the choice of a finite field $\mathbb{F}_p$ with $p = 2^n + 1$ and its primitive elements which are the form 3^{2m+1} modulo p for all $m \in \mathbb{N}$.

3 Description of Our Approach

The scheme of our conception consists of three entities: 1) Kerberos client that belongs to the KDC realm; 2) Browser that supports HTTPS for a more secure data exchange and cryptographic primitives, virtualisation functions and hash functions; 3) KDC server, which is the key distribution center, provides symmetric cryptographic primitives, virtualization functions and hash functions. It is composed of a basic three storing identification parameters assigned to each user identified by ID. These parameters are used to authenticate users during the communication phase and can be easily changed in the renewal phase, and which are successively rated:

ID	*pwv*	*N*

- *ID*: User identification.
- *pwv*: Footprint of the password. In our proposal, it will be used for generating keys encryption / decryption to ensure:
  1) The user identification during communication and renewal phases.
  2) The confidentiality of messages exchanged between users and the KDC server.
  3) The confidentiality of the new password chosen by users in the renewal phase.
- *N*: Integer number regenerated from the password.

3.1 Conception of Our Approach

Our authentication scheme is based on three phases: registration, communication and renewal phases.

3.1.1 Registration Phase

This phase, regenerates its own authentication settings using a username and password of user not shared between the browser and the KDC, as described in Figure 2.

In this process, the KDC generates for each user three authentication parameters, based on a salt generator which generates different salts for each user. At the client side, each client must have a valid password and a unique ID that does not exist in the database. The dialogue of the registration phase is described as follows:

- The client sends its ID and ID_R of releam Which he wants to register to the KDC.
- The KDC server checks the existence of the *ID*.
  - If it exists, it returns an error message.
  - Otherwise, it
    - Generates a first *salti*.
    - Calculates M that is equal to the sum of the bits of *salti* and *salti* length.
    - Chooses two prime numbers p and q with p upper than M and q lower than M.

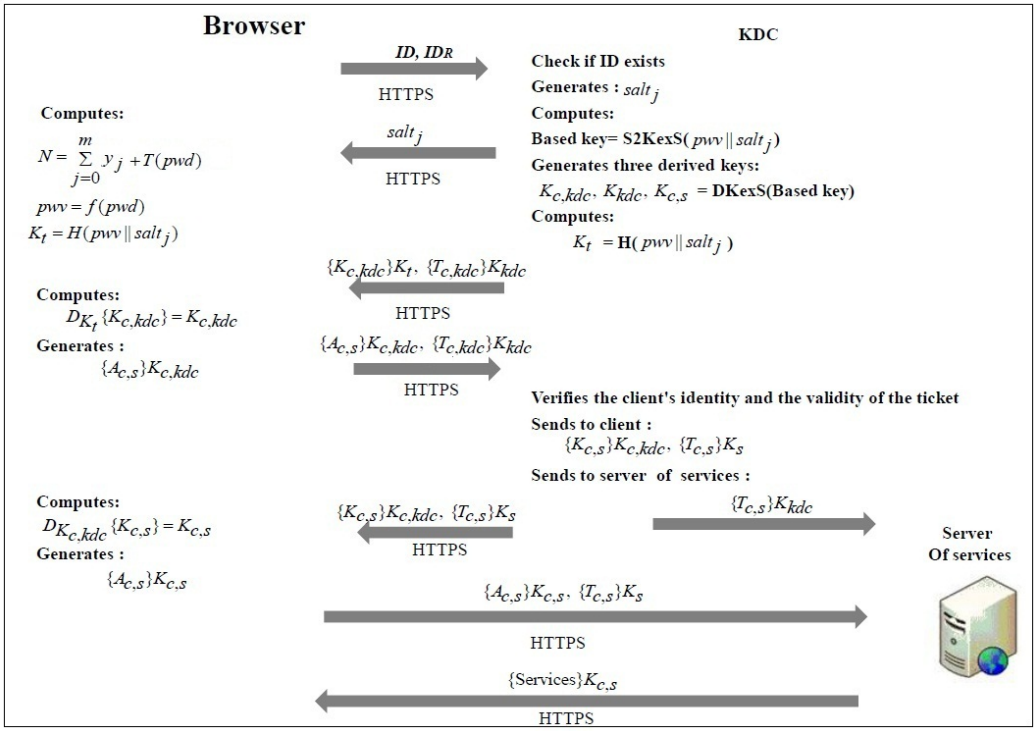


Figure 3: Description of the new communication phase

* Chooses a number g in order that g is a divisor of M.
* Calculates $q_{kdc} = g^q \bmod p$
* Sends q_{kdc}, g and q to the client.

- The client:
  - Calculates N, which is equal to the sum of the password bits and the password length.
  - Chooses a prime number r upper than N.
  - Calculates $q_c = g^r \bmod p$.
  - Calculates $pwv = f(pwd)$ where f is a virtualization function.
  - Calculates the key $K_s = q_{kdc}^r \bmod p$.
  - Sends q_c, $\{pwv, N\}K_s$ to KDC .
- The KDC server:
  - Calculates the key $K_s = q_c^q \bmod p$.
  - Decrypts $\{pwv, N\}K_s$ and obtains pwv, and N.
  - Stores ID, pwv and N.

3.1.2 Authentication and Identification Phase (Communication)

In this phase each user must prove his identity (ID) to the KDC server, specifically the KDC that must authenticate the user because the Kerberos system is based on a trusted third party [8, 20, 24, 25]. For this reason the client sends his ID and the ID_R of its realm (authentication without sending the password [12, 21]) to the KDC server. it last checks the ID in the database, if it exists, the KDC generates a basic key from client authentication parameters stored in the database if not it returns a message error. The key generation has been enhanced by new features $S2KexS$ and $DKexS$ [30] to make the generation key dynamic.

The dialogue of the communication phase is described as follows (see Figure 3):

- Client :
  - Sends his ID and the ID_R of its releam to KDC server.
  - Calculates N.
  - Enters the password pwd and calculates $pwv = f(pwd)$.
- The KDC server:
  - Verifies his own ID_R and checks the existence of user ID, if doesn't exist the KDC sends an error message, otherwise.
  - Calculates a based key with the function $S2KexS$ from pwv stored in the database and a new regenerated $salt_i$
  - Calculates three derived keys $K_{c,kdc}$ K_{kdc} and $K_{c,s}$ with the key derivation function from the based key $DKexS(besed\ key)$.
  - Calculates a temporary key $K_t = H(pwv||salt_i)$.
  - Encrypts $K_{c,kdc}$ with K_t.

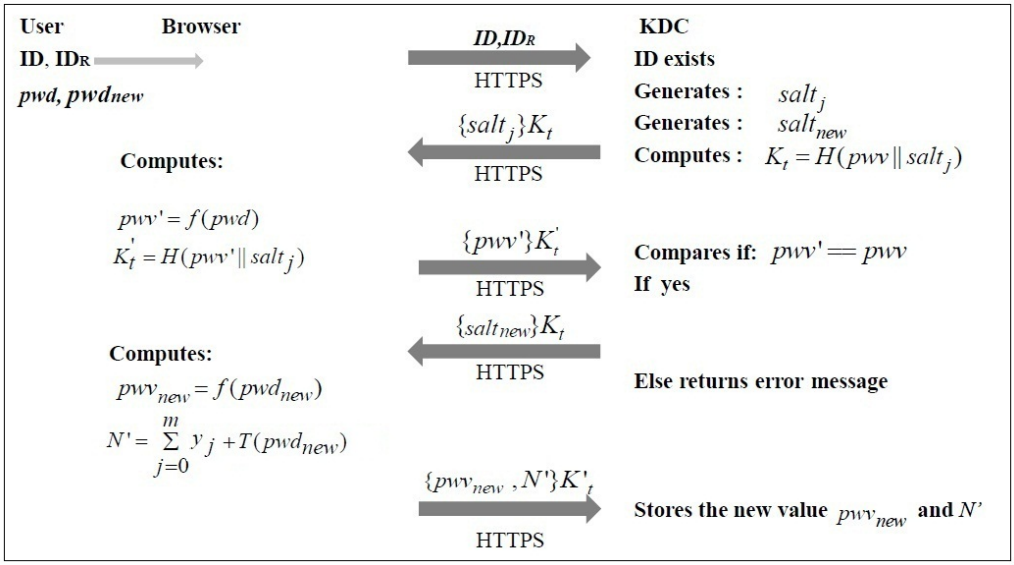


Figure 4: Description of the renewal phase

- Encrypts $T_{c,kdc}$ with K_{kdc}.
- Sends $\{T_{c,kdc}\}K_{kdc}$, $\{K_{c,kdc}\}K_t$ and $salt_i$ to the client.

- The client:
  - Calculates $K_t = H(pwv||salt_i)$.
  - Finds the $K_{c,kdc}$.
  - Generates an authenticator $A_{c,kdc}$ which contains the requested service, the calculated number N and others authentication parameters.
  - Sends $\{A_{c,kdc}\}K_{c,kdc}$ and $\{T_{c,kdc}\}K_{kdc}$ to KDC.
- The KDC server:
  - Finds $A_{c,kdc}$ and $T_{c,kdc}$.
  - Checks the validity of the ticket time and the client's identity from $A_{c,kdc}$ parameters.
  - Encrypts $K_{c,s}$ with $K_{c,kdc}$.
  - Creates a ticket $T_{c,s}$ that will be shared between the client and the server of services.
  - Encrypts $T_{c,s}$ with K_s.
  - Sends $\{T_{c,s}\}K_s$ and $\{K_{c,s}\}K_{c,kdc}$ to the client.
- Client:
  - Decrypts $\{K_{c,s}\}K_{c,kdc}$ and gets $K_{c,s}$.
  - Generates $A_{c,s}$ which contains the service and client authentication parameters and encrypts it with $K_{c,s}$.
  - Sends $\{A_{c,s}\}K_{c,s}$ and $\{T_{c,s}\}K_s$ to server of services.
- Server of services:
  - Decrypts $\{T_{c,s}\}K_s$ with K_s and checks the client's identity and the validity of the ticket time..
  - If the identification is successful, it encrypts the requested service with the key $K_{c,s}$ and sends the message to the client. Otherwise the server of services sends an error message.

3.1.3 Renewal Phase

This phase allows the renewal of client authentication parameters. It represents the most important phase especially for new users, because it enables the exchange of the new parameters in an environment more secure than the registration phase. In this phase, we must ensure the identity of the user, mutual authentication and validity of the new password as described in Figure 4.

In this phase, it should be noted that the client is already logged into his session so the encryption keys are already shared. So the client must enter his old password to validate the authentication parameters with the KDC server, then he enters his new password.

- Client sends his ID and ID_R of releam to the KDC.
- The KDC server:
  - Verifies his ID_R and checks the existence of ID. if doesn't exist, it returns an error message, otherwise:
  - Generates two new salts $salt_{new}$ and $salt_j$.
  - Calculates K_t which is equal to hashed pwv concatenated with $salt_j$.
  - Sends $salt_j$ to the client.
- Client:
  - Enters his pwd.

Figure 5: Behavioral study of the registration phase for three iterations

  - Calculates $pwv' = f(pwd)$ and $K_t = H(pwv'||salt_j)$.
  - Sends $\{pwv'\}K_t$ to the KDC.

- The KDC server:
  - Decrypts $\{pwv'\}K_t$ and compares pwv' with pwv:
  - If $pwv' == pwv$ then $K_t == K'_t$ therefore the server sends the $salt_i$ encrypted with K_t to the client. Otherwise it sends an error message asking him to send his ID.

- Client:
  - Gets the $salt_{new}$ using K_t.
  - Calculates the new value of N.
  - Calculates $pwv_{new} = f(pwd_{new})$.
  - Sends $\{pwv_{new}, N'\}K_t$ to KDC.

- Server KDC:
  - Decrypts $\{pwv_{new}, N'\}K_t$ with K_t.
  - Updates the server database with new values of pwv_{new} and N'.

4 Behavior Study

After presenting the purpose of the integrated regenerator RGSCS [2] in different phases, in this section, we focus on behavioral study of registration and communication phases to test the influence of RGSCS on our proposal. We begin then by studying the impact of salts regenerated by RGSCS for a low and redundant given password 'aaaaaa', on the generation of encryption keys in the registration phase, and generation of session key in the communication phase by studying the correlation of the generated binary sequences.

4.1 Behavioral Study of the Registration Phase

In the figure Figure 5, we have implemented the registration phase using PHP 5 to program the various functions; DES mode CBC as an encryption algorithm, a virtualization function based on the dynamic rotation [1], the RGSCS generator and the Diffie Helmman protocol based on a finite field $\mathbb{F}_p$ of characteristic $p = 2^n + 1$ [7] having as primitive elements the numbers 3^{2m+1} modulo p for all $m \in \mathbb{N}$. The hardware used in our experiments is a AMD $E-300$ CPU 1.3 GHz and 4Go as RAM running under Windows 7. In our case, we took the number 3 as a small primitive element of $\mathbb{F}_p$ to evaluate our results even this primitive element make our protocol dynamic

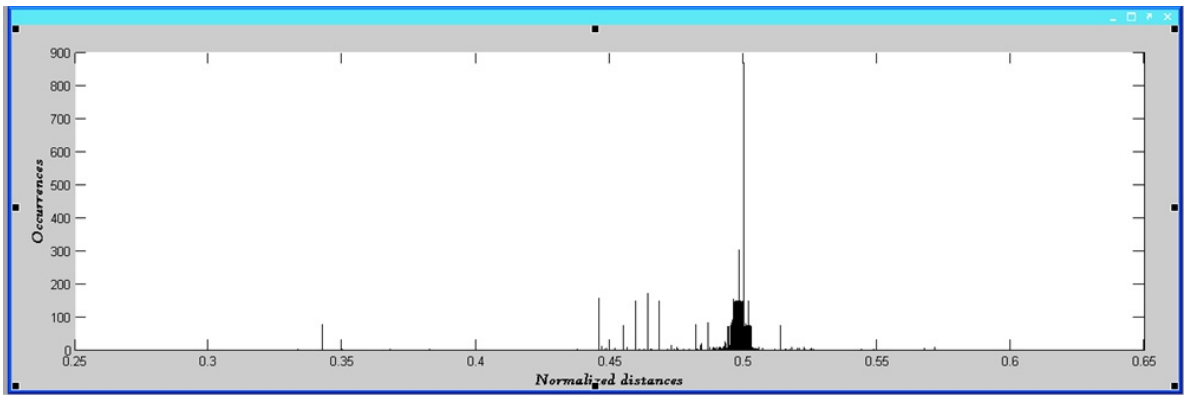


Figure 6: Study of the no correlation of the keys K_t for 200 sessions with the same password, and different salts

and per user.

For three iterations and a given password, we analyze the entities regenerated for the client and the KDC server and we deduce the following results:

- The sent messages are not related to the original password.
- The footprint of the password is unpredictable.
- The encryption keys are dynamic and per session.

4.2 Behavioral Study of the Communication Phase

In our approach the session keys are dynamic, per session and have a variable size. However, the behavioral study of these keys requires a normalized Hamming distance, named D, defined in [2] by:

$$D(S,S') = \frac{\sum_{i=0}^{k-1} ((S(i \ mod \ K) + S'(i \ mod \ K')) \ mod \ 2)}{k} \quad (1)$$

with S and S' are two binary strings having period successively K and K' not necessary the same and $k = lcm(K, K')$. This function D allows the estimation of correlation between binary sequences not necessary with the same length. Asimi et al [2] found that two binary strings S and S' are weakly correlated if $D \simeq 0.5$.

Propriety 1. *: Let S and S' be two periodic binary strings. we say that S and S' are weakly correlated if $D(S,S') \simeq 0.5$.*

In Figure 6, even under restricted cases the results are accumulated in the vicinity of 0.5, which means that the keys used and associated to the same password are not correlated. Therefore, knowledge of information on the key gives no information on the other. This is due to the uncorrelation of binary signals calculated by the hash function applied to the fingerprint password concatenated with a dynamic salt per session.

5 Security Analysis

The evolution of the computer system and the development of new technologies, the attacks become increasingly efficient. For these reasons, Kerberos has known several modifications to the levels of performance and functionality against these attacks. However Kerberos V5, the current version, with all its amelioration, was discussed by several security analysis [9, 13, 33, 31], those show its weaknesses specifically against the dictionary attack only in the communication phase.

In this section, we evaluate the security of our protocol by analyzing the level of influence using addition salt to the password, and the impact of the Diffie Hellman principle [7] against different types of attacks. Further, we discuss the impact of adding dynamic salt per session to the password in both client side and KDC server side. In the client side, the addition of a dynamic salt per session to the password, and the application of virtualization function make the authentication process by password unbreakable. They reduce the chance of password divination attacks such as brute force and dictionary attacks. In the other hand, storing the password footprint (dynamic password disturbed by salt in our case) is stronger than storing the clear password in KDC database server. This makes storage of password more reliable in the KDC server side.

5.1 Impact of Salt Upon Password

The majority of the applications users are conscious of authentication by passwords. It requires the storage of simple passwords in most cases [20, 26]. In parallel, other authentication alternatives have been proposed [34]. However, their use is too limited especially in web applications [32]. The description of Kerberos integrated a static salt (client address or the domain name) to disrupt the password used for the generation of encryption keys [12, 13].

This technique does not solve the problem of dictionary attack that represents a real challenge against the Kerberos authentication techniques [17, 31, 33]. To address this type of attack, our approach is based on the

Table 2: Comparison between our protocol and previous versions of Kerberos

Parameters	Previous version of Kerberos	Our protocol
Mutual authentication	OK	OK
Portability	OK	OK
Use of ticket	OK	OK
Use of expiration time	OK	OK
Use of Deffie Helman		OK
salt	static and per user	dynamic and per session
Session key	$K_s = H(pwd)$	$pwv = f(pwd)$ and $K_s = H(pwv\|\|salt_i)$
Based key	based on string-to-key function	based on S2KexS function
Derived key	based on derived key function	based on DKexS function
N		New authentication number calculated from password

RGSCS regenerator making the use of keys generation functions more robust, and who's their different outputs from a session to another.

As for the registration phase, the impact of salt used to disrupt the password makes it communication phase more reliable (registration of password footprint). Therefore the guess of original password either by listening to requests exchanged between the client and the KDC or by brute force is almost impossible.

5.2 Impact of the Diffie Hellman Principle

The principle of Diffie Hellman solved several types of attacks such as man in the middle [7]. It has undergone several changes [5, 6, 10] with the development of computers (computing speed, performance processors). The conjunction of this principle and the dynamic salt per session made the parameters used in our protocol more complicated and indefinable. This allows us to create a secure channel to a more secure password exchange. with this technology we have reduced the chance that a password will be guessed from the parameters stored or exchanged between client and KDC.

5.3 Robustness to the Dictionary Attack

Most password crackers are provided with standard dictionaries [23]. The experience allows that the Kerberos realm had already the strength of the password, reflecting authentication without sending it [12]. Although, the description of registration phase is not written in any reference, and the communication phase is based on a clearly stored password [18, 25, 28]. Our principle reduces the probability of finding the password is in the registration phase and communication phase. It is caused by disturbance by adding the dynamic salt per session and application virtualization function. Even if a hacker succeeded in capturing several messages, he will not have the opportunity to find the password in question by the dictionary attack.

6 Comparison Between Our Protocol and Previous Versions

Our protocol, which is a Kerberos V5 improvement, aims to ensure the confidential exchange between clients, authentication servers and services server. For these reasons our approach is based on tickets, the Diffie Hellman protocol and other functions namely: S2KexS function, DKexS function [30].

- Diffie Hellman algorithm allows confidential exchange of credentials authentication without requirement HTTPS.
- S2KexS function calculates a more robust and undevinable base key from a dynamic salt and a password digital print.
- DKexS function calculates three encryption keys used in the communication phase to ensure the confidentiality and integrity of data exchanged between clients, servers and services.

However the adding of pseudorandom regenerator, S2KexS function, DKexS function and Diffie Hellman protocol makes our protocol more robust. The comparison between our approach and the traditional Kerberos defined in [25] is described as follows:

7 Conclusion

Several extended authentication protocols have been described for strong password authentication [20, 27, 29]. For Kerberos, several solutions have been proposed such as using the smart card [16] or public keys [17] etc, but these techniques do not reduce the chance that the password is guessed and the rest of the protocol becomes breakable. In this article, we presented a new protocol based on the principle of Diffie Hellman [7] and the regenerator of salt RGSCS [2] cryptographically secure and per session. Our principal objective, however, was to protect users even with weak passwords. This leads us to use

these techniques to face the current known attacks by Kerberos V5 such as dictionary attack [33]. Our authentication scheme provides a more reliable model with uncorrelated authentication parameters between different clients in the same realm even if they have identical passwords. This is proved by the behavioral study who presented an encouraging results with unpredictable keys even with the use of a weak password.

References

[1] Y. Asimi, A.Amghar, A. Asimi, and Y. Sadqi, "Strong zero-knowledge authentication based on the session keys (SAAK)," *International Journal of Network Security & Its Applications*, vol. 7, no. 1, p. 51, 2015.

[2] Y. Asimi, A.Amghar, A. Asimi, and Y. Sadqi, "New random generator of a safe cryptographic salt per session," *International Journal of Network Security*, vol. 18, no. 3, pp. 445–453, 2016.

[3] Y. Asimi, A. Amghar, A. Asimi, and Y. Sadqi, "Strong zero-knowledge authentication based on virtual passwords," *International Journal of Network Security*, vol. 18, no. 4, pp. 601–616, 2015.

[4] S. M. Bellovin and M. Merritt, "Encrypted key exchange: Password-based protocols secure against dictionary attacks," in *IEEE Computer Society Symposium on Research in Security and Privacy*, pp. 72–84, Oakland, May 1992.

[5] D. Boneh, *The Decision Diffie-Hellman Problem*, pp. 48–63, Springer, Berlin, Heidelberg, 1998.

[6] D. Cash, E. Kiltz, and V. Shoup. *The Twin Diffie-Hellman Problem and Applications*, pp. 127–145, Springer, Berlin, Heidelberg, Apr. 2008.

[7] W. Diffie and M. Hellman, "New directions in cryptography," *IEEE Transactions on Information Theory*, vol. 22, no. 6, pp. 644–654, 1976.

[8] P. Dourish, R. E. Grinter, J. D. De La Flor, and M. Joseph, "Security in the wild: user strategies for managing security as an everyday, practical problem," *Personal and Ubiquitous Computing*, vol. 8, no. 6, pp. 391–401, 2004.

[9] E. El-Emam, M. Koutb, H. Kelash, and O. Faragallah, "An authentication protocol based on kerberos 5.," *International Journal of Network Security*, vol. 12, no. 3, pp. 159–170, 2011.

[10] N. Fazio, R. Gennaro, I. M. Perera, and W. E. Skeith, *Hard-Core Predicates for a Diffie-Hellman Problem over Finite Fields*, pp. 148–165, Springer, Berlin, Heidelberg, 2013.

[11] S. Gaw and E. W. Felten, "Password management strategies for online accounts," in *Proceedings of the second symposium on Usable privacy and security*, pp. 44–55, Pittsburgh, Pennsylvania, USA, July 2006.

[12] J. Kohl and C. Neuman, "The kerberos network authentication service (v5)," Tech. Rep. RFC 1510, Sep. 1993.

[13] J. Y. Kohl, B. C. Neuman, and Y. Theodore, "The evolution of the kerberos authentication service," 1994.

[14] C. C. Lee, C. H. Liu, and M .S.Hwang, "Guessing attacks on strong-password authentication protocol," *International Journal of Network Security*, vol. 15, no. 1, pp. 64–67, 2013.

[15] R. Lidl and H. Niederreiter, "Finite fields: Encyclopedia of mathematics and its applications," *Computers and Mathematics with Applications*, vol. 7, no. 33, p. 136, 1997.

[16] N. Mavrogiannopoulos, A. Pashalidis, and B. Preneel, "Toward a secure kerberos key exchange with smart cards," *International Journal of Information Security*, vol. 13, no. 3, pp. 217–228, 2014.

[17] M.Backes, I. Cervesato, A. D. Jaggard, A. Scedrov, and J. K. Tsay, "Cryptographically sound security proofs for basic and public-key kerberos," *International Journal of Information Security*, vol. 10, no. 2, 2011.

[18] A. Melnikov, "The kerberos v5 (gssapi) simple authentication and security layer (SAAL) mechanism," Nov. 2006.

[19] A. J. Menezes, P. C. Van Oorschot, and S. A. Vanstone, *Handbook of Applied Cryptography*, CRC press, 1996.

[20] R. Morris and K. Thompson, "Password security: A case history," *Communications of the ACM*, vol. 22, no. 11, pp. 594–597, 1979.

[21] R. M. Needham and M. D. Schroeder, "Using encryption for authentication in large networks of computers," *Communications of the ACM*, vol. 21, no. 12, pp. 993–999, 1978.

[22] G. Notoatmodjo and C. Thomborson, "Passwords and perceptions," in *Proceedings of the Seventh Australasian Conference on Information Security*, pp. 71–78, Wellington, New Zealand, Jan. 2009.

[23] D. P.Jablon, "Extended password key exchange protocols immune to dictionary attack," in *Sixth IEEE Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises*, pp. 248–255, June 1997.

[24] K. Raeburn, "Encryption and checksum specifications for kerberos 5," Feb. 2005.

[25] K. Raeburn, "Network working group c. neuman request for comments: 4120 USC-ISI obsoletes: 1510 t. Yu category: Standards track s. hartman," July 2005.

[26] Y. Sadqi, A. Asimi, and Y. Asimi, "A cryptographic mutual authentication scheme for web applications," *arXiv preprint arXiv:1412.2908*, 2014.

[27] Y. Sadqi, A. Asimi, and Y. Asimi. "Short: A lightweight and secure session management protocol,". in *Networked Systems*, pp. 319–323. Springer, Marrakech, Morocco, May 2014.

[28] B. Schneier, *Applied cryptography: protocols, algorithms, and source code in C*, john wiley and sons, 2007.

[29] J. G. Steiner, B. C. Neuman, and J. I. Schiller, "Kerberos: An authentication service for open network systems," in *USENIX Winter*, pp. 191–202, Dallas, TX, Feb 1988.
[30] Z. Tbatou, A. Asimi, Y. Asimi, and Y. Sadqi, "Kerberos v5: Vulnerabilities and perspectives," in *Third World Conference on Complex Systems (WCCS'15)*, pp. 1–5, Marrakech, Morocco, Nov. 2015.
[31] J. K. Tsay, *Formal Analysis of the Kerberos Authentication Protocol*, PhD thesis, University of Pennsylvania, 2008.
[32] R. Tso, "Security analysis and improvements of a communication-efficient three-party password authenticated key exchange protocol," *The Journal of Supercomputing*, vol. 66, no. 2, pp. 863–874, 2013.
[33] T. D. Wu, "A real-world analysis of kerberos password security.," in *NDSS*, Feb. 1999.
[34] Q. Xie, B. Hu, and T. Wu, "Improvement of a chaotic maps-based three-party password-authenticated key exchange protocol without using server's public key and smart card," *Nonlinear Dynamics*, vol. 79, no. 4, pp. 2345–2358, 2015.

Biography

Tbatou Zakariae received his Master's degree in Computer Science and Distributed Systems in 2013 from Departments of Mathematics and Computer Science, Faculty of Science, University Ibn Zohr, Agadir, Morocco. He is currently Ph.D student in Departments of Mathematics and Computer Sciences, Information Systems and Vision Laboratory, Morocco. His research interests include Authentication Protocols, distributed systems, cloud computing, Computer and Network Security and Cryptography.

ASIMI Ahmed is a full professor at the Faculty of Science, Agadir, Morocco. He received his Ph.D degree in Number theory from Department of Mathematics, Faculty of Science, University Mohammed V, Agdal in 2001, Morocco. He is reviewer at the International Journal of Network Security (IJNS) and at the journal of Computer and Information Science. He is a speaker in national and international conferences on the topics of cryptology and computer security. His main areas of research interests include Number theory, Code theory, Computer Cryptology, Computer and Network Security.

Younes Asimi received his Ph.D. in Strong Zero-Knowledge Authentication Based on virtual passwords per session and the Session Keys in 2015. He is currently pursuing Ph.D in Departments of Mathematics and Computer Sciences, Information Systems and Vision Laboratory, Morocco. His research interests include Authentication Protocols, Computer and Network Security and Cryptography.

Yassine SADQI received his Ph.D in the security of Computer Science and Distributed Systems at the Ibn Zoher University in 2015. Agadir, Morocco. His main ?eld of research interest is computer security, cryptography and authentication in Web applications.

Guezzaz Azidine received his Master's degree in the field of Computer Science and Distributed Systems in 2013 from Departments of Mathematics and Computer Science, Faculty of Science, University Ibn Zohr, Agadir, Morocco. He is currently Ph.D student in Departments of Mathematics and Computer Sciences, Information Systems and Vision Laboratory, Morocco. His main field of research interest is Intrusion Detection and Prevention, Computer and Network Security and Cryptography.

Further Characterization of $\mathcal{H}$ Vectorial Functions

Yuwei Xu[1,2], Chuankun Wu[1]

(Corresponding author: Yuwei Xu)

State Key Laboratory of Information Security, Institute of Information Engineering[1]
Chinese Academy of Sciences, Beijing 100093, China
School of Cyber Security, University of Chinese Academy of Sciences, Beijing 100049, China[2]
(Email: xuyuwei@iie.ac.cn)

(Received June 30, 2016; revised and accepted Sept. 3 & Sept. 25, 2016)

Abstract

Vectorial Boolean bent functions, which possess the maximal nonlinearity and the minimum differential uniformity, contribute to optimum resistance against linear cryptanalysis and differential cryptanalysis. $\mathcal{H}$ vectorial functions is an infinite class of vectorial Boolean bent functions presented by S. Mesnager. This paper is devoted to further characterization of the $\mathcal{H}$ vectorial functions. It is shown that the EA-equivalent relationships among vectorial Boolean functions may be characterized by their component functions. As a result, the EA-equivalent relationships among $\mathcal{H}$ vectorial functions induced by many projectively equivalent o-polynomials of a given o-polynomial are obtained.

Keywords: Bent Functions; Cryptography; EA-equivalence; $\mathcal{H}$ Functions; O-polynomials

1 Introduction

Vectorial Boolean functions, which are widely used in block ciphers, stream ciphers and Hash functions, paly an important role in cryptography [1, 2, 14, 15, 16, 19]. The security of the cryptographic algorithms, adopting vectorial Boolean functions as nonlinear components, usually depends on the cryptographic properties of the vectorial Boolean functions adopted [12]. The nonlinearity and the differential uniformity of the adopted vectorial Boolean functions are two parameters that measure the resistence of the cryptographic algorithms against linear cryptanalysis [3, 18] and differential cryptanalysis [4, 17] respectively. The vectorial Boolean functions possessing the maximal nonlinearity, which is the optimal nonlinearity, are referred to as *vectorial Boolean bent functions*. The concept bent of vectorial Boolean functions, which is an extension of Boolean bent functions [24], was first considered by Nyberg in [22], where it was shown that bent (n, m)-functions (i.e., the vectorial Boolean functions from $\mathbb{F}_{2^n}$ to $\mathbb{F}_{2^m}$) exist if and only if n is even and $n \geq 2m$. Vectorial Boolean bent functions are also named as *perfect nonlinear functions* [11, 22], for the reason possessing the minimum differential uniformity, which is the optimal differential uniformity. Thus, the study of vectorial Boolean bent functions are of great significance.

In [20], an infinite class of vectorial Boolean bent functions named as $\mathcal{H}$ *vectorial functions* was presented. More precisely, it was shown in [20] that, if G is an o-polynomial on $\mathbb{F}_{2^k}$, then the function $xG(yx^{2^k-2})$ is bent, where $(x, y) \in \mathbb{F}_{2^k} \times \mathbb{F}_{2^k}$. In [20], it is proved that $\mathcal{H}$ vectorial functions induced by the projectively equivalent o-polynomials $G(x)$, $\mu G(x) + \nu$, $G(\mu x + \nu)$, $xG(x^{2^k-2})$ and $(G(x^{2^s}))^{2^{k-s}}$ are EA-equivalent, where G is an o-polynomial on $\mathbb{F}_{2^k}$, $\mu \in \mathbb{F}_{2^k}^*$ and $\nu \in \mathbb{F}_{2^k}$. However, whether G is an o-polynomial is necessary for $xG(yx^{2^k-2})$ to be bent is unknown. And the EA-equivalent relationships among the $\mathcal{H}$ vectorial functions induced by other projectively equivalent o-polynomials is unclear.

This paper shows that, for $m \mid k$, the function $Tr_m^k(xG(yx^{2^k-2}))$ is bent if and only if G is an o-polynomial on $\mathbb{F}_{2^k}$, where $(x, y) \in \mathbb{F}_{2^k} \times \mathbb{F}_{2^k}$. This paper also shows that the EA-equivalent relationships among vectorial Boolean functions may be characterized by their component functions. Subsequently, the EA-equivalent relationships among the $\mathcal{H}$ vectorial functions induced by 27 projectively equivalent o-polynomials are characterized.

The rest of this paper is organized as follows. Section 2 provides some preliminaries for the description of the paper. Section 3 characterizes $\mathcal{H}$ vectorial functions. And Section 4 concludes this paper.

2 Preliminaries

Throughout this paper, let k, m be two positive integers, $\mathbb{F}_{2^k}$ denote the Galois field $GF(2^k)$ and $\mathbb{F}_{2^k}^* = \mathbb{F}_{2^k} \setminus \{0\}$.

For $m \mid k$, the trace function $Tr_m^k : \mathbb{F}_{2^k} \to \mathbb{F}_{2^m}$ is defined as

$$Tr_m^k(x) = x + x^{2^m} + x^{2^{2m}} + \cdots + x^{2^{(\frac{k}{m}-1)m}}.$$

In particular, $Tr_1^k(x)$ is called the absolute trace function on $\mathbb{F}_{2^k}$. Note that the trace function has the well known properties that $Tr_m^k(x) = Tr_1^m \circ Tr_m^k(x)$ and $Tr_m^k(x) = Tr_m^k(x^2)$.

A mapping $G : \mathbb{F}_{2^k} \to \mathbb{F}_{2^m}$ is referred to as a vectorial Boolean function, which is also known as a (k,m)-function, a multiple output Boolean function or an S-box. Particularly, G is a k-variable Boolean function if $m = 1$. A (k,m)-function G can be represented as $G = (g_1, g_2, \cdots, g_m)$, where $g_1, g_2, \cdots, g_m$ are m Boolean functions on $\mathbb{F}_{2^k}$ and called the *coordinate functions* of G. Any nonzero linear combination of the coordinate functions is called a *component function* of G, and can be represented as $Tr_1^m(\lambda G)$, where $\lambda \in \mathbb{F}_{2^m}^*$.

A (k,m)-function G can be uniquely represented in the univariate polynomial representation as $G(x) = \sum_{i=0}^{2^k-1} a_i x^i$, where $a_i \in \mathbb{F}_{2^k}$. The algebraic degree of G, denoted by $deg(G)$, is defined as $deg(G) = \max\{wt(i) : 0 \le i \le 2^k-1, a_i \neq 0\}$, where $wt(i)$ denotes the *Hamming weight* of i, i.e., the number of 1's of i in its 2-adic representation. G is called an *affine vectorial Boolean function* if $deg(G) \le 1$. Particularly, a *linear vectorial Boolean functions* is a affine vectorial Boolean functions with algebraic degree 1 and constant term null, or with algebraic degree 0 (i.e., constant function). For $m \mid k$, G can also be represented in a non-unique way as

$$G(x) = Tr_m^k(P(x)),\ P(x) \in \mathbb{F}_{2^k}[x].$$

An (n,m)-function F with $n = 2k$ can be uniquely represented in the bivariate polynomial representation as $F(x,y) = \sum_{0 \le i_1, i_2 \le 2^k-1} a_{i_1,i_2} x^{i_1} y^{i_2}$, where $(x,y) \in \mathbb{F}_{2^k} \times \mathbb{F}_{2^k}$ and $a_{i_1,i_2} \in \mathbb{F}_{2^k}$. The algebraic degree of F is $deg(F) = \max\{wt(i_1)+wt(i_2) : 0 \le i_1, i_2 \le 2^k-1, a_{i_1,i_2} \neq 0\}$. For $m \mid k$, F can also be represented non-uniquely as

$$F(x,y) = Tr_m^k(P(x,y)),\ P(x,y) \in \mathbb{F}_{2^k}[x,y].$$

The *nonlinearity* of a k-variable Boolean function g, denoted by $nl(g)$, is defined as $nl(g) = \min_{g' \in \mathbb{A}_k} d(g,g')$, where $\mathbb{A}_k$ is the set of all the k-variable affine Boolean functions and $d(g,g')$ is the *Hamming distance* between g and g', i.e., the cardinality of the set $\{x \in \mathbb{F}_{2^k} : g(x) \neq g'(x)\}$. The nonlinearity of g can be measured by $nl(g) = 2^{k-1} - \frac{1}{2}\max_{\omega \in \mathbb{F}_{2^k}} W_g(\omega)$, where $W_g(\omega) = \sum_{x \in \mathbb{F}_{2^k}} (-1)^{g(x)+Tr_1^k(\omega x)}$ is the *Walsh transform* of g. The *Walsh spectrum* of g is the set $\{W_g(\omega) : \omega \in \mathbb{F}_{2^k}\}$. The well known Parseval's equation $\sum_{\omega \in \mathbb{F}_{2^k}} (W_g(\omega))^2 = 2^{2k}$ implies that $nl(g) \le 2^{k-1} - 2^{\frac{k}{2}-1}$. An n-variable Boolean function f with n even is referred to as a *Boolean bent function* if and only if $nl(f) = 2^{n-1} - 2^{\frac{n}{2}-1}$.

The *nonlinearity* of a (k,m)-function G, denoted by $nl(G)$, is defined as $nl(G) = \min\{nl(Tr_1^m(\lambda G)) : \lambda \in \mathbb{F}_{2^m}^*\}$. The nonlinearity of G can be measured by $nl(G) = 2^{k-1} - \frac{1}{2}\max_{\omega \in \mathbb{F}_{2^k}} \max_{\lambda \in \mathbb{F}_{2^m}^*} W_G(\omega, \lambda)$, where $W_G(\omega,\lambda) = \sum_{x \in \mathbb{F}_{2^k}} (-1)^{Tr_1^m(\lambda G(x)) + Tr_1^k(\omega x)}$ is the *Walsh transform* of G. The *Walsh spectrum* of G is the set $\{W_G(\omega,\lambda) : \omega \in \mathbb{F}_{2^k}, \lambda \in \mathbb{F}_{2^m}^*\}$. The Parseval's equation also implies that, for the (k,m)-function G, $nl(G) \le 2^{k-1} - 2^{\frac{k}{2}-1}$. An (n,m)-function F with n even is referred to as a *vectorial Boolean bent function* if and only if $nl(F) = 2^{n-1} - 2^{\frac{n}{2}-1}$. The bent property of vectorial Boolean functions can be characterized by their component functions.

Definition 1. *An (n,m)-function F with n even is bent if and only if all of its component functions are Boolean bent functions (i.e., $Tr_1^m(\lambda F)$ is bent for every $\lambda \in \mathbb{F}_{2^m}^*$).*

The extended affine equivalence (EA-equivalence) and the Carlet-Charpin-Zinoviev equivalence (CCZ-equivalence) are two greatly useful tools to study the existence, constructions and various properties of vectorial Boolean functions. Although EA-equivalence is a particular case of CCZ-equivalence [6, 9], the two concepts of equivalent relations are coincident in some special cases [5], such as Boolean functions [6] and vectorial Boolean bent functions [7]. Note that the nonlinearity is an EA-invariant parameter [9]. Here, we recall the definition of EA-equivalence.

Definition 2 ([5, 9, 23])**.** *Let G, G' be two (k,m)-functions and*

$$G' = A_3 \circ G \circ A_2 + A_1.$$

The corresponding concepts of equivalence between G and G' are called:

- *Linear equivalence, if A_3 and A_2 are two linear permutations on $\mathbb{F}_{2^m}$ and $\mathbb{F}_{2^k}$ respectively, and A_1 is null.*
- *Affine equivalence, if A_3 and A_2 are two affine permutations on $\mathbb{F}_{2^m}$ and $\mathbb{F}_{2^k}$ respectively, and A_1 is null.*
- *Extended affine equivalence (EA-equivalence), if A_3 and A_2 are two affine permutations on $\mathbb{F}_{2^m}$ and $\mathbb{F}_{2^k}$ respectively, and A_1 is an affine (k,m)-function.*

We recall the definition of o-polynomials.

Definition 3 ([10])**.** *A permutation polynomial G on $\mathbb{F}_{2^k}$ is called an oval polynomial (o-polynomial), if the function*

$$x \in \mathbb{F}_{2^k} \mapsto \begin{cases} \frac{G(x+\gamma)+G(\gamma)}{x}, & if\ x \neq 0 \\ 0, & if\ x = 0 \end{cases}$$

is a permutation on $\mathbb{F}_{2^k}$ for every $\gamma \in \mathbb{F}_{2^k}$.

In the end of this section, we recall two useful lemmas.

Lemma 1 ([10])**.** *The function $Tr_1^k(xG(yx^{2^k-2}))$ is bent if and only if G is an o-polynomial on $\mathbb{F}_{2^k}$, where $(x,y) \in \mathbb{F}_{2^k} \times \mathbb{F}_{2^k}$.*

Lemma 2 ([10]). *Let G be an o-polynomial on $\mathbb{F}_{2^k}$. For every $\lambda \in \mathbb{F}_{2^k}^*$, $Tr_1^k(xG(yx^{2^k-2}))$ and $Tr_1^k(\lambda xG(yx^{2^k-2}))$ are EA-equivalent, where $(x,y) \in \mathbb{F}_{2^k} \times \mathbb{F}_{2^k}$.*

3 Further Characterization of $\mathcal{H}$ vectorial functions

In [20], S. Mesnager shown that, if G is an o-polynomial on $\mathbb{F}_{2^k}$, then the function $xG(yx^{2^k-2})$ is bent, $(x,y) \in \mathbb{F}_{2^k} \times \mathbb{F}_{2^k}$, which is referred to as $\mathcal{H}$ vectorial functions. Here we give the following conclusion.

Theorem 1 ($\mathcal{H}$ vectorial functions). *Let $m \mid k$. Then the function*

$$Tr_m^k(xG(yx^{2^k-2}))$$

is bent if and only if G is an o-polynomial on $\mathbb{F}_{2^k}$, where $(x,y) \in \mathbb{F}_{2^k} \times \mathbb{F}_{2^k}$.

Proof. According to Lemma 1, the necessity is obvious.

Assume G is an o-polynomial on $\mathbb{F}_{2^k}$. According to Lemma 2, for any $\lambda_1, \lambda_2 \in \mathbb{F}_{2^m}^*$, the bent properties of $Tr_1^k(\lambda_1 xG(yx^{2^k-2}))$ and $Tr_1^k(\lambda_2 xG(yx^{2^k-2}))$ are the same. According to Definition 1 and Lemma 1, the sufficiency holds. □

Proposition 2 in [20] showed that, the function $xG(yx^{2^k-2})$ is EA-equivalent to every one of $\mu xG(yx^{2^k-2}) + \nu)$, $xG(\mu yx^{2^k-2} + \nu)$, $yG(y^{2^k-2}x)$ and $x(G(y^{2^s}x^{2^k+2^s-2}))^{2^{k-s}}$, where G is an o-polynomial on $\mathbb{F}_{2^k}$, $s \in \mathbb{N}$, $\mu \in \mathbb{F}_{2^k}^*$ and $\nu \in \mathbb{F}_{2^k}$. That is, S. Mesnager's $\mathcal{H}$ vectorial functions induced by the projectively equivalent o-polynomials $G(x)$, $\mu G(x)+\nu$, $G(\mu x+\nu)$, $xG(x^{2^k-2})$ and $(G(x^{2^s}))^{2^{k-s}}$ are EA-equivalent. Recall that two o-polynomials G and G' are called *projectively equivalent* [8] if $G^\alpha = \frac{G(x)+G(0)}{G(1)+G(0)}$ and $G'^\alpha = \frac{G'(x)+G'(0)}{G'(1)+G'(0)}$ define equivalent hyperovals. However, the proof of Proposition 2 in [20] is based on special forms of the four projectively equivalent o-polynomials $\mu G(x)+\nu$, $G(\mu x+\nu)$, $xG(x^{2^k-2})$ and $(G(x^{2^s}))^{2^{k-s}}$, which is not suitable for the general case.

Here, we introduce a new technique for studying the EA-equivalent relationships among the $\mathcal{H}$ vectorial functions induced by projectively equivalent o-polynomials. That is, the EA-equivalent relationships among vectorial Boolean functions may be characterized by their component functions. By this means, the EA-equivalent relationships among $\mathcal{H}$ vectorial functions induced by more projectively equivalent o-polynomials of a given o-polynomial can be characterized.

Lemma 3. *Let G, G' be two (k,m)-functions. Then there exist some affine (k,m)-function A_1 and some affine permutation A_2 on $\mathbb{F}_{2^k}$ such that $G' = G \circ A_2 + A_1$ if and only if $Tr_1^m(G)$ and $Tr_1^m(G')$ are EA-equivalent.*

Proof. The necessity is obvious. In the following, we prove the sufficiency.

By Definition 2, $Tr_1^m(G)$ and $Tr_1^m(G')$ are EA-equivalent if and only if there exist some affine permutation A_2 on $\mathbb{F}_{2^k}$ and some k-variable affine Boolean function g such that $Tr_1^m(G'(x)) = Tr_1^m(G(A_2(x))) + g(x)$. For the k-variable affine Boolean function g, there exists some affine function $P(x) \in \mathbb{F}_{2^k}[x]$ such that $g(x) = Tr_1^k(P(x)) = Tr_1^m \circ Tr_m^k(P(x))$. Let $A_1(x) = Tr_m^k(P(x))$. Then A_1 is an affine (k,m)-function. Thus, $Tr_1^m(G'(x)) = Tr_1^m(G(A_2(x))) + Tr_1^m(A_1(x))$, i.e., $Tr_1^m(G'(x) + G(A_2(x)) + A_1(x)) \equiv 0$. Then $G' = G \circ A_2 + A_1$. □

Following from the discussions in [10, 8], we divide 27 projectively equivalent o-polynomials into four classes.

Lemma 4. *Let G be an o-polynomial on $\mathbb{F}_{2^k}$. Denote $\tau_1 = G(x)$, $\tau_2 = G^{-1}(x)$, $\tau_3 = (xG(x^{2^k-2}))^{-1}$, $\tau_4 = (x + xG(x^{2^k-2}+1))^{-1}$, and*

$$\begin{aligned}
S_{\tau_1} &= \{G(x), (G(x^{2^s}))^{2^{k-s}}, \mu G(x) + \nu, G(\mu x + \nu),\\
&\quad xG(x^{2^k-2}), G(x+1)+1, x(G(x^{2^k-2}+1)+1),\\
&\quad x + (x+1)G(x(x+1)^{2^k-2}),\\
&\quad (x+1)G((x+1)^{2^k-2})+1\},\\
S_{\tau_2} &= \{G^{-1}(x), zG^{-1}(x^{2^k-2}), G^{-1}(x+1)+1,\\
&\quad x(G^{-1}(x^{2^k-2}+1)+1),\\
&\quad x + (x+1)G^{-1}\cdot(x(x+1)^{2^k-2}),\\
&\quad (x+1)G^{-1}((x+1)^{2^k-2})+1\},\\
S_{\tau_3} &= \{\ (xG(x^{2^k-2}))^{-1},\\
&\quad (xG^{-1}(x^{2^k-2}))^{-1},\\
&\quad ((x+1)G^{-1}((x+1)^{2^k-2})+1)^{-1},\\
&\quad (x(x^{2^k-2} + (x^{2^k-2}+1)G((x+1)^{2^k-2}))^{-1})^{-1},\\
&\quad ((x+1)G((x+1)^{2^k-2})+1)^{-1},\\
&\quad (x(x^{2^k-2} + (x^{2^k-2}+1)G^{-1}((x+1)^{2^k-2}))^{-1})^{-1}\},\\
S_{\tau_4} &= \{(x + xG(x^{2^k-2}+1))^{-1},\\
&\quad (x + xG^{-1}(z^{2^k-2}+1))^{-1},\\
&\quad (x + (x+1)G^{-1}(x\cdot(x+1)^{2^k-2}))^{-1},\\
&\quad x(x^{2^k-2} + (x^{2^k-2}+1)G^{-1}((x+1)^{2^k-2}))^{-1},\\
&\quad (x + (x+1)G(x(x+1)^{2^k-2}))^{-1},\\
&\quad x(x^{2^k-2} + (x^{2^k-2}+1)G((x+1)^{2^k-2}))^{-1}\},
\end{aligned}$$

where $s \in \mathbb{N}$, $\mu \in \mathbb{F}_{2^k}^*$ and $\nu \in \mathbb{F}_{2^k}$. Let $i_1, i_2 \in \{\tau_1, \tau_2, \tau_3, \tau_4\}$, $G_1 \in S_{i_1}$ and $G_2 \in S_{i_2}$. Then $Tr_1^k(xG_1(yx^{2^k-2}))$ and $Tr_1^k(xG_2(yx^{2^k-2}))$, where $(x,y) \in \mathbb{F}_{2^k} \times \mathbb{F}_{2^k}$,

1) are EA-equivalent if $i_1 = i_2$;

2) may be EA-inequivalent if $i_1 \neq i_2$.

According to Lemma 3 and Lemma 4, we deduce

Theorem 2. *Let the parameters be identified with those in Lemma 4. Then* $Tr_m^k(xG_1(yx^{2^k-2}))$ *and* $Tr_m^k(xG_2(yx^{2^k-2}))$, *where* $(x,y) \in \mathbb{F}_{2^k} \times \mathbb{F}_{2^k}$, $G_1 \in S_{i_1}$ *and* $G_2 \in S_{i_2}$,

1) are EA-equivalent if $i_1 = i_2$;

2) may be EA-inequivalent if $i_1 \neq i_2$.

Note that $\mathcal{H}$ vectorial functions viewed in univariate representation are Niho vectorial Boolean bent functions. Indeed, the result of Lemma 4 in [10] can be extend to (n,m)-functions with $n = 2k$, which indicates that the restrictions of $\mathcal{H}$ vectorial functions to the vector space $\omega\mathbb{F}_{2^k}$ are linear for all $\omega \in \mathbb{F}_{2^n}^*$. Recall that a positive integer d (in the sense of modulo 2^n-1) is named as a Niho exponent and x^d a Niho power function if the restriction of x^d to $\mathbb{F}_{2^k}$ is linear [13, 21], i.e., $d \equiv 2^s \pmod{2^k-1}$ for some nonnegative integer $s < n$. A bent function is named as a Niho bent function if the exponents of all its non-constant terms are Niho exponents, when it is viewed in the univariate representation.

4 Conclusions

In this paper, $\mathcal{H}$ vectorial functions are further characterized. In [20], it was shown that G is an o-polynomial on $\mathbb{F}_{2^k}$ is sufficient for $xG(yx^{2^k-2})$ is bent, $(x,y) \in \mathbb{F}_{2^k} \times \mathbb{F}_{2^k}$ to be bent. However, the necessity is unknown. This paper proves that $Tr_m^k(xG(yx^{2^k-2}))$ is bent if and only if G is an o-polynomial on $\mathbb{F}_{2^k}$.

Based on special forms of the four projectively equivalent o-polynomials $\mu G(x) + \nu$, $G(\mu x + \nu)$, $xG(x^{2^k-2})$, $(G(x^{2^s}))^{2^{k-s}}$ of a given o-polynomial G, Proposition 2 in [20] showed that the $\mathcal{H}$ vectorial functions corresponding to the five projectively equivalent o-polynomials $G(x)$ $\mu G(x) + \nu$, $G(\mu x + \nu)$, $xG(x^{2^k-2})$, $(G(x^{2^s}))^{2^{k-s}}$ are EA-equivalent. In this paper, we introduce a new technique for studying the EA-equivalent relationships among vectorial Boolean functions, i.e Lemma 3. According to Lemma 3, the EA-equivalent relationships among the $\mathcal{H}$ vectorial functions corresponding to 27 projectively equivalent o-polynomials are characterized.

As we can see from Theorem 2, new projectively equivalent o-polynomials may derive new EA-inequivalent $\mathcal{H}$ vectorial functions, thus the identification and classification of new projectively equivalent o-polynomials of a given o-polynomial is very interesting, which is our future work.

Acknowledgments

This study was supported by National Natural Science Foundation of China (Grant No. 61173134) and Strategic Priority Research Program of the Chinese Academy of Sciences (Grant No. XDA06010701). The authors gratefully acknowledge the anonymous reviewers for their valuable comments.

References

[1] A. Aboshosha, K. A. ElDahshan, E. K. Elsayed, and A. A. Elngar, "Ea based dynamic key generation in RC4 ciphering applied to CMS," *International Journal of Network Security*, vol. 17, no. 4, pp. 405–412, 2015.

[2] M. Alam and S. Ray, "Design of an intelligent SHA-1 based cryptographic system: A CPSO based approach," *International Journal of Network Security*, vol. 15, no. 6, pp. 465–470, 2013.

[3] T. Baigneres, P. Junod, and S. Vaudenay, "How far can we go beyond linear cryptanalysis?," in *Advances in Cryptology (ASIACRYPT'04)*, pp. 432–450, Jeju Island, Korea, July 2004.

[4] E. Biham and A. Shamir, "Differential cryptanalysis of DES-like cryptosystems," *Journal of Cryptology*, vol. 4, no. 1, pp. 3–72, 1991.

[5] L. Budaghyan, *Construction and Analysis of Cryptographic Functions*, Springer, 2015.

[6] L. Budaghyan and C. Carlet, "CCZ-equivalence of single and multi-output boolean functions," in *Postproceedings of the Ninth International Conference on Finite Fields and Their Applications*, vol. 9, pp. 43–54, Dublin, Ireland, July 2009.

[7] L. Budaghyan and C. Carlet, "CCZ-equivalence of bent vectorial functions and related constructions," *Designs, Codes and Cryptography*, vol. 59, no. 1-3, pp. 69–87, 2011.

[8] L. Budaghyan, C. Carlet, T. Helleseth, and A. Kholosha, "On o-equivalence of Niho bent functions," in *Arithmetic of Finite Fields*, pp. 155–168, 2014.

[9] L. Budaghyan, C. Carlet, and A. Pott, "New classes of almost bent and almost perfect nonlinear polynomials," *IEEE Transactions on Information Theory*, vol. 52, no. 3, pp. 1141–1152, 2006.

[10] C. Carlet and S. Mesnager, "On Dillon's class $\mathcal{H}$ of bent functions, Niho bent functions and O-polynomials," *Journal of Combinatorial Theory, Series A*, vol. 118, no. 8, pp. 2392–2410, 2011.

[11] F. Chabaud and S. Vaudenay, "Links between differential and linear cryptanalysis," in *Advances in Cryptology (EUROCRYPT'94)*, pp. 356–365, Perugia, Italy, May 1995.

[12] C. Chen, X. Yu, Y. Xiang, X. Li, and T. Li, "An improved DPA attack on DES with forth and back random round algorithm," *International Journal of Network Security*, vol. 19, no. 2, pp. 285–294, 2017.

[13] H. Dobbertin, G. Leander, A. Canteaut, C. Carlet, P. Felke, and P. Gaborit, "Construction of bent functions via Niho power functions," *Journal of Combinatorial Theory, Series A*, vol. 113, no. 5, pp. 779–798, 2006.

[14] T. Gulom, "The encryption algorithms AES-PES16-1 and AES-RFWKPES16-1 based on networks PES16-1 and RFWKPES16-1," *International Journal of Electronics and Information Engineering*, vol. 3, no. 2, pp. 53–66, 2015.

[15] M. Hwang, C. Chang, and K. Hwang, "A watermarking technique based on one-way hash functions," *IEEE Transactions on Consumer Electronics*, vol. 45, no. 2, pp. 286–294, 1999.

[16] M. Hwang and P. Sung, "A study of micro-payment based on one-way Hash chain.," *International Journal of Network Security*, vol. 2, no. 2, pp. 81–90, 2006.

[17] J. Kim, S. Hong, B. Preneel, E. Biham, O. Dunkelman, and N. Keller, "Related-key boomerang and rectangle attacks: theory and experimental analysis," *IEEE Transactions on Information Theory*, vol. 58, no. 7, pp. 4948–4966, 2012.

[18] M. Matsui, "Linear cryptanalysis method for DES cipher," in *Advances in Cryptology (EUROCRYPT'93)*, pp. 386–397, Lofthus, Norway, May 1994.

[19] A. Mersaid and T. Gulom, "The encryption algorithm AES-RFWKIDEA32-1 based on network RFWKIDEA32-1," *International Journal of Electronics and Information Engineering*, vol. 4, no. 1, pp. 1–11, 2016.

[20] S. Mesnager, "Bent vectorial functions and linear codes from o-polynomials," *Designs, Codes and Cryptography*, vol. 77, no. 1, pp. 99–116, 2015.

[21] Y. Niho, "Multi-valued cross-correlation functions between two maximal linear recursive sequences," Tech. Rep., DTIC Document, 1972.

[22] K. Nyberg, "Perfect nonlinear S-boxes," in *Advances in Cryptology (EUROCRYPT'91)*, pp. 378–386, Brighton, UK, April 1991.

[23] K. Nyberg, "Differentially uniform mappings for cryptography," in *Advances in Cryptology (EUROCRYPT'93)*, pp. 55–64, Lofthus, Norway, May 1994.

[24] O. S. Rothaus, "On 'bent' functions," *Journal of Combinatorial Theory, Series A*, vol. 20, no. 3, pp. 300–305, 1976.

Biography

Yuwei Xu is currently pursuing Ph.D degree in State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, & University of Chinese Academy of Sciences. His research interest is cryptographic function.

Chuankun Wu received his PhD degree in Engineering in 1994. He has been working in the area of information security since. He has just recently joined Beijing Kuangn Pty Ltd specializing in security of industry control systems. Before that he was a research professor at the Institute of Information Engineering, Chinese Academy of Sciences. His research interests include cryptography, security protocols, and security techniques in Internet of Things.

A Robust and Efficient Remote Authentication Scheme from Elliptic Curve Cryptosystem

Guifa Hou, Zhijie Wang
(Corresponding author: Guifa Hou)
Department of Computer Science and Information Engineering, Anyang Institute of Technology
Huanghe Avenue, Anyang 455000, China
(Email: houguifa01@outlook.com)
(Received June 18, 2016; revised and accepted Sept. 3 & Sept. 25, 2016)

Abstract

Along with the extensive prevalence of the network and the portable equipments, people can access network resources conveniently. The protection of participants' privacy and data confidentiality is significant. Authentication mechanism is essential to assure the authenticity of all participants and forbid the illegal accessing. In this paper, we propose a robust remote authentication scheme with privacy protection, which achieves the efficiency. Besides, we prove the completeness of the proposed scheme through BAN-logic. The performance comparisons show that our proposal is sufficiently robust and suitable to the practical application environment.

Keywords: Anonymity; Authentication; BAN-logic

1 Introduction

With the large-scale proliferation of Internet and network technologies, users can conveniently obtain the desire resources by kinds of portable devices such as (e.g., mobile phones, PDAs and notebook computers) at any time and any place. On the other hand, it also brings kinds of network security problems due to the open nature of the Internet. In order to solve these security problems, the password based authentication schemes using smart cards have been widely deployed to verify the legitimacy of remote users in the login process. Since the computation capacity of these potable devices is limited, these authentication schemes should be more efficient for suiting to the practical application environment.

In 1981, Lamport [20] proposed a remote authentication scheme based on static login identity (ID). Until now, ample of remote authentication schemes based on Lamport's scheme have been published in the literatures [1, 5, 8, 10, 11, 13, 26]. These schemes can be further divided into static ID and dynamic ID schemes, the main drawback of the former schemes is that users should login to the remote server with the fixed ID. However, the latter kind of schemes can eliminate the risk of ID-theft and protect users' privacy. In 2004, Das et al. [8] presented a remote user authentication scheme based on dynamic ID using smart cards, which allowed users to choose and change their passwords freely, and need not servers to maintain the verifier table. However, in 2004, Awashti [2] analyzed several weaknesses of Das et al.'s scheme and showed that their scheme was completely insecure. Later on, many dynamic-ID authentication schemes based on Das et al.'s scheme are published to achieve better security and efficiency [1, 5, 10, 11, 13, 16, 17, 19, 26].

Because of the convenience and secure computation of smart cards, a number of password authentication schemes using smart cards have been proposed [3, 6, 7, 9, 12, 14, 18, 21, 23, 24]. Most of the previous authentication schemes assume that smart cards are tamper-resistant (i.e., secret information stored in the smart card cannot be revealed). However, recent research results have shown that the sensitive data stored in the smart card could be extracted by monitoring the power consumption and analyzing the leaked information about the cardholder [15, 22]. Thus, such schemes rely on the tamper-resistance assumption are prone to types of attacks, such as impersonation attack, server spoofing attack, and off-line password guessing attack, etc.. And hence, a secure authentication scheme should be able to withstand a series of attacks rely on stolen smart card attack.

Most of the schemes proposed in the literatures do not achieve the revocation of smart cards. This problem may lead to the abuse of lost smart cards to login the system successfully. Thereby, to avoid the misuse of smart cards, the remote server should allow users' revocation. In 2005, Fan et al. [9] proposed a robust authentication scheme based on the factoring problem. In their scheme, the smart cards revocation problem is solved. However, in 2009, Rhee et al. [23] pointed out Fan et al.'s scheme is vulnerable to server spoofing attack. At the same time, Wang et al. [25] presented an authentication scheme tried to solve smart cards revocation problem. Unfortunately, their scheme is susceptible to the known key attack and

Table 1: Notations

Notation	Meaning
U_i	The ith user
S	The remote server
ID_i	The identity of the user U_i
PW_i	The password of the user U_i
x	The master secret key of S
SK	The session key shared among U_i and S
$H(\cdot)$	A one-way hash function
$E_k(M)/D_k(C)$	The symmetric encryption/decryption
$\oplus$	Exclusive-OR operation
$\|$	String concatenation operation

the stolen smart card attack. In 2011, Wang et al. [24] proposed an improved scheme with key agreement based on the elliptic curve discrete logarithm problem. Nevertheless, in the same year, Chang et al. [6] pointed out Wang et al.'s scheme cannot withstand server spoofing attack and presented an improved authentication scheme.

In this paper, we propose a comparatively secure dynamic identity authentication scheme which achieves the criterion listed in Table II. Noticeably, in the security analysis, BAN-logic [4] is employed to prove the completeness of the proposal. From the performance and functionality comparisons, our scheme is superior for suiting the practical environment.

The structure of our paper is organized as follows. In Section 2, we propose an improved robust authentication scheme. Subsequently, we analyze the security of our proposal in Section 3 and compare the performance with the previous related protocols in Section 4. At last, Section 5 presents the overall conclusion.

2 Our Scheme

In this section, we propose an authentication scheme which can remedy a range of network attacks. It is composed five basic phases: registration phase, login phase, authentication and session key exchange phase, smart card revocation phase and off-line password change phase. The notations used in our scheme are summarized in Table 1.

2.1 Preliminaries

In this section, we introduce the basic knowledge about CAPTCHA in brief. More details about CAPTCHA are referenced in [27].

2.1.1 Related Concepts

Completely Automated Public Turing test to tell Computer and Humans Apart(CAPTCHA) is an automated test that humans can pass, but difficult for computers to pass. For example, CAPTCHA requires users to identify a series of letters that may be warped or obscured by distracting backgrounds and other noise in the image. Using CAPTCHA, S can distinguish legitimate users from computer bots while requiring minimal effort by human user.

2.2 Registration Phase

Initially, S stores a large number of CAPTCHA puzzles which correspond to answers in a database with the format $(puzzle, answer)$. Then the remote server S selects a large prime number p and two integer elements a, b, where $p > 2^{160}$ and $4a^3 + 27b^2 mod p \neq 0$. Then S chooses an elliptic curve equation $E_p(a,b) : y^2 = x^3 + ax + b mod p$. Let G be a base point of the elliptic curve, where n multiplies G is equal to O and $n > 2^{160}$.

Step 1: U_i selects his/her identity ID_i and password PW_i. After that, he/she registers in S with sending $\{ID_i, A_i\}$ over a secure communication channel, where $A_i = H(ID_i \| PW_i)$.

Step 2: Upon receiving the registration request, S computes $B_i = E_{A_i}(H(x\|n_i), n_i \cdot G)$, where x is the master secret key and n_i is a unique random number for U_i. Note that, the public key of S is $Pub_S = x \cdot G$.

Step 3: After that, S maintains a registration table which includes $(H(ID_i \oplus x) \cdot G, n_i)$. S can retrieve n_i from the registration table by $H(ID_i \oplus x) \cdot G$ in the revocation phase and in the authentication and key agreement phase.

Step 4: Then S writes $\{B_i, H(\cdot), G, E_k()/D_k()\}$ into the smart card and issues it to the client U_i through a secure channel.

2.3 Login Phase

When the user U_i wants to login S, he/she should insert the smart card to the terminal and key in ID_i with PW_i, then the smart card performs the following steps:

Step 1: The smart card computes $A_i = H(ID_i \| PW_i)$ to decrypt B_i and obtains $H(x\|n_i)$, $n_i \cdot G$. Afterwards, it generates a random nonce t in Z_p^* and computes

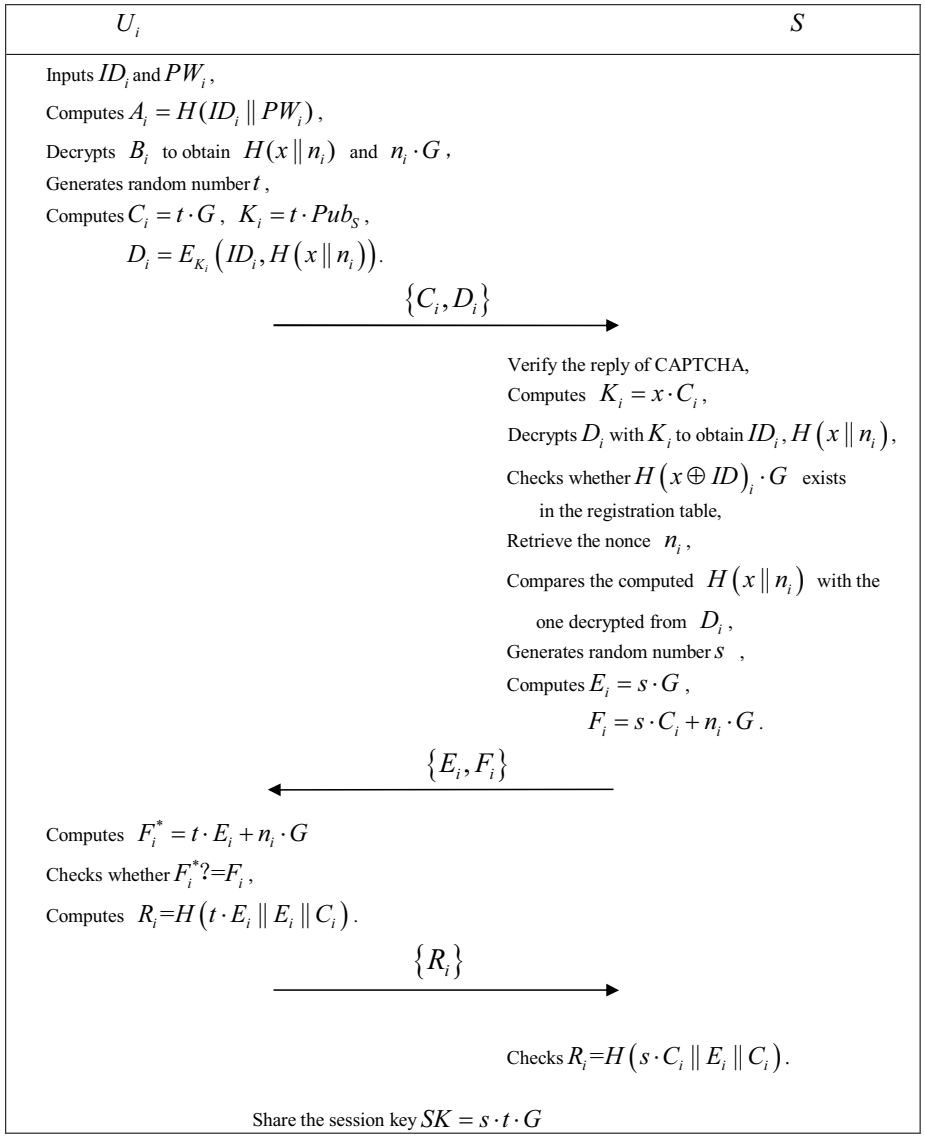


Figure 1: Login phase and authentication and session key exchange phase

$C_i = t \cdot G$, $K_i = t \cdot Pub_S$, $D_i = E_{K_i}(ID_i, H(x\|n_i))$, then sends the login request message $\{C_i, D_i\}$ to the remote server S.

Step 2: When S receives the message $\{C_i, D_i\}$, it randomly selects a CAPTCHA puzzle in its database and sends to U_i. If S receives the incorrect reply from U_i which is not corresponding to the puzzle transmitted to U_i, the login request will be terminated.

2.4 Authentication and Session Key Exchange Phase

Step 1: After checking the reply of CAPTCHA puzzle from U_i, S computes $K_i = x \cdot C_i$. Then it decrypts D_i with K_i to obtain ID_i and $H(x\|n_i)$. Then it calculates $H(x \oplus ID_i) \cdot G$ and checks it whether exists in the registration table. If so, S can retrieve the nonce n_i; otherwise, S aborts the messages.

Step 2: S calculates $H(x\|n_i)$ with the retrieved n_i. If the computed value is equal to the decrypted $H(x\|n_i)$ from D_i, S will execute the following steps; otherwise, the login request will be rejected.

Step 3: After the verification of U_i, S generates a random number s in Z_p^* and computes $E_i = s \cdot G$ and $F_i = s \cdot C_i + n_i \cdot G$. Then S transmits the replied message $\{E_i, F_i\}$ to U_i.

Step 4: Upon receiving the replication, the smart card computes $F_i^* = t \cdot E_i + n_i \cdot G$ and checks $F_i^* ? = F_i$. If the equation holds, the legitimacy of S is authentic. After that, U_i computes $R_i = H(t \cdot E_i \| E_i \| C_i)$ and transmits the session key verification message $\{R_i\}$ to S.

Step 5: Upon receiving the reply, S verifies whether R_i equals to the computed value $H(s \cdot C_i \| E_i \| C_i)$. If the equivalence holds, the mutual authentication is achieved; else, the entire authentication is failed.

After finishing the mutual authentication, U_i and S agree on the common session key $SK = s \cdot t \cdot G$.

2.5 Password Changing Phase

When U_i wants to update his/her password without the help of S. U_i inserts his/her smart card into a card reader and inputs ID_i with PW_i.

Step 1: U_i computes $A_i = H(ID_i \| PW_i)$ to decrypt B_i and obtains $H(x \| n_i)$, $n_i \cdot G$.

Step 2: U_i can be allowed to input the new password PW_i^{new}.

Step 3: The smart card computes $A_i^{new} = H(ID_i \| PW_i^{new})$, $B_i^{new} = E_{A_i^{new}}(H(x \| n_i), n_i \cdot G)$, and stores B_i^{new} into the smart card to replace B_i.

2.6 Smart Card Revocation Phase

In case of lost or stolen smart cards, U_i could request S for its revocation. In our scheme, U_i should transmit ID_i to S via a secure communication channel, then S computes $H(ID_i \oplus x) \cdot G$ and checks it whether exists in the registration table or not. If so, S removes the entry $(H(ID_i \oplus x) \cdot G, n_i)$ from the registration table.

3 Secure Analysis of Our Scheme

3.1 Completeness Proof Based on BAN-logic

In this section, we prove that the authentication goals using BAN-logic [4], which is a logic of belief focuses on the beliefs of the legitimate principals involved in the protocol. Let define the notations below:

- $\mathcal{P} \mid\equiv X$: The principal $\mathcal{P}$ believes a statement X or $\mathcal{P}$ would be entitled to believe X.
- $\sharp(X)$: The formula X is fresh.
- $\mathcal{P} \Rightarrow X$: The principal $\mathcal{P}$ has jurisdiction over the statement X.
- $\mathcal{P} \triangleleft X$: The principal $\mathcal{P}$ sees the statement X.
- $\mathcal{P} \mid\sim X$: The principal $\mathcal{P}$ once said the statement X.
- (X, Y): The formula X or Y is one part of the formula (X, Y).
- $\langle X \rangle_Y$: The formula X is combined with the formula Y.
- $\{X\}_Y$: The formula X is encrypted under the key Y.
- $\mathcal{P} \overset{k}{\longleftrightarrow} \mathcal{Q}$: The principals $\mathcal{P}$ and $\mathcal{Q}$ use the shared key k to communicate. Here, k will never be discovered by any principal except for $\mathcal{P}$ and $\mathcal{Q}$.
- $\mathcal{P} \overset{k}{\rightleftharpoons} \mathcal{Q}$: k is shared secret known to $\mathcal{P}$, $\mathcal{Q}$, and possibly to one trusted by them.
- SK: The session key used in the current session.

In the following, we introduce Some main logical postulates used in the demonstration:

- The message-meaning rule: $\frac{\mathcal{P} \mid\equiv \mathcal{Q} \overset{k}{\longleftrightarrow} \mathcal{P}, \mathcal{P} \triangleleft \{X\}_k}{\mathcal{P} \mid\equiv \mathcal{Q} \mid\sim X}$, $\frac{\mathcal{P} \mid\equiv \mathcal{Q} \overset{k}{\rightleftharpoons} \mathcal{P}, \mathcal{P} \triangleleft \langle X \rangle_k}{\mathcal{P} \mid\equiv \mathcal{Q} \mid\sim X}$.
- The freshness-conjuncatenation rule: $\frac{\mathcal{P} \mid\equiv \sharp(X)}{\mathcal{P} \mid\equiv \sharp(X, Y)}$.
- The nonce-verification rule: $\frac{\mathcal{P} \mid\equiv \sharp(X), \mathcal{P} \mid\equiv \mathcal{Q} \mid\sim X}{\mathcal{P} \mid\equiv \mathcal{Q} \mid\equiv X}$.
- The jurisdiction rule: $\frac{\mathcal{P} \mid\equiv \mathcal{Q} \Rightarrow X, \mathcal{P} \mid\equiv \mathcal{Q} \mid\equiv X}{\mathcal{P} \mid\equiv X}$, $\frac{\mathcal{P} \mid\equiv (X, Y)}{\mathcal{P} \mid\equiv X}$, $\frac{\mathcal{P} \triangleleft (X, Y)}{\mathcal{P} \triangleleft X}$, $\frac{\mathcal{P} \mid\equiv \mathcal{Q} \mid\sim (X, Y)}{\mathcal{P} \mid\equiv \mathcal{Q} \mid\sim X}$.

For proving the proper mutual authentication and the agreement of session key, we list the verification goals as follows:

Goal 1: $U_i \mid\equiv (U_i \overset{SK}{\longleftrightarrow} S)$.

Goal 2: $S \mid\equiv (U_i \overset{SK}{\longleftrightarrow} S)$.

Next, we list the idealized form transformed from the proposed scheme in the following:

Message 1: $U_i \rightarrow S$: $(C_i, \{ID_i, C_i\}_{\langle n_i \rangle_x})$.

Message 2: $S \rightarrow U_i$: $(E_i, \{S \mid\equiv (S \overset{SK}{\rightleftharpoons} U_i), C_i, E_i\}_{n_i})$.

Message 3: $U_i \rightarrow S$: $\langle C_i, E_i \rangle_{SK}$.

The following assumptions are presented to further analyze our scheme:

A.1: $U_i \mid\equiv (U_i \overset{n_i}{\longleftrightarrow} S)$;

A.2: $S \mid\equiv (S \overset{\langle n_i \rangle_x}{\longleftrightarrow} U_i)$;

A.3: $U_i \mid\equiv \sharp(C_i)$;

A.4: $S \mid\equiv \sharp(E_i)$;

A.5: $S \mid\equiv U_i \Rightarrow (ID_i, C_i)$;

A.6: $S \mid\equiv U_i \Rightarrow (C_i, E_i)$;

A.7: $U_i \mid\equiv S \Rightarrow (S \mid\equiv (S \overset{SK}{\rightleftharpoons} U_i), C_i, E_i)$;

A.8: $U_i \mid\equiv t$;

A.9: $S \mid\equiv s$.

According to the above-mentioned logical postulates and assumptions, we demonstrate the validity of our scheme in the following:

- According to Message 1, we obtain:
  $S \triangleleft (C_i, \{ID_i, C_i\}_{\langle n_i \rangle_x})$.

- According to the jurisdiction rule, we obtain:
  $S \triangleleft \{ID_i, C_i\}_{\langle n_i \rangle_x}$.
- According to Assumption A.2 and the message-meaning rule, we obtain:
  $S \mid\equiv U_i \mid\sim (ID_i, C_i)$.
- According to the jurisdiction rule, we obtain:
  $S \mid\equiv U_i \mid\sim C_i$.
- According to Message 2, we obtain:
  $U_i \triangleleft (E_i, \{S \mid\equiv (S \overset{SK}{\rightleftharpoons} U_i), C_i, E_i\}_{n_i})$.
- According to the jurisdiction rule, we obtain:
  $U_i \triangleleft \{S \mid\equiv (S \overset{SK}{\rightleftharpoons} U_i), C_i, E_i\}_{n_i}$.
- According to Assumption A.1 and the message-meaning rule, we obtain:
  $U_i \mid\equiv S \mid\sim (S \mid\equiv (S \overset{SK}{\rightleftharpoons} U_i), C_i, E_i)$.
- According to Assumption A.3 and the freshness-conjuncatenation rule, we obtain:
  $U_i \mid\equiv \sharp(S \mid\equiv (S \overset{SK}{\rightleftharpoons} U_i), C_i, E_i)$.
- According to $U_i \mid\equiv S \mid\sim (S \mid\equiv (S \overset{SK}{\rightleftharpoons} U_i), C_i, E_i)$ and the nonce-verification rule, we obtain:
  $U_i \mid\equiv S \mid\equiv (S \mid\equiv (S \overset{SK}{\rightleftharpoons} U_i), C_i, E_i)$.
- According to Assumption A.7 and the jurisdiction rule, we obtain:
  $U_i \mid\equiv (S \mid\equiv (S \overset{SK}{\rightleftharpoons} U_i), C_i, E_i)$.
- According to the jurisdiction rule, we obtain:
  $U_i \mid\equiv (S \mid\equiv (S \overset{SK}{\rightleftharpoons} U_i))$, $U_i \mid\equiv E_i$.
- According to $SK = t \cdot E_i$ and Assumption A.8, we obtain:
  $U_i \mid\equiv (U_i \overset{SK}{\longleftrightarrow} S)$ (**Goal 1**).
- According to Message 3, we obtain:
  $S \triangleleft \langle C_i, E_i \rangle_{SK}$.
- According to $S \mid\equiv S \overset{SK}{\rightleftharpoons} U_i$ and message-meaning rule, we obtain:
  $S \mid\equiv U_i \mid\sim (C_i, E_i)$.
- According to Assumption A.4 and the freshness-conjuncatenation rule, we obtain:
  $S \mid\equiv \sharp(C_i, E_i)$.
- According to $S \mid\equiv U_i \mid\sim (C_i, E_i)$ and the nonce-verification rule, we obtain:
  $S \mid\equiv U_i \mid\equiv (C_i, E_i)$.
- According to Assumption A.6 and the jurisdiction rule, we obtain:
  $S \mid\equiv (C_i, E_i)$.
- According to the jurisdiction rule, we obtain:
  $S \mid\equiv C_i$.
- According to $S \mid\equiv U_i \mid\sim C_i$, $SK = s \cdot C_i$ and Assumption A.9, we obtain:
  $S \mid\equiv (U_i \overset{SK}{\longleftrightarrow} S)$ (**Goal 2**).

3.2 Discussion on Possible Attacks

In the following, we demonstrate that our scheme is able to withstand DoS attack, off-line password guessing attack, replay attack, server spoofing attack, parallel session attack and impersonation attack. Moreover, our scheme achieves mutual authentication and users' anonymity.

We assume that the computation Diffie-Hellman problem (CDHP) in the elliptic curves is difficult to be solved in polynomial time.

CDHP: Given two points $s \cdot P, t \cdot P$, where $s, t \in Z_p^*$, the computation Diffie-Hellman problem (CDHP) is to find the point $(s \cdot t)P$ on $E_p(a, b)$.

3.2.1 DoS Attack

Completely Automated Public Turing test to tell Computer and Humans Apart(CAPTCHA) technique is used in our proposed scheme which makes the malicious attacker cannot launch DoS attack. When users login in the remote server S, they must reply S an answer responding to the CAPTCHA puzzle. These puzzles are difficult for computers to solve, and thus the DoS attack which launched by computers is resisted effectively.

3.2.2 Off-line Password Guessing Attack

In off-line password guessing attack, the adversary attempts to guess the identity ID_i and password PW_i from the intercepted messages transmitted between U_i and S. If an adversary eavesdrops U_i's login request message $\{C_i, D_i\}$, which $C_i = t \cdot G$, $D_i = E_{K_i}(ID_i, H(x \| n_i))$. It is impossible to obtain ID_i in real polynomial time due to the difficulty of CDHP in elliptic curve cryptosystem.

3.2.3 Mutual Authentication and Users' Anonymity

In the authentication and session key exchange phase, the remote server and users can authenticate each other such that no adversary can impersonate any participant in this system. Besides, the message transmitted between users and the server will be updated in each session, therefore, no one can trace the user by eavesdropping. Thus, our proposal provides perfect forward security, mutual authentication and users' anonymity.

Table 2: Comparisons of functionality

	Li et al.'s [21]	Chen et al.'s [7]	Jiang et al.'s [14]	Wei et al.'s [26]	Ours
Prevention of impersonation attack	No	No	Yes	No	Yes
Prevention of off-line password guessing attack	No	Yes	Yes	Yes	Yes
Prevention of server spoofing attack	Yes	Yes	Yes	Yes	Yes
Prevention of replay attack	Yes	Yes	Yes	Yes	Yes
Preserving user anonymity	No	No	No	No	Yes
Parallel session attack	Yes	Yes	Yes	Yes	Yes
Mutual authentication	Yes	Yes	Yes	Yes	Yes
Perfect forward secrecy	Yes	No	No	Yes	Yes

Table 3: Performance comparisons: Computation cost

Type of operations	Li et al.'s [21]	Chen et al.'s [7]	Jiang et al.'s [14]	Wei et al.'s [26]	Ours
T_H	7	8	6	12	5
T_{sym}	0	0	0	0	2
T_{asy}	8	6	6	4	8

3.2.4 Replay Attack

The replay attack is that attackers re-submit the login message transmitted between users and the server to impersonate users. In our scheme, neither the replay of an old login message $\{C_i, D_i\}$ in the login phase nor the replay of the response message $\{E_i, F_i\}$ of the server in the authentication and session key exchange phase, it will fail in Step 2 and Step 4 of authentication and session key exchange phase, due to the random numbers are updated for every session and the adversary cannot get the real one. Therefore, our scheme can withstand replay attack.

3.2.5 Parallel Session Attack

The parallel session attack is impossible to be launched in our scheme, due to message structure transmitted between users and the server is different. Both $\{D_i, E_i, F_i\}$ and $\{M_i, N_i\}$ have different structures, so the adversary is not able to perform such an attack.

3.2.6 Perfect Forward Secrecy

Suppose the long-term secret key x is revealed by an adversary, he/she cannot derive U_i's previous session key $SK = s \cdot t \cdot G$ since they are contributed by two selected random numbers. Moreover, even the user's previous login request message $\{C_i, D_i\}$ is eavesdropped by the attacker, he/she also cannot obtain s and t. Thus, the proposed scheme is able to ensure perfect forward secrecy.

3.2.7 Impersonation Attack

An adversary can obtain $B_i = E_{A_i}(H(x\|n_i), n_i \cdot G)$, which is stored in U_i's smart card. Then, he/she needs to forge a valid login request $\{C_i, D_i\}$, in which $C_i = t \cdot G$, $D_i = E_{t \cdot Pub_S}(ID_i, H(x\|n_i))$. Nevertheless, it is impossible for the adversary to compute them without password and identity of U_i. Further, we have demonstrated that our proposed scheme could achieve the security of identity and password in the above. Thus, the attacker cannot forge the valid login request to impersonate U_i and launch such an attack.

3.2.8 Server Spoofing Attack

As illustrated above, our enhanced scheme achieves mutual authentication between users and the remote server. Moreover, the attacker cannot obtain t, s and n_i from the message $\{E_i, F_i\}$. And hence, he/she has no ability to calculate the session key $SK = s \cdot t \cdot G$. Thus, our scheme can avoid the server spoofing attack.

4 Performance Evaluation

In this section, we will evaluate the performance and functionality of the proposed scheme, and then make comparisons with Li et al.'s [21], Chen et al.'s [7], Jiang et al.'s [14] and Wei et al.'s [26] schemes. Let T_H be the time complexity for one-way hash function operations; T_{sym} indicates the time complexity of asymmetric encryption and T_{asy} is defined as the time complexity of the symmetric encryption.

Table 2 lists the functionality comparisons of the proposed scheme and other related schemes. We can see that Li et al.'s, Chen et al.'s, Jiang et al.'s and Wei et al.'s schemes satisfy only five, five, six and six requirements list in Table 2, respectively. While the proposed scheme can achieve all requirements list in Table 2. As a result, the proposed scheme is more secure and has more functionalities compared with these related schemes.

From Table 3, we can find that the total computation cost of Li et al.'s, Chen et al.'s, Jiang et al.'s, Wei et al.'s and our proposed schemes are $7T_H + 8T_{asy}$, $8T_H + 6T_{asy}$,

$6T_H + 6T_{asy}$, $12T_H + 4T_{asy}$, $5T_H + 2T_{sym} + 8T_{asy}$. Compared with other related schemes, our scheme is slightly efficient than Li et al.'s scheme and needs more computational cost than other schemes. Nevertheless, these schemes are insecure and our scheme can satisfy more admired criterion compared with them.

5 Conclusions

In this paper, we propose a secure authentication scheme using CAPTCHA technique. Then, we present its formal proof using the BAN-logic. Furthermore, the discussions on possible attacks shows that the robustness of the proposal. By comparing with several related schemes, our scheme satisfies many admired criterion to suit for practical application.

Acknowledgments

The authors gratefully acknowledge the anonymous reviewers for their valuable comments.

References

[1] D. S. AbdElminaam, H. M. A. Kader, M. M. Hadhoud and S. M. EI-Sayed, "Increase the Performance of Mobile Smartphones using Partition and Migration of Mobile Applications to Cloud Computing," *International Journal of Electronics and Information Engineering*, vol. 1, no. 1, pp. 34–44, 2014.

[2] A. K. Awasthi, "Comment on a dynamic ID-based remote user authentication scheme," *Transactions on Cryptology*, vol. 1, no. 2, pp. 15–16, 2004.

[3] G. T. Becher, *Intentional and Unintentional Sidechannels in Embedded Systems*, University of Massachusetts Amherst, 2014.

[4] M. Burrows, M. Abadiand R. Needham, "A logic of authentication," *ACM Transactions on Computer Systems*, vol. 8, no. 1, pp. 18–36, 1990.

[5] C. C. Chang and C. Y. Lee, "A smart card-based authentication scheme using user identify cryptography," *International Journal of Network Security*, vol. 15, no. 2, pp. 139–147, 2013.

[6] Y. F. Chang and P. Y. Chang, "An improved user authentication and key agreement scheme providing user anonymity," *Journal of Electronic Science and Technology*, vol. 4, no. 9, pp. 352–358, 2011.

[7] B. L. Chen, W. C. Kuo and L. C. Wuu, "Robust smart-card-based remote user password authentication scheme," *International Journal of Communication Systems*, vol. 27, no. 2, pp. 377–389, 2014.

[8] M. L. Das, A. Saxena, and V. P. Gulati, "A dynamic ID-based remote user authentication scheme," *IEEE Transactions on Consumer Electronics*, vol. 50, no.2, pp. 629–631, 2004.

[9] C. I. Fan, Y. C. Chan and Z. K. Zhang, "Robust remote authentication scheme with smart cards", *Computers & Secutity*, vol. 24, no. 8, pp. 619–628, 2005.

[10] D. L. Guo and F. T. Wen, "A more robust authentication scheme for roaming service in global mobility networks using ECC," *International Journal of Network Security*, vol. 18, no. 2, pp. 217–233, 2016.

[11] D. L. Guo, Q. Y. Wen, W. M. Li, H. Zhang and Z. P. Jin, "A Novel Authentication Scheme Using Self-certified Public Keys for Telecare Medical Information Systems," *Journal of Medical Systems*, vol. 39, no. 6, pp. 1–8, 2015.

[12] D. L. Guo, Q. Y. Wen, W. M. Li, H. Zhang and Z. P. Jin, "Analysis and Improvement of Chaotic Map Based Mobile Dynamic ID Authenticated Key Agreement Scheme," *Wireless Personal Communications*, vol. 83, no. 1, pp. 35–48, 2015.

[13] M. S. Hwang, C. C. Lee, S. K. Chong and J. W. Lo, "A Key Management for Wireless Communications," *International Journal of Innovative Computing, Information and Control*, Vol. 4, No. 8, pp. 2045–2056, 2008.

[14] Q. Jiang, J. Ma, G. Li and X. Li, "Improvement of robust smart-card-based password authentication scheme," *International Journal of Communication Systems*, vol. 28, no. 2, pp. 383–393, 2015.

[15] P. Kocher, J. Jaffe and B. Jun, "Differential power analysis," *Proceedings of Advances in Cryptology*, Santa Barbara, CA, USA, pp. 388–397, 1999.

[16] S. Kumari, "Design flaws of "An anonymous two-factor authenticated key agreement scheme for session initiation protocol using elliptic curve cryptography," *Multimedia Tools and Applications*, vol. 76, no. 2, pp. 1801–1815, Jan. 2017.

[17] S. Kumari, M. K. Khan and M. Atiquzzaman, "User authentication schemes for wireless sensor networks: A review," *Ad Hoc Networks*, vol. 27, pp. 159–194, 2015.

[18] S. Kumari, M. K. Khan and X. Li, "A more secure digital rights management authentication scheme based on smart card," *Multimedia Tools and Applications*, vol. 75, no. 2, pp. 1135–1158, 2016.

[19] S. Kumari, X. Li, F. Wu, A. K. Das, H. Arshad and M. K. Khan, "A user friendly mutual authentication and key agreement scheme for wireless sensor networks using chaotic maps," *Future Generation Computer Systems*, vol. 63, pp. 56–75, 2016.

[20] L. Lamport, "Password authentication with insecure communication," *Communications of the ACM*, vol. 24, no. 11, pp. 770–772, 1981.

[21] X. Li, J. W. Niu, M. K. Khan and J. G. Liao, "An enhanced smart card based remote user password authentication scheme," *Journal of Network and Computer Applications*, vol. 36, no. 5, pp. 1365–1371, 2013.

[22] T. S. Messerges, E. A. Dabbish, and R. H. Sloan, "Examining smart-card security under the threat

of power analysis attacks," *IEEE Transactions on Computers*, vol. 5, no.51, pp. 541–552, 2002.

[23] H. S. Rhee, J. O. Kwon and D. H. Lee, "A remote user authentication scheme without using smart cards," *Computer Standards & Interfaces*, vol. 31, no. 1, pp. 6–13, 2009.

[24] R. C. Wang, W. S. Juang, and C. L. Lei, "Robust authentication and key agreement scheme preserving the privacy of secret key," *Computer Communications*, vol. 3, no. 34, pp. 274–280, 2011.

[25] X. M. Wang, W. F. Zhang, J. S. Zhang and M. K. Khan. "Cryptanalysis and improvement on two efficient remote user authentication scheme using smart cards," *Computer Standards & Interfaces*, vol. 5, no. 29, pp. 507–512, 2007.

[26] J. H. Wei, W. F. Liu and X. X. Hu, "Secure and efficient smart card based remote user password authentication scheme," *International Journal of Network Security*, vol. 18, no. 4, pp. 782–792, 2016.

[27] L. von Ahn, M. Blum, N. Hopper and J. Langford, "CAPTCHA: Using Hard AI Problems for Security," *Advances in Cryptology (EUROCRYPT'03)*, pp. 294–311, Warsaw, Poland, 2003.

Biography

Guifa Hou received the M.S. degree in Computer software & theory from University of Science & Technology Beijing, Beijing, China, in 2007. He is now an associate professor in Anyang Institute of Technology, Henan, China. His research works focus on information integration and information security.

Zhijie Wang received the B.S. degree in Electrical Technology from Henan University, Kaifeng, Henan, China, in 1999, and the M.S. degree in Computer Applied Technology from Jiangsu University, Zhenjiang, Jiangsu, China, in 2008. His research interests include data mining and information security.

Secure Data Outsourcing on Cloud Using Secret Sharing Scheme

Arup Kumar Chattopadhyay[1], Amitava Nag[2] and Koushik Majumder[3]
(Corresponding author: Arup Kumar Chattopadhyay)
Department of Computer Science and Engineering, Academy of Technology[1]
Adisaptagram, Aedconagar, Hooghly 712121, West Bengal, India
Department of Information Technology, Central Institute of Technology[2]
Kokrajhar 783370, BDAT, Assam, India
Maulana Abul Kalam Azad University of Technology, West Bengal, India[3]
(Email: ardent.arup@gmail.com)
(Received July 12, 2016; revised and accepted Nov. 15 & Dec. 25, 2016)

Abstract

Data Outsourcing in Cloud (DOC) has its exclusive benefits like low-cost, lower management overhead, elasticity of storage etc and these encourage organizations to use cloud computing to outsource massive amount of data to the cloud providers. The outsourced environment of the cloud and its inherent loss of control cause risk of exposing highly sensitive data to internal or external attacks. Traditionally, the data are kept encrypted to have secure authorized-only access. But, encrypting and decrypting large data files are computationally costly. Hence, secret sharing based DOC schemes have emerged due to their low complexity. Here, the proposed scheme uses simple Boolean based encryption and decryption of the data files (only image-files are considered in this paper) which is low in computational cost. The encrypted data files will be shared on the cloud. A threshold (t, n)-secret sharing scheme applied on the symmetric key of the encryption algorithm. The n share-keys will be generated from the secret key and will be distributed among participants. If t or more ($\leq n$) shared-keys are submitted, the original data files can be retrieved. Hence, it allows a threshold authorized-group of t or more data-users.

Keywords: Cloud Computing; DOC; Image Encryption; Secret Sharing

1 Introduction

Cloud computing was primarily developed for resource sharing with the motive of high availability, scalability, efficiency, cost-effectiveness of deploying utilities on network. Today cloud is a cost effective, flexible and on demand service delivery platform for providing business online. The key benefits of cloud computing are as follow:

1) Rapid Deployment: Deployment of hardware and software resources on cloud is on demand.
2) Availability: Resources on cloud is available anytime and anywhere.
3) Cost Reduction: It reduces the cost invested on hardware and software.
4) Scalability: Cloud can scale up or down its availability of resources like storage, computing depending on the varying needs of client.
5) Efficiency: Sharing the resources on cloud also provides an optimal utilization of resources.
6) Easier Collaboration: The resources shared on cloud can be accessed by many users at the same time from heterogeneous platforms. Thus it provides a collaborative approach for resource use.

Authors in [9] have described the key characteristics of cloud technology such as Multi-tenancy, Massive Scalability, Rapid Elasticity, Measured Service etc.

Data Outsourcing in Cloud (DOC) is a new paradigm where data are stored onto a trusted third-party service provider, such as cloud file server, cloud data server etc. The benefits of DOC as described in [5, 35] are on-demand and high quality service, universal data access by data-users regardless of their location and cost reduction in hardware and software resources.

Security remains the critical issue for the data outsourced on cloud. Data owner leverages on the service provider's hardware and software for storing and managing the outsourced data because it is a cost-effective and efficient solution, but at the same time the data owner loses control over the sensitive or confidential data which may be disclosed to unauthorized users [8]. As a result some customers are unwilling or unable to entrust their raw sensitive data to cloud providers. Authors in [16, 21]

have discussed the security issues for cloud. These issues include - cloud integrity, privacy, confidentiality and availability. The unique security requirements for cloud computing are identified in [43]. In such situation, different solutions for securing the confidentiality of the outsourced data has been developed. The authors in [11] discussed different types of attack on cloud which can be security threat for the outsourced data at cloud and authors in [4] have summarized the key strategies to be used to enhance the security of data stored at cloud environment and compare those techniques [29].

The traditional approaches used to secure the outsourced data on cloud are based on encryption. Both symmetric and asymmetric encryption schemes are used to conceal the original data and provide access structure to authorize the user to access the data. The use of encryption techniques to secure the data on cloud are discussed in [15, 20, 28, 34, 40, 46]. Vimercati et al. in [40] proposed a scheme where a two layer encryption is to be imposed on data to manage the access control on outsourced data. In 2007, Chase [10] has put forward Multi-Authority Authority Based Encryption (ABE), in this scheme a user is identified by a set of attributes and functions to determine the ability of the user to decrypt the cipher and it provides a fine grain access control. The scheme in [10] was improved by Li et al. [22] and defined with lesser ciphertexts and user's secret keys, along with ability to the encrypter to determine the number of attributes are desired for each ciphertext. The attribute based access control proposed in [20] is based ciphertext-policy attribute based encryption (CP-ABE) to enforce access control policies with efficient attribute and user revocation capability. Lu et al. in [28] proposed fine grained content level access control by hiding the paintext data content and issuing search token with decryption keys. Raykova et al. in [34] proposed a two-level access control scheme which handles both read and write access control for data users and data owners. Zhou et al. in [46] have proposed a tree-based key management scheme that allows the outsourced data to be accessed by multiple parties who hold different rights. Tree-based key management method of access hierarchies for data outsourcing is also discussed by W. Wang et al. in [41]. The authors in [12] have proposed a new public-key cryptosystems (that compress secret keys) that produce constant size ciphertext and enable efficient delegation of decryption rights for any set of ciphertexts. Giweli et al. [15] proposed a secure DOC method based on a combination of cryptography techniques, including the Chines Remainder Theorem, symmetric and asymmetric encryption. Fu et al. in [14] proposed a verifiable outsourced ciphertext decryption based on prime order bilinear group which made the scheme efficient for data consumer.

Achieving confidentiality of outsourced data using encryption is not computationally efficient as the encryption/decryption operations are time consuming. Most of these schemes use specialized encryption with complex mechanism, ranging from order-preserving encryption [2], which has limited security but is highly efficient, to oblivious RAM [37], which has provable access pattern and is highly secure but with poor performance. In recent years, a number of DOC schemes have been proposed with secret sharing. The authors in [1, 18, 39] proposed secure data outsourcing with Shamir's threshold secret sharing scheme [36], in which the secret (the data file) will be encoded into n shares or pieces and those n shares or pieces will be kept on n different data-storage-servers in the cloud; knowledge of any t or more shares will retrieve the whole secret (data file). But, Dautrich and Ravishankar [13] have shown that these schemes are vulnerable to the collusive attack in which any t colluding servers can recover all the files outsourced to the cloud. Muhammad et al. in [30] proposed secure data outsourcing based on Asmuth-Bloom secret sharing scheme (Asumuth-Bloom used Chinese Remainder Theorem in their proposed scheme [6]). Authors in [3, 31, 33] proposed secret sharing for securing multi-cloud i.e. a collection of several cloud infrastructures. Another fast and secure data outsourcing scheme is proposed by Liu at el. [27] based on Shamir's secret sharing scheme. It split the data files into public and private shares. Public shares are available in the cloud where as private shares are with data users. If valid data users submit their private shares, then only full data file can be retrieved from the cloud.

We assume honest-but-curious server threat model, commonly used for data outsourcing in cloud. The objective of the scheme described in this paper, is to secure the outsourced data with the combined effect of encryption and secret sharing. The data to be outsourced will be encrypted with a key and the key will be shared by a secret sharing algorithm to generate a threshold access structure to retrieve the key in full. Hence, the permissible group having threshold number of members will be able to retrieve all the secrets, encrypted using the key.

The rest of the paper is organized as follows: Section 2 defines the entities from cloud computing and secret sharing to be used in our scheme, Section 3 describes the algorithms to be used at proposed scheme, Section 4 describes the proposed model, a few experimental results are shown in Section 5, the security analysis is there in Section 6, and the scheme in concluded at Section 7.

2 Preliminaries

The entities from the cloud computing and secret sharing schemes to be used in the proposed model are described as follows.

2.1 Entities of Cloud Computing

Data Owner. Data owner is the actual possessor of the sensitive and confidential data files to be outsourced on the cloud. The data owner does not prefer to store the raw data with trusted third-party server, rather the encrypted version - cipher-steam will be

outsourced and store at cloud data-storage-server.

Data User. Data user is the end-user trying to access the secure outsourced data on cloud. The data user must have enough privilege such that he/she will be called an authorized user permitted to access the secret data on cloud.

Cloud Storage Server. These are the collection of data-storage-servers in the cloud that stores the data files outsourced on the cloud and provides location transparency to the data owner and data users.

2.2 Entities of (t, n)-threshold Secret Sharing

Secret. In secret-sharing schemes, the secret S is the data or data file which must the secured from the unauthorized user or unauthorized group. In the proposed scheme, the secret S is the encrypted version of the data-file to the stored in cloud.

Shares. The secret S will be encoded in n pieces called shares or shadows. Knowledge of t or more shares ($\leq n$) reveal the secret in full, any less than t shares will not reveal any secret.

Dealer. Dealer is the owner of the secret, who generates the shares and distributes them among n participants.

Participants. The participants are the users seeking the secret. If t or more participants submit their shares the secret can be revealed.

Combiner. The combiner is responsible for decoding of the secret. If t or more shares are submitted to the combiner then the combiner can decode the secret. Otherwise no information about the secret can be revealed.

3 Related Works

The secret sharing scheme used in our proposed model is based on Shamir's (t, n)-threshold secret sharing (1979) [36] and Thien-Lin image secret sharing scheme (2002) [38] (which extended Shamir's scheme for secret image sharing (SIS)). In the following subsections we discuss Shamir's scheme and the image encryption scheme using affine transformation and XOR operations as proposed by Nag et al. [32].

3.1 Shamir's Secret Sharing Scheme

Shamir in [36] has proposed a $(t, n)-$ threshold secret sharing scheme based on Lagrange interpolation. In a $(t, n)-$ threshold secret sharing scheme, a secret S is encoded into n parts called shadows or shares and distributed among n participants or players. If any t (or more) shares are obtained then the full secret S can be reconstructed and no less than t shares can reconstruct the secret or expose any information about the secret. As the decoding criteria depends on minimum number (t) of participants who can reconstruct the secret, t is called threshold. As proposed by Shamir, the (t, n)-threshold secret sharing scheme requires $(t-1)$ degree of polynomial:

$$q(x) = (a_0 + a_1 x + a_2 x^2, \cdots, a_{t-1} x^{t-1}) mod\ p, \quad (1)$$

where $a_0 = S$, the secret, p is a large prime, $a_0, a_1, \cdots, a_{t-1}$ are coefficients are in $GF(p)$.

Construction of Shares: The n shares s_i where $i = 1, 2, \cdots, n$ can be generated as follows:

$$s_1 = q(1), \cdots, s_i = q(i), \cdots, s_n = q(n).$$

Reconstruction of Secret: If t or more shares are submitted then the polynomial $q(x)$ can be regenerated by Lagrange interpolation theorem as follows:

$$q(x) = \sum_{j=1}^{t} s_j \prod_{m=1, m\neq j}^{t} \frac{x - x_m}{x_j - x_m} mod\ p.$$

The secret S can be determine as

$$S = q(0).$$

Several improvements have been proposed on the Shamir's scheme to detect and identify cheaters - disloyal participants who provide faked shares to deceive the honest participants (because they obtain an incorrect secret) or dishonest dealer supplies fake shares to the participants. Harn et al. [19] and Liu et al. [26] proposed verifiable secret sharing scheme based CRT and extended Asmuth-Bloom scheme. Those schemes are further improved by Liu et al. [25] by introducing single way hash function to ensure the secrecy maintained for both the shares and secret.

3.2 Image Encryption Using Affine Transform and XOR Operation

Nag et al. proposed an encryption scheme [32] in 2011 effective for images. The scheme considered eight 8-bit keys, say $K_0, K_1, K_2, K_3, K_4, K_5, K_6$ and K_7. The two-phase encryption process is as follows.

Phase 1: Redistribution of the pixels to the new locations using affine transformation with four keys breaks the strong correlation between the pixels. Let the old pixel position be (x, y), the translated location (x', y') can be calculated as

$$\begin{aligned} x' &= (K_0 + K_1 \times x) mod\ n \\ y' &= (K_2 + K_3 \times y) mod\ n. \end{aligned}$$

Phase 2: The image will be decomposed into $\frac{n}{2} \times \frac{n}{2}$ blocks recursively till each block will be of size (2×2). Encrypt the block pixels (p_1, p_2, p_3 and p_4) using keys - K_4, K_5, K_6 and K_7 as follows:

$$\begin{aligned} cp_1 &= p_1 \oplus K_4 \\ cp_2 &= p_2 \oplus K_5 \\ cp_3 &= p_3 \oplus K_6 \\ cp_4 &= p_4 \oplus K_7. \end{aligned}$$

Apply the procedure for all the blocks. The decryption can be done when all 8 keys are available.

Phase 1: The encrypted pixels in the blocks will be encrypted as

$$\begin{aligned} p_1 &= cp_1 \oplus K_4 \\ p_2 &= cp_2 \oplus K_5 \\ p_3 &= cp_3 \oplus K_6 \\ p_4 &= cp_4 \oplus K_7. \end{aligned}$$

Phase 2: The image pixels will be re-positioned as

$$\begin{aligned} x &= ((x' + (-K_0)) \times K_1^{-1}))mod\ n \\ y &= ((y' + (-K_2)) \times K_3^{-1}))mod\ n. \end{aligned}$$

The same scheme has been applied for the proposed model with some modification - (1) the key size is 16-bit consist of four 4-bit subkeys used for affine transformation, and (2) a key matrix of size the same size as the target image ($w \times h$) is used rather (2×2) matrix with four 8-bit keys; the key matrix directly gets XORed with the target image.

4 Proposed Model

Let the m secret images are $SI_1, SI_2, \cdots, SI_m$ of fixed size $w \times h$ and a random matrix R (Combiner Secret) of same size. Let number of participants be n. The objective is that if any t or more participants ($\leq n$) submit their keys then all the images can be retrieved. The phases are as follows.

4.1 Encryption of Secret Images and Distribution of Secret Keys

Step 1: The data owner submit m secret images to the dealer. The dealer encrypts the secret images SI_i, $\{i = 1\ to\ m\}$ with combiner secret R and a 16-bit key K (having four 4-bit subkeys - k_1, k_2, k_3, and k_4). The encoded images EI_i are generated as follows: For each image SI_i in $i = 1, 2, \cdots, m$ apply affine transformation as

$$\begin{aligned} x' &= (k_1 + k_2 \times x)mod\ w \\ y' &= (k_3 + k_4 \times y)mod\ h \end{aligned}$$

where (x, y) is the old pixel position and (x', y') is the new position. Then, the translated images, SI_i' will be XORed with R as

$$EI_i = SI_i' \oplus R, \quad \{for\ i = 1\ to\ m\}.$$

Step 2: The encoded images will be stored in cloud and the combiner secret R will be encrypted (with affine transformation) to C_R with a key k_1, k_2, k_3, k_4 and will be stored with the combiner. For each pixel position (x, y) of R, calculate the translated position (x', y') using affine translation.

$$\begin{aligned} x' &= (k_1 + k_2 \times x)mod\ w \\ y' &= (k_3 + k_4 \times y)mod\ h. \end{aligned}$$

Step 3: Generate n shares $sk_1, sk_2, \cdots, sk_n$ from the 16-bit key, K using *Shamir Secret Sharing Scheme*.

Step 4: Distribute key shares sk_i to P_i for $i = 1, 2, \cdots, n$.

4.2 Retrieval of Secret Images

Step 1: Let t or more participants ($\leq n$) submit their key shares sk_i to the Combiner.

Step 2: Combiner uses *Shamir's (t, n)-threshold Secret Sharing Scheme* to reconstruct the secret key, K. Thus, from K, four 4-bit subkeys- k_1, k_2, k_3, k_4 become available now.

Step 3: The encrypted combiner secret C_R will be decrypted with k_1, k_2, k_3, k_4 and the combiner key R will be retrieved as

$$\begin{aligned} x &= (x' + (-k_1)) \times k_2^{-1})\ mod\ w \\ y &= (y' + (-k_3)) \times k_4^{-1})\ mod\ h. \end{aligned}$$

Step 4: Combiner fetches all m encoded images, EI_i and decodes them as follows: First, EI_i will be XORed with R:

$$SI_i' = EI_i \oplus R, \quad \{for\ i = 1\ to\ m\}$$

Then, apply affine re-translation for each SI_i' for $i = 1, 2, \cdots, m$ as

$$\begin{aligned} x &= (x' + (-k_1)) \times k_2^{-1})\ mod\ w \\ y &= (y' + (-k_3)) \times k_4^{-1})\ mod\ h \end{aligned}$$

where (x, y) is the original location of the pixel.

Step 5: The decoded images will be made available to the t or more requesting participants.

A comparative study between different encryption schemes proposed for DOC is show in Table 1.

Table 1: Comparisons between proposed encryption scheme with other related scheme

Schemes	Decryption key size	Cipher-text size	Encryption type
Proposed Scheme	key-size for affine transformation is fixed; key-size for XOR is in $O(plaintext)$	$O(plaintext)$	symmetric-key
Key assignment schemes for a predefined hierarchy in [24]	depends on the hierarchy	$O(plaintext)$	symmetric or public-key
Symmetric-key encryption with Compact Key in [7]	fixed	$O(plaintext)$	symmetric-key
Attribute-Based Encryption in [17]	variable size	$O(plaintext)$	public-key
Key-Aggregate Cryptosystem based Encryption in [12]	fixed	$O(plaintext)$	public-key

5 Observations

Consider the secret image SI (as shown in Figure 1a) which has to be uploaded in the cloud. Consider a 16-bit key, $K = 20715$ (0110101110000101) as secret, $S = K$. Using Shamir's (t, n)-threshold secret sharing scheme where $n = 5, t = 3$, the shares are generated as follows:

$$s_1 = 20748 \quad (0001100001000101),$$
$$s_2 = 20805 \quad (0101000101000101),$$
$$s_3 = 20886 \quad (0011010011000101),$$
$$s_4 = 20991 \quad (0111111111000101),$$
$$s_5 = 21120 \quad (0000000010100101).$$

The subkeys used for encryption from key K are - k_1, k_2, k_3 and k_4 as follows:

$$\begin{aligned} k_1 &= 0110 = 6, \\ k_2 &= 1011 = 11, \\ k_3 &= 1000 = 8, \\ k_4 &= 0101 = 5. \end{aligned}$$

Consider combiner secret R (random matrix of the size of original image) as shown in Figure 1b. Use the combiner secret R to encrypt the original image SI. The result encrypted image shown in Figure 1c.

Applying Affine transformation to translate the pixel positions with keys k_1, k_2, k_3 and k_4 we get the encrypted image shown in Figure 1d. The Combiner Secret, R is also encrypted by keys, k_1, k_2, k_3 and k_4 to generate the encrypted version of Combiner Secret, C_R and will be stored with the Combiner.

Considering any 3 shares say, s_1, s_2, s_3 the key K and the subkeys are reconstructed as: $K = 20715$ $(k_1 = 0110 = 6;\ k_2 = 1011 = 11;\ k_3 = 1000 = 8;\ k_4 = 0101 = 5)$.

Using the keys Combiner perform affine transformation to re-translate the pixels to reconstruct the Combiner Secret R from C_R. The same keys are used to re-translate the encrypted image, as shown in Figure 2a.

Use Combiner Secret R to farther decrypt the image. The obtained image is shown in Figure 2b.

6 Security Analysis

The original image and the final encrypted image are as shown in Figure 3a and Figure 3b respectively.

6.1 Key Space Analysis

To resist brute force attack the key space should be large enough to make the attack infeasible. In proposed scheme, the key used for the XOR operation is the same size of the image i.e. $w \times h$. Hence, the key space will be $2^{w \times h}$. So, for a very small image of (64×64), they key space is $2^{64 \times 64}$, which is quite large. Affine cipher is created by 4 keys each of size 4 bits, key space for affine translation is 2^{16}. So, total key space will be $2^{(16 \times w \times h)}$, which is large enough if the image is not extremely small.

6.2 Statistical Analysis

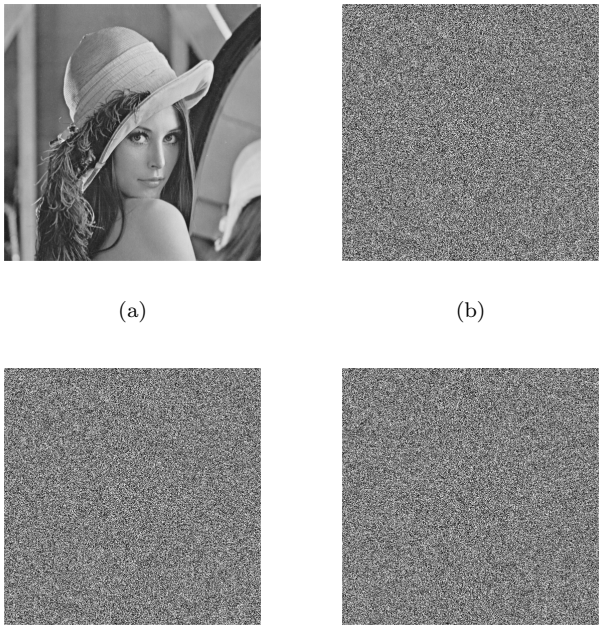
To resist statistical attacks, large amount of diffusion and confusion needs to be introduced in the cipher image. The following are the statistical analysis to ensure that it can prevent statistical attacks.

6.2.1 Histograms of Corresponding Images

Histogram of a given image reflects the distribution information of the pixel values. The histogram of an ideal encrypted image should have a uniform distribution and it will be completely different from the histogram of the plain-image. Histograms of the (512×512) grayscale image (of Lena.tif) and the cipher image are as shown in figures 4a and 4b. It is clearly visible that the histogram of the ciphered image in Figure 4b is fairly uniformly distributed, this is important in resisting statistical analysis attack.

6.2.2 Correlations of Two Adjacent Pixels

In a plain-image the adjacent pixels are highly correlated in either horizontal, vertical or diagonal direction. It is an important challenge for an encryption technique to reduce the correlation significantly in cipher image. To test the correlation of plain-image and cipher-image, the



(a) (b)

(c) (d)

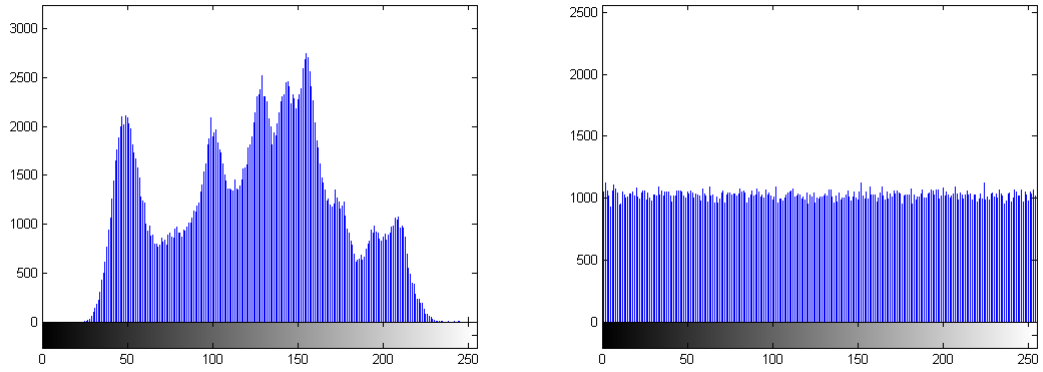
Figure 1: Construction of the encrypted image - (a) is the original image; (b) Combiner Secret R; (c) image after XORed with Combiner Secret R; (d) image after affine translation (this is the final image to be uploaded on cloud)

(a) (b)

Figure 2: Reconstruction of the secret image.(a) image after affine re-translation of the pixels; (b) image after XORed with Combiner Secret R

(a) (b)

Figure 3: (a) is the original image; (b) final encrypted image



(a) (b)

Figure 4: (a) is the histogram of the original image; (b) is the histogram of the final encrypted image

correlation coefficients of the adjacent pixels in vertical, horizontal and diagonal directions are evaluated by using following equations:

$$
\begin{aligned}
r_{xy} &= \frac{cov(x,y)}{\sqrt{D(x)}\sqrt{D(y)}}, \\
E(x) &= \frac{1}{S}\sum_{i=1}^{S} x_i, \\
D(x) &= \frac{1}{S}\sum_{i=1}^{S}(x_i - E(x))^2, \\
cov(x,y) &= \frac{1}{S}\sum_{i=1}^{S}(x_i - E(x))(y_i - E(y)),
\end{aligned}
$$

where x and y are gray values of adjacent pixels and S is total number of duplets (x, y) obtained from image. $E(x)$ and $D(x)$ are the expectation and variance of x, respectively. As shown in Table 2, there is high correlation in the original image, but correlation in cipher image is negligible. Comparison between proposed scheme with few important encryption schemes (with respect to correlation-values) is shown in Table 2.

6.2.3 Information Entropy Analysis

The information entropy is the indicator of the randomness in the image, which can be calculated as

$$H(s) = \sum_{i=0}^{2^{N-2}} p(s_i) log_2 \frac{1}{p(S_i)},$$

where $p(s_i)$ is the probability of variable $s(i)$. For true random source producing 2^L, the entropy should be L. For, a gray-scale image with 8-bit pixels can have 2^8 different values (that is 0 to 255). Thus, entropy of a true random image must be 8. A close value of 8 of information entropy for a cipher image indicates, it is random enough.
Information Entropy Analysis for Lena.tif (512×521)

Entropy value of Original Image	=	7.4451
Entropy value of Cipher Image	=	7.9992

6.2.4 Sensitivity Analysis

Sensitivity analysis is another measure of randomness, have two indicators (1) NPCR (number of pixels change rate) and (2) UACI (unified average changing intensity). NPCR and UACI are computed by the following equations:

$$
\begin{aligned}
NPCR &= \frac{\sum_{i,j} D(i,j)}{W \times H} \times 100\% \\
UACI &= \frac{1}{W \times H}\sum_{i,j}\frac{|c_1(i,j) - c_2(i,j)|}{255} \times 100
\end{aligned}
$$

where c_1 and c_2 are two images of size $W \times H$. If $c_1(i,j) \neq c_2(i,j)$, then $D(i,j) = 1$, otherwise $D(i,j) = 0$.
Sensitivity Analysis for Lena.tif (512×521)

NPCR Score	=	0.996105194091797
UACI Score	=	0.286168850169462

NPCR and UACI values of proposed image encryption and few other important encryption schemes are as shown in Table 3.

6.3 Security Analysis of Data Outsourced at Cloud

The following are few security requirements (known as CIA) to be satisfied by DOC.

Data Confidentiality. The outsourced data on cloud should not be revealed to an unauthorized data user or data group. In the proposed (t,n)-scheme, a threshold group of t or more data-users only can access the outsourced data. But, as the scheme is applied on the symmetric keys, not on the data-file, the scheme cannot provides access transparency for

Table 2: Correlation coefficients of two adjacent pixels in three directions

Image (Scheme)	Horizontal	Vertical	Diagonal
Original image	0.97192828	0.98502945	0.95933083
Cipher image (using Proposed Scheme)	0.00201878	0.00136410	0.00186629
Cipher image (using Zhu's Scheme ([47])	0.00201613	0.00091642	0.00165094
Cipher image (using Zhang's Scheme ([45]))	0.00243875	0.00064593	0.00124402
Cipher image (using Wang's Scheme ([42]))	0.00190641	0.00381759	0.00194828

Table 3: Performance of different algorithms

Algorithm	NPCR	UACI
Proposed algorithm	0.996105194	0.286168850
Wang's algorithm in [42]	0.995864868	0.332533000
Zhang's algorithm in [44]	0.995998382	0.310221067
Zhu's algorithm in [47]	0.811958313	0.273860931
Lian's algorithm in [23]	0.033283200	0.007984160

data-files (data-user know the file which content the desired data).

Data Integrity. The outsourced data must be revealed in full to an authorized data-user or data-group whereas no part of the secret data will be revealed to the unauthorized-user or data-group. In the proposed scheme, a group of t or more data-users are authorized to access the data-files, whereas no part of data-file will be revealed to a group consists of less than t users.

Data Availability. Availability of data-files means that the data outsourced on the cloud must be available when required. In our proposed scheme, the same can be achieved by replicating the data-files in cloud.

A common threat to the DOC using secret sharing is described as Collusive attack (in [13]), where multiple cloud storage-servers can collude and can reveal the secret which none of them can do individually. A DOC scheme must be *Collusion-resistance*. In our proposed scheme the data-file is encrypted, where as the key is encoded in n shares and secretly distributed to the n data-users, so collusive attack is not applicable. A comparative study between DOC schemes [1, 18, 27, 39] using secret sharing is as shown in Table 4.

7 Conclusion

The confidentiality and security of the outsourced data on cloud is the concern of the day. Encryption based approaches are computationally inefficient and secret sharing based approaches are threatened by limited security of colluding servers. Thus, the proposed model is a combination of encryption and secret sharing. The data files (only images are considered in the proposed model) are encrypted by simple XOR and affine transformation of pixels' positions which are efficient in term of computation and the encryption key is shared by Shamir's threshold secret sharing scheme, which creates a threshold access structure of the data users. If the threshold or more number of participants submit their keys the data-files can be retrieved in full.

References

[1] D. Agrawal, A. E. Abbadi, F. Emekci, A. Metwally, and S. Wang, "Secure data management service on cloud computing infrastructures," in *Proceeding of the Service and Application Design Challenges in Cloud*, pp. 57–80, 2011.

[2] R. Agrawal, J. Kiernan, R. Srikant, and Y. Xu, "Order preserving encryption for numeric data," in *Proceedings of the 2004 ACM SIGMOD international conference on Management of data*, pp. 563–574, Paris, France, June 2004.

[3] M. K. Alam and S. Banu, "An approach secret sharing algorithm in cloud computing security over single to multi clouds," *International Journal of Scientific and Research Publications*, vol. 3, no. 4, pp. 1–5, 2013.

[4] S. L. Anil and R. Thanka, "A survey on security of data outsourcing in cloud," *International Journal of Scientific and Research Publications*, vol. 3, no. 2, pp. 1–3, 2013.

[5] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. H. Katz, A. Konwinski, G. Lee, D. A. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, "Above the clouds: a berkeley view of cloud computing," Tech. Rep. UCB/EECS-2009-28, Feb. 2009.

[6] C. Asmuth and J. Bloom, "A modular approach to key safeguarding," *IEEE Transactions on Information Theory*, vol. 29, no. 2, pp. 208–210, 1983.

[7] J. Benaloh, M. Chase, E. Horvitz, and K. Lauter, "Patient controlled encryption: Ensuring privacy of

Table 4: Comparison between different SSS-based DOC schemes

Schemes	Data Confidentiality	Data Correctness	Collusion Resistance	Number of public shares	Round of transmission	Types
Proposed Scheme	Yes	Yes	NA	1	1	(t,n)-SSS based
[27]	Yes	Yes	Yes	$n-1$	l	(t,n)-SSS based
[39]	Yes	Yes	No	$l \times t$	l	(t,n)-SSS based
[1]	Yes	Yes	No	$l \times t$	l	(t,n)-SSS based
[18]	Yes	Yes	No	$l \times t$	l	(t,n)-SSS based

electronic medical records," in *Proceedings of the 2009 ACM workshop on Cloud computing security (CCSW'09)*, pp. 103–114, Chicago, IL, USA, Nov. 2009.

[8] 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.

[9] S. Carlin and K. Curran, "Cloud computing technologies," *International Journal of Cloud Computing and Services Science*, vol. 1, no. 2, 2012.

[10] M. Chase, "Multi-authority attribute based encryption," in *Proceeding of the 4th conference on Theory of cryptography (TCC'07)*, pp. 515–534, Amsterdam, Netherlands, Feb. 2007.

[11] R. Chow, P. Golle, M. Jakobsson, E. Shi, J. Staddon, R. Masuoka, and J. Molina, "Controlling data in the cloud: Outsourcing computation without outsourcing control," in *Proceedings of the 2009 ACM workshop on Cloud computing security (CCSW'09)*, pp. 85–90, Chicago, IL, USA, Nov. 2009.

[12] C. K. Chu, S. S. M. Chow, W. G. Tzeng, J. Zhou, and R. H. Deng, "Key-aggregate cryptosystem for scalable data sharing in cloud storage," *IEEE Transactions on Parallel and Distributed System*, vol. 25, no. 2, pp. 468–477, 2014.

[13] J. L. Dautrich and C. V. Ravishankar, "Security limitations of using secret sharing for data outsourcing," in *Proceeding of the IFIP Annual Conference on Data and Application Security and Privacy*, pp. 145–160, Paris, France, 2012.

[14] 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.

[15] N. Giweli, S. Shahrestani, and H. Cheung, "Enhancing data privacy and access anonymity in cloud computing," *Communications of the IBIMA*, vol. 2013, no. 462966, pp. 1–10, 2013.

[16] A. Goel and S. Goel, "Security issues in cloud computing," *International Journal of Application or Innovation in Engineering & Management*, vol. 4, no. 1, pp. 121–124, 2012.

[17] V. Goyal, O. Pandey, A. Sahai, and B. Waters, "Multi-identity single-key decryption without random oracles," in *Proceedings of the 3rd SKLOIS Conference (Inscrypt 2007)*, pp. 89–98, Alexandria, Virginia, USA, Oct. 2006.

[18] M. A. Hadavi and R. Jalili, "Secure data outsourcing based on threshold secret sharing; towards a more practical solution," in *Proceedings of the 36th International Conference on Very Large Data Bases*, pp. 54–59, Singapore, Sept. 2010.

[19] L. Harn, M. Fuyou, and C. C. Chang, "Verifiable secret sharing based on the chinese remainder theorem," *Security and Communication Networks*, vol. 7, no. 6, pp. 950–959, 2014.

[20] J. Hur and D. K. Noh, "Attribute-based access control with efficient revocation in data outsourcing systems," *IEEE Transactions on Parallel and Distributed Systems*, vol. 22, no. 7, pp. 1214–1221, 2011.

[21] R. Kumar, "Cloud computing and security issue," *International Journal Of Engineering And Computer Science*, vol. 5, no. 11, pp. 18823–18826, 2016.

[22] K. Li and H. Ma, "Outsourcing decryption of multi-authority abe ciphertexts," *International Journal of Network Security*, vol. 16, no. 4, pp. 286–294, 2014.

[23] S. Lian, J. Sun, and Z. Wang, "A block cipher based on a suitable use of the chaotic standard map," *Chaos, Solitons & Fractals*, vol. 26, no. 1, p. 117–129, 2005.

[24] S. Lian, J. Sun, and Z. Wang, "Dynamic and efficient key management for access hierarchies," *ACM Transactions on Information and System Security*, vol. 12, no. 3, pp. 18:1–18:43, 2009.

[25] Y. Liu and C. C. Chang, "An integratable verifiable secret sharing mechanism," *International Journal of Network Security*, vol. 18, no. 4, pp. 617–624, 2016.

[26] Y. Liu, L. Harn, and C. C. Chang, "A novel verifiable secret sharing mechanism using theory of numbers and a method for sharing secrets," *International Journal of Communication Systems*, vol. 28, no. 7, pp. 1282–1292, 2015.

[27] Y. Liu, H. L. Wu, and C. C. Chang, "A fast and secure scheme for data outsourcing in the cloud," *KSII Transactions on Internet and Information Systems*, vol. 8, no. 8, pp. 2708–2721, 2014.

[28] Y. Lu and G. Tsudik, "Enhancing data privacy in the cloud," in *Proceedings of the IFIP International Conference on Trust Management*, Singapore, 2011.

[29] 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.

[30] Y. I. Muhammad, M. Kaiiali, A. Habbal, A. S. Wazan, and A. S. Ilyasu, "A secure data outsourcing scheme based on asmuthbloom secret sharing," *Enterprise Information Systems*, vol. 16, no. 4, pp. 1–23, 2016.

[31] M. Muhil, U. H. Krishna, R. K. Kumar, and E. A. M. Anita, "Securing multi-cloud using secret sharing algorithm," *Procedia Computer Science*, vol. 50, pp. 421–426, 2015.

[32] A. Nag, J. P. Singh, S. Khan, S. Biswas, D. Sarkar, and P. P. Sarkar, "Image encryption using affine transform and xor operation," in *Proceedings of the International Conference on Signal Processing, Communication, Computing and Networking Technologies (ICSCCN'11)*, Singapore, July 2011.

[33] P. Pareek, "Cloud computing security from single to multi-clouds using secret sharing algorithm," *International Journal of Advanced Research in Computer Engineering & Technology*, vol. 2, no. 12, pp. 3261–3264, 2013.

[34] M. Raykova, H. Zhao, and S. M. Bellovin, "Privacy enhanced access control for outsourced data sharing," in *Proceeding of the Financial Cryptography and Data Security*, pp. 223–238, Kralendijk, Bonaire, Feb. 2012.

[35] M. N. O. Sadiku, S. M. Musa, and O. D. Momoh, "Cloud computing: opportunities and challenges," *IEEE Potentials*, vol. 33, no. 1, pp. 34–36, 2014.

[36] A. Shamir, "How to share a secret," *Communications of the ACM*, vol. 22, no. 11, pp. 612–613, 1979.

[37] E. Stefanov, E. Shi, and D. Song, "Towards practical oblivious ram," *Cryptography and Security*, pp. 1–40, 2012.

[38] C. C. Thien and J. C. Lin, "Secret image sharing," *Computers and Graphics*, vol. 26, no. 5, p. 765–770, 2002.

[39] X. Tian, C. F. Sha, X. L. Wang, and A. Y. Zhou, "Privacy preserving query processing on secret share based data storage," in *Proceeding of the 16th International Conference on Database Systems for Advanced Applications, (DASFAA'11)*, pp. 108–122, Hong Kong, China, Apr. 2011.

[40] S. D. C. Vimercati, S. Foresti, S. Jajodia, S. Paraboschi, and P. Samarati, "Over-encryption: Management of access control evolution on outsourced data," in *Proceedings of the 33rd International Conference on Very Large Data Bases (VLDB'07)*, pp. 123–134, Vienna, Austria, Sept. 2007.

[41] W. Wang, Z. Li, R. Owens, and B. Bhargava, "Secure and efficient access to outsourced data," in *Proceedings of the 2009 ACM Workshop on Cloud Computing Security (CCSW'09)*, pp. 55–66, Chicago, IL, USA, Nov. 2009.

[42] X. Wang, L. Liu, and Y. Zhang, "A novel chaotic block image encryption algorithm based on dynamic random growth technique," *Optics and Lasers in Engineering*, vol. 66, pp. 10–18, 2015.

[43] M. Zareapoor, P. Shamsolmoali, and M. A. Alam, "Establishing safe cloud: Ensuring data security and performance evaluation," *International Journal of Electronics and Information Engineering*, vol. 1, no. 2, pp. 88–99, 2014.

[44] Y. Q. Zhang and X. Y. Wang, "A symmetric image encryption algorithm based on mixed linearnonlinear coupled map lattice," *Information Science*, vol. 273, no. 20, pp. 329–351, 2014.

[45] Y. Q. Zhang and X. Y. Wang, "A new image encryption algorithm based on non-adjacent coupled map lattices," *Applied Soft Computing*, vol. 26, pp. 10–20, 2015.

[46] M. Zhou, Y. Mu, W. Susilo, J. Yan, and L. Donga, "Privacy enhanced data outsourcing in the cloud," *Journal of Network and Computer Applications*, vol. 35, no. 4, pp. 1367–1373, 2012.

[47] Z. L. Zhua, W. Zhangc, K. W. Wongb, and H. Yua, "A chaos-based symmetric image encryption scheme using a bit-level permutation," *Information Science*, vol. 181, no. 6, p. 1171–1186, 2011.

Biography

Arup Kumar Chattopadhyay has received his BE degree from Visvesvaraya Technological University, Belgaum in 2002 and MTech. degree from University of Calcutta, Kolkata in 2011. He is currently associated with Academy of Technology, India as Assistant Professor. His current research interest is in Information Security.

Dr. Amitava Nag is having 12 years of academic experience and at present is serving as Associate Professor at the Dept. of IT, Central Institute of Technology (CIT), Kokrajhar, India. Prior to joining CIT, he was associated with Academy of Technology, Hooghly as Associate Professor.He received his PhD in Engineering from the University of Kalyani, India. He holds M.Tech in Information Technology from the University of Calcutta, India. His research focuses on Information Security, Cloud Computing and Internet of Things (IoT). He has contributed to numerous research articles in various journals and conferences of repute and is also one of the authors of 5 books. He is a member of the Institution of Engineers.

Koushik Majumder has received his Ph.D from Jadavpur University, Kolkata. He obtained his B. Tech and M. Tech degrees from the University of Calcutta, Kolkata. He is currently associated with Maulana Abul Kalam Azad University of Technology (MAKAUT) as Assistant Professor. His current research interest includes Mobile Adhoc Network, Information Security, Cloud Computing etc.

A Novel Weighted Visual Cryptography Scheme with High Visual Quality

I-Chun Weng, Tzung-Her Chen
(Corresponding author: Tzung-Her Chen)

Department of Computer Science and Information Engineering, National Chiayi University
Chiayi 600, Taiwan, R.O.C.
(Email: thchen@mail.ncyu.edu.tw)
(Received Sept. 6, 2016; revised and accepted Jan. 15, 2017)

Abstract

Visual Cryptography (VC) has been developed to encode a secret image into n shares for n participants in the past decades, in which each share is treated with the same priority. However, the privilege for participants in a group is not always the same. In this paper, a weighted visual cryptography scheme is proposed such that each participant obtains her/his share with different weight according to the different group of predefined privilege. The secret can be disclosed only if stacking predefined k or more shares in which containing predefined some specific shares from the specific groups. Otherwise, no information about the secret can be revealed. It is worthwhile to note that the higher value of total weight of stacking shares; the more information about the secret revealed from the stacked result. The experimental results demonstrate that the proposed scheme does work.

Keywords: Random Grid; Visual Secret Sharing; Weighted Visual Cryptography

1 Introduction

With the technology continually upgrading and improving, people communicate each other on the Internet conveniently. However, the security for the transmission of information through the public channel will be taken into consideration seriously. To this end, traditional cryptography, including DES, AES and RSA [12, 16], has been well-defined to guarantee confidentiality, authentication, integrity, etc. Security of these computational cryptographic tools is determined by the strength of encryption/decryption key′s length. This draws the security to be relative but not absolute.

Visual Secret Sharing (VSS) has drawn much attention in academia, in which the concept of k-out-of-n threshold is presented to encode a secret into n share images by a codebook and, then, recover the secret by recognizing the stacked result of at least k share images ($2 \leq k \leq n$). Compared with traditional encryption, VSS offers unbreakable encryption if less than k meaningless share images are collected. Furthermore, VSS provides the decryption operations without the needs of cryptographic knowledge and computational devices. It is worthwhile to note that VSS is suitable for the applications of high-security needed and without computational devices given.

Within VSS there are three categories: (1) Visual Cryptography (VC) [11]; (2) Probabilistic VC (PVC) [19]; and (3) Random Grid-based (RG) [3, 8]. It is a trade-off between adopting VC or PVC/RG techniques. It is well-known that VC has two main disadvantages: codebook design and pixel expansion [3] while size-invariant PVC/RG has one main concern compared with VC, i.e., lower visual quality of reconstructed secret image. However, the size of pixel expansion seems not always a problem nowadays since of the rapid development of Internet bandwidth and display resolution including HDTV, smart phone, digital camera/recorder, etc. At this stage, VC does not have its own grievances in terms of pixel expansion.

VSS is further extended in various application, such as Progressive VSS [5], General Access Structure (GAS) [1], (k, n)-threshold based VSS [13], Meaningful VSS [4], Multi-secret VSS [6], etc. By the way, Secret Image Sharing (SIS) [15] performs secret sharing like VSS, but computational devices are needed to encode and decode.

A traditional VSS mechanism assumes that each share is treated as the same priority. However, the privilege for participants is not always the same. Hence, the assumption does not reflect the actual situation in real life because everyone stays in different level of job with different duty. For example, the company is composed of the different level of employees. Simultaneously, these employees can be further subdivided into the general manager, the manager and the staff. It is reasonable that a staff of k or more has the same privilege as the manager′s. Likewise, only k or more managers have the same privilege as the general manager′s.

In the literature, there are several VSS or SIS schemes

taking privilege into account. Chen et al. [2] proposed a weighted SIS method, in which the encoding process is divided into two phases. Firstly, sharing shares generated by a (k, n) threshold SIS scheme obtain the same weight. Secondly, these shares are further divided into some groups. The final shares of each group with different privilege are generated respectively. Inspired by the previous SIS method, Lin et al. [10] and Shyu et al. [14] presented their weighted SIS schemes.

Li et al. [9] proposed the (t, s, k, n)-Essential secret image sharing (ESIS) scheme, in which the traditional threshold property is combined with the essential property in the decoding process. ESIS generates n shares containing s essential shares and $(n - s)$ non-essential shares. Within k or more shares including t or more essential shares, the hidden secret can be disclosed computationally; otherwise, no secret is revealed.

Yan et al. [18] further extend Li et al.′s SIS scheme to form (k_0, n_0, k, n)-Essential RG-based scheme. In Yan et al.′s scheme, participants are divided into two group: *essential* and *non-essential*. In the decoding process, the essential share is essential to disclose the secret; otherwise, no secret is revealed without any *essential* share. There are n shares which contain n_0 essential shares and $n - n_0$ non-essential shares. Only k or more shares within containing at least k_0 essential shares can be used to recover the secret. Since Yan et al. adopt a size-invariant VSS scheme, for instance, Wu and Sun′s RG-based VSS scheme [17], the visual quality of disclosed secret is potentially concerned.

In 2015, the concept of privilege is introduced into VC by Hou et al. [7] to benefit from the following advantages: (1) each share with an predefined capability to disclose the secret according to the participant′s privilege; (2) the disclosed secret with a better contrast; and (3) shares with size of the original secret. However, Hou et al.′s scheme does not support the concept and the related property of "essential group".

In order to benefit the superiority of visual quality compared to the RG-based VSS [18] and of supporting the property of "essential group", this paper proposes a weighted visual cryptography scheme such that each participant obtains the share with different weight according to the different-privilege group predefined. It is worthwhile to note the proposed scheme achieves progressive-recovery of secret. That is, the higher value of total weight of stacking shares; the more information about the secret revealed from the stacked result. The experimental results demonstrate that the proposed scheme does work.

The rest of this paper is organized as follows. The related works are given in the next section. The present scheme is described in Section 3. Section 4 demonstrates the experimental results, respectively. Further discussions and conclusions are given in Sections 5 and 6.

2 Related Works

This section gives the brief review of traditional VC and weighted RG-based VSS.

2.1 Visual Cryptography

Visual cryptography (VC), proposed by Naor and Shamir [11] in 1995, encodes a secret image into a number of meaningless shares. The secret can be recognized by the human visual system and disclosed by stacking share upon satisfying recovering threshold condition.

Table 1: The codebook of (2, 2) VC scheme and the stacking results

Secret pixel	□		■	
Share S_1	■□	□■	■□	□■
Share S_2	■□	□■	□■	■□
Stacking result	■□	□■	■	■

Table 1 shows the codebook of a (2, 2) VSS scheme. The dealer uses the codebook to generate two shares, S_1 and S_2. Each secret pixel is corresponding to the defined 1×2 pixels according the codebook. In such a way, the size of a share is twice as big as the secret. When S_1 and S_2 are stacked together, the secret can be disclosed.

2.2 Weighted Visual Secret Sharing Based on Random Grid

Inspired from Li et al.′s SIS scheme [9], Yan et al. [18] propose a (k_0, n_0, k, n)-essential and non-essential RG-VSS scheme, in which the encoding process is divided into two steps.

Step 1. Encoding a binary secret image into $n_0 + 1$ essential shares by $(k_0 + 1, n_0 + 1)$ RG-based VSS, like $SC_1, SC_2, ..., SC_{n_0}, \widetilde{SC_{n_0+1}}$. The share $\widetilde{SC_{n_0+1}}$ is further encoded to generate non-essential shares.

Step 2. Encoding $\widetilde{SC_{n_0+1}}$ into $n - n_0$ non-essential shares by $(k - k_0, n - n_0)$ RG-based VSS, like $SC_{n_0+1}, ..., SC_{n-1}, SC_n$.

Therefore, $\widetilde{SC_{n_0+1}}$ can be restored by stacking some $k - k_0$ or more non-essential shares. The secret image can be disclosed by stacking k or more shares in which containing k_0 or more essential shares.

Note that the VC-based VSS has its superiority of visual quality compared to the RG-based VSS. It benefits to present a VC-based weighted VSS scheme.

3 Proposed Method

The proposed VC-based weighted VSS scheme consists of encoding and decoding phases. An example is given to clear the processes.

Assume there are n participants $P_1, P_2, ..., P_n$ classified into two sub-group: n_0 *specific* participants and $n - n_0$ *general* participants. And at least k_0 out of n_0 specific participants are essentially asked to take part in the secret disclosing operations. Furthermore, $k - k_0$ or more out of $n - n_0$ *general* participants have the same privilege as one *specific* participant. Note that there are at least k participants asked to disclose the secret.

Encoding process is divided into four steps: pre-processing, temporary-share generation, *specific*-share generation, and *general*-share generation.

Step 1. Pre-processing: According predefined k_0 and n_0 (resp. $k - k_0$ and $n - n_0$), design the codebook CB_A of $m \times (n_0 + 1)$ matrix C_A^0 andC_A^1 (resp. CB_B of $m \times (n - n_0)$ matrix C_B^0 and C_B^1) for decoding the white (resp. black) pixels of a secret image and within each row is presented a sharing method in which a participant is assigned. Tables 3 and 4 show the codebook examples of CB_A and CB_B.

Within Table 2 (Resp. 3) there are two groups: a white group and a black group. If a certain pixel of a secret image is white, participants are assigned codewords by randomly selecting one of the cases in the white group. On the contrary, if the secret pixel is black, the codewords are selected from the black group.

Table 2: The codebook CB_A of 4×3 sharing metrics used for (3,3) VC case

C_A^0	C_A^1
$\begin{bmatrix}1&1&0&0\\0&1&1&0\\1&0&1&0\end{bmatrix}_{4\times3}$	$\begin{bmatrix}1&1&0&0\\0&1&1&0\\0&1&0&1\end{bmatrix}_{4\times3}$

Table 3: The codebook CB_B of 4×4 sharing metrics used for (2,4) VC case

C_B^0	C_B^1
$\begin{bmatrix}1&0&0&0\\1&0&0&0\\1&0&0&0\\1&0&0&0\end{bmatrix}_{4\times4}$	$\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}_{4\times4}$

Step 2. Temporary-share generation: A secret image S with size of $w \times h$ is decoded to generate $n_0 + 1$ temporary shares by decoding a secret pixel 0/1 (denoted white/black) according to C_A^0 and C_A^1 by a $(k_0 + 1, n_0 + 1)$ VC scheme. Finally, the $n_0 + 1$ temporary shares of size $4w \times h$ are generated.

Step 3. Specific-share generation: Carry on Step 2. n_0 temporary shares are further expanded into the shares of size $4w \times 4h$ assigned to specific participants. Note that here a pixel is extended into four ones by duplicating a pixel into four.

Step 4. General-share generation: The last temporary share is used to generate $n - n_0$ shares assigned to all general participants. Here, a pixel of the temporary share image is decoding according to C_B^0 and C_B^1 by a $(k - k_0, n - n_0)$ VC scheme. Finally, the $n - n_0$ general shares of size $4w \times 4h$ are generated.

Example 1. *Assume the access structure is defined* $(k_0, n_0, k, n) = (2, 2, 4, 6)$. *There are* $n = 6$ *participants in which two members have the higher privilege and the others have the general privilege. First, a* 512×512 *secret image* S *is encoded into three temporary shares of size* 2048×512 *by (3,3) VC in Table 2. If a secret pixel is white (black), the sharing matrix* $C_A^0(C_A^1)$ *is referred. When the generated two shares of size* 2048×2048 *are assigned to the two participants of higher privilege. Finally, the third temporary share is encoded below. If a pixel of the temporary share is white (black), the sharing matrix* $C_B^0(C_B^1)$ *in Table 3 is referred to generate four shares with size of* 2048×2048 *by (2,4) VC and, then, assign the generated three shares to the general members.*

4 Experimental Results

Carry on the example given in Section 3 and the access structure is defined $(k_0, n_0, k, n) = (2, 2, 4, 6)$. The 512×512 secret image S in Figure 1(a) is encoded into six 2048×2048 shares as $S_1, S_2, S_3, S_4, S_5, S_6$ in Figure 1 (b)-(g), in which S_1, S_2 for specific members with higher privilege, and S_3, S_4, S_5, S_6 for the other general members.

As the access structure is defined $(k_0, n_0, k, n) = (2, 2, 4, 6)$, if stacking any two or three shares, no secret is revealed as shown in Figures 2 and 3.

As the access structure is defined $(k_0, n_0, k, n) = (2, 2, 4, 6)$, if stacking both S_1 and S_2 with at least two shares out from S_3, S_4, S_5,and S_6, the secret can be disclosed as shown in Figure 4(a)-(f) and Figure 5(a)-(d) and (g). Otherwise, the stacked results are noise-like as shown in Figure 4(g)-(o) and Figure (e)-(f).

The experimental results presented above demonstrate that the proposed scheme does work and the secret can be disclosed and recognized by the human visual system under the predefined privilege policy.

5 Discussions

The proposed scheme presents a new VC-based VSS scheme with the following properties.

1) **Privilege:** The participants are classified into different groups. The high-privilege members are asked to essentially participate to disclose the secret.

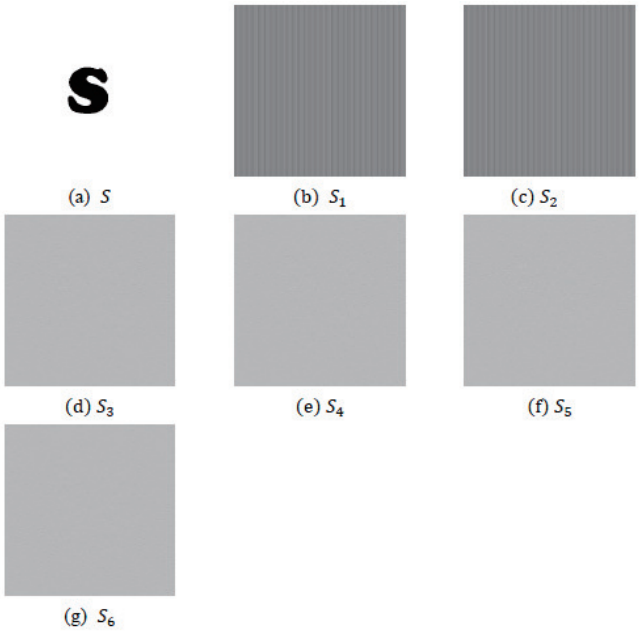


Figure 1: (a) secret image S; (b)-(c) special shares with higher weight; and (d)-(g) general shares

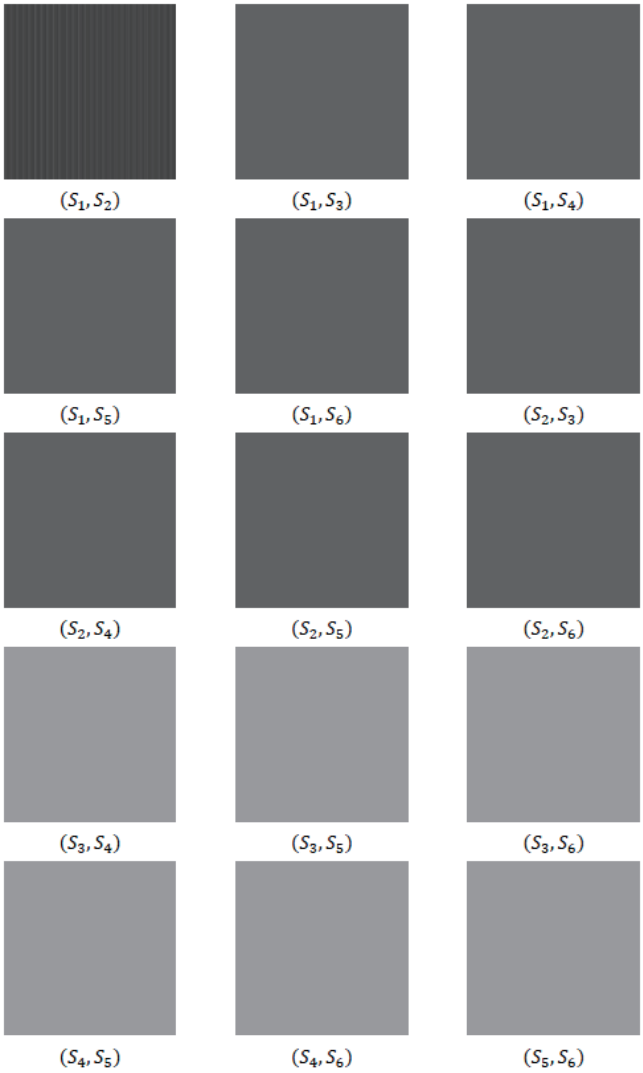


Figure 2: The fifteen stacked results of any two from six shares

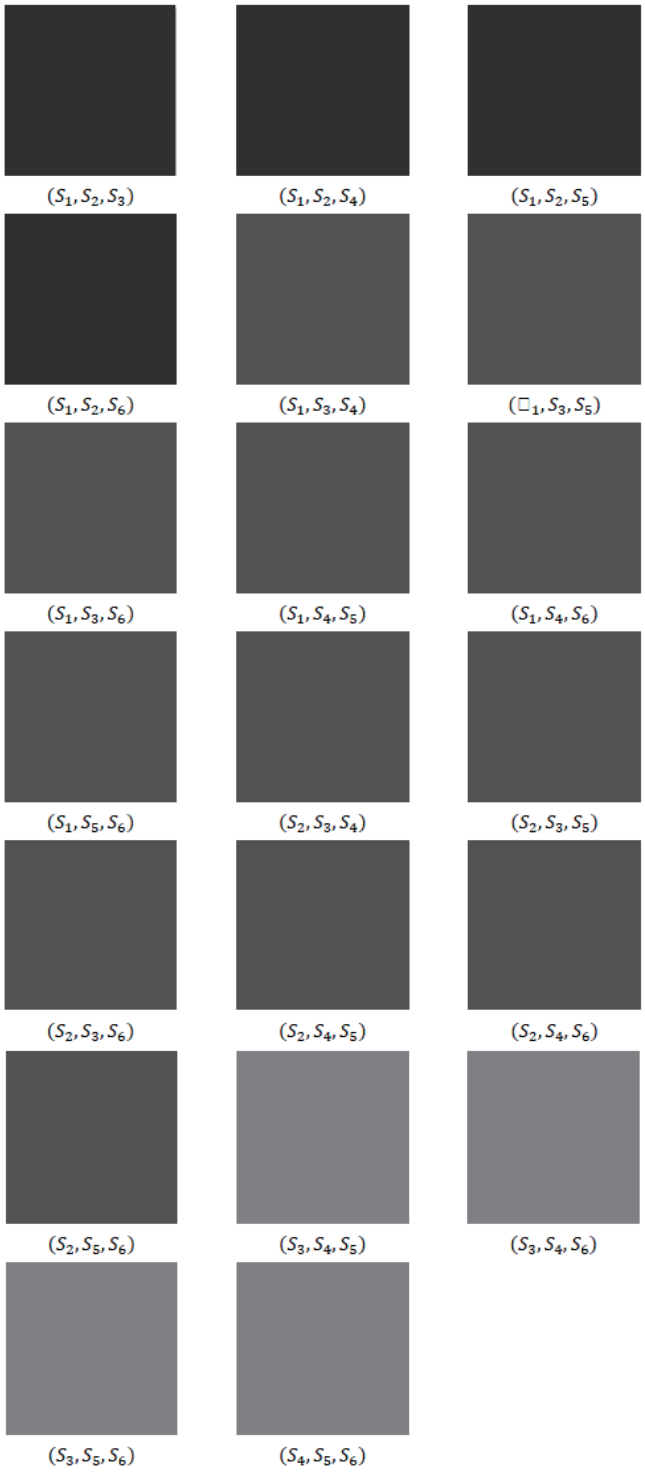


Figure 3: The twenty stacked results of stacking any three shares

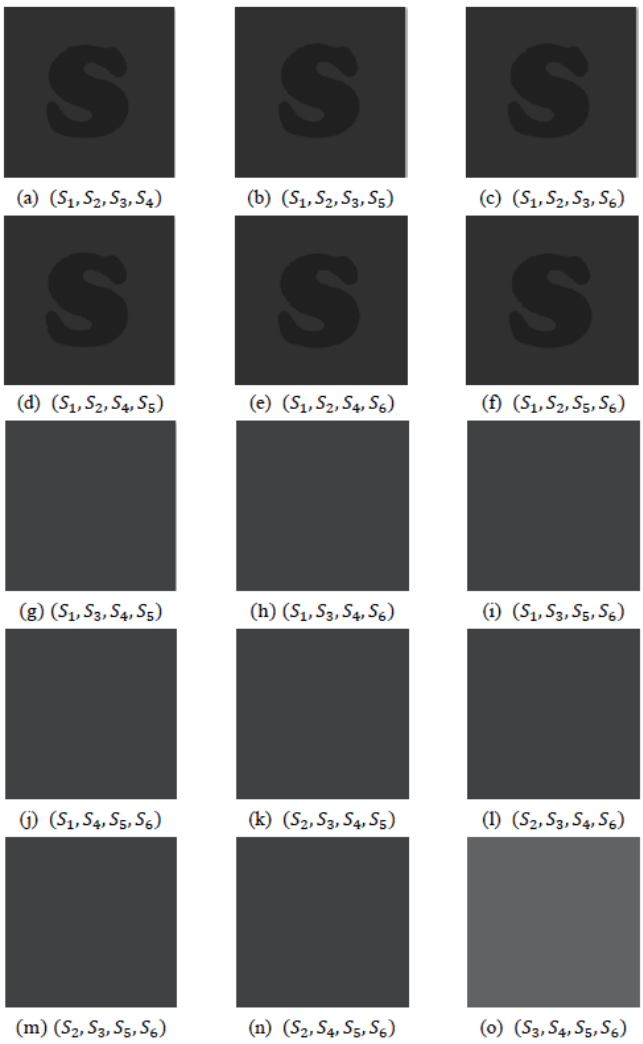


Figure 4: (a)-(f) the stacked results including two specific shares and two general ones; (g)-(o) the stacked noise-like results

Table 4: The contrast values in the experimental results

Stacked results	Contrast
(S_1, S_2, S_3, S_4)	0.0553059
(S_1, S_2, S_3, S_5)	0.0552645
(S_1, S_2, S_3, S_6)	0.0553177
(S_1, S_2, S_4, S_5)	0.055266
(S_1, S_2, S_4, S_6)	0.0553192
(S_1, S_2, S_5, S_6)	0.0552778
$(S_1, S_2, S_3, S_4, S_5)$	0.117063
$(S_1, S_2, S_3, S_4, S_6)$	0.117122
$(S_1, S_2, S_3, S_5, S_6)$	0.117076
$(S_1, S_2, S_4, S_5, S_6)$	0.117078
$(S_1, S_2, S_3, S_4, S_5, S_6)$	0.186576

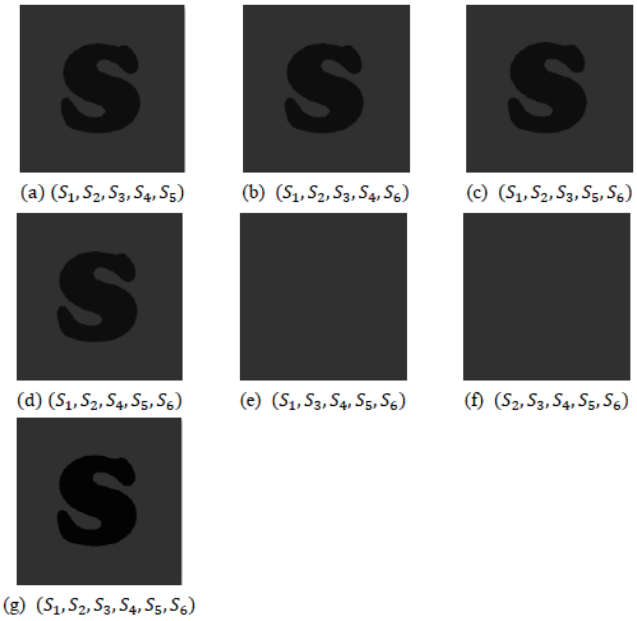


Figure 5: (a)-(d) the stacked results including two specific shares and three general ones; (e)-(f) the stacked noise-like results; and (g) the one by stacking all shares

2) **Visual quality:** The disclosed secret can be clearly recognized in the experiments. Furthermore, the contrast is defined as $\frac{H(V_1)-H(V_0)}{m}$, where m is the pixel expansion rate, $H(V_0)$ and $H(V_1)$ are the Hamming weights of all white and all black area in the disclosed image corresponding to the white and black areas in the original image. Table 4 demonstrates the contrast values of all the cases in which the secret can be disclosed in the experimental results.

3) **Progressive:** It is obvious by Table 4 that the secret in the proposed scheme can be progressively disclosed.

4) **Security:** Any information of the secret cannot be revealed without satisfying the access structure. Here, less than k shares or less than $k_0 specific$ shares cannot be stacked to reveal the secret.

Table 5 gives the comparison between the related works and the proposed.

6 Conclusions

In order to enable the practical end of assigning shares with the specific privilege, this paper proposes a new weighted VC scheme. Each participant in a different group with distinct privilege is delivered a share with specific weight. The disclosed secret is revealed only under the predefined policy of privilege. The experimental results demonstrate the proposed scheme works well.

Table 5: Comparisons of the weighted VSS schemes

Schemes	VC	RG	PVC	SIS	Pixel expansion	Visual quality	Essential
Chen et al. [2]				✓	N/A	N/A	No
Lin et al. [10]				✓	N/A	N/A	No
Shyu et al. [14]				✓	N/A	N/A	No
Li et al. [9]				✓	N/A	N/A	Yes
Yan et al. [18]		✓	✓		No	Low	Yes
Hou et al. [7]	✓				No	Low	No
The proposed	✓				Yes	High	Yes

Acknowledgments

This research was partially supported by Ministry of Science and Technology, R.O.C., under contrast No. MOST 103-2221-E-415 -017.

References

[1] G. Ateniese, C. Blundo, A. De Santis and D. R. Stinson, "Visual cryptography for general access structures," *Information and Computation*, vol. 129, pp. 86–106, 1996.

[2] C. C. Chen, C. C. Chen and Y. C. Lin, "Weighted modulated secret image sharing method," *Journal of Electronic Imaging*, vol. 18, pp. 043011-1–043011-6, Oct. 2009.

[3] T. H. Chen, and K. H. Tsao, "User-friendly random-grid-based visual secret sharing," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 21, pp. 1693–1703, 2011.

[4] T. H. Chen, K. H. Tsao, and Y. T. Yang, "Friendly color visual secret sharing by random grids," *Fundamenta Informaticae*, vol. 96, pp.61–70, 2009.

[5] W. P. Fang, and J. C. Lin, "Progressive viewing and sharing of sensitive images," *Pattern Recognition and Image Analysis*, vol. 16, pp.632–636, 2006.

[6] J. B. Feng, H. C. Wu, C. S. Tsai, Y. F. Chang, and Y. P. Chu, "Visual secret sharing for multiple secrets," *Pattern Recognition*, vol. 41, pp. 3572–3581, 2008.

[7] Y. C. Hou, Z. Y. Quan, and C. F. Tsai, "A privilege-based visual secret sharing model," *Journal of Visual Communication and Image Representation*, vol.33, pp. 358–367, 2015.

[8] O. Kafri and E. Keren, "Encryption of pictures and shapes by random grids," *Optics Letters*, vol. 12, pp. 377–379, 1987.

[9] P. Li, C. N. Yang, C. C. Wu, Q. Kong and Y. Ma, "Essential secret image sharing scheme with different importance of shadows," *Journal of Visual Communication and Image Representation*, vol. 24, no. 7, pp. 1106–1114, Oct. 2013.

[10] S. J. Lin, L. S. T. Chen and J. C. Lin, "Fast-weighted secret image sharing," *Optical Engineering*, vol. 48, pp. 077008-1–077008-7, July 2009.

[11] M. Naor and A. Shamir, "Visual cryptography," in *Proceedings of Advances in Cryptography (Eurocrypt'94)*, Lecture Notes in Computer Science, vol. 950, pp. 1–12, 1995.

[12] B. Schneier, *Applied Cryptography*, 2nd ed, 1996.

[13] S. J. Shyu, and M. C. Chen, "Optimum pixel expansions for threshold visual secret sharing schemes," *IEEE Transactions on Information Forensics and Security*, vol. 6, pp. 960–969, 2011.

[14] S. J. Shyu, C. C. Chuang, Y. R. Chen and A. F. Lai, "Weighted threshold secret image sharing," in *Proceedings of The Third Pacific-Rim Symposium on Image and Video Technology*, LNCS 5414, Springer, Jan. 2009.

[15] C. C. Thien, and J. C. Lin, "Secret image sharing," *Computers and Graphics*, vol. 26, pp. 765–770, 2002.

[16] G. Tuychiev, "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.

[17] X. Wu, and W. Sun, "Improving the visual quality of random grid-based visual secret sharing," *Signal Processing*, vol. 93, pp. 977–995, 2013.

[18] X. Yan, S. Wang, X. Niu and C. N. Yang, "Essential visual cryptographic scheme with different importance of shares," in *Proceedings of the* 21^{st} *International Conference on Neural Information Processing*, LNCS 8836, pp. 636–643, Springer, Nov. 2014.

[19] C. N. Yang, "New visual secret sharing schemes using probabilistic method," *Pattern Recognition Letters*, vol. 25, pp. 481–494, 2004.

Biography

I-Chun Weng received his M.S. in Department of Computer Science and Information Engineering from National Chiayi University in 2016. Her research interest is visual cryptography and information security.

Tzung-Her Chen was born in Tainan, Taiwan, Republic of China, in 1967. He received the B.S. degree in Department of Information & Computer Education from National Taiwan Normal University in 1991 and the M.S. degree in Department of Information Engineering from Feng

Chia University in 2001. In 2005, he received his Ph.D. degree in Department of Computer Science from National Chung Hsing University. He has been with Department of Computer Science and Information Engineering at National Chiayi University as Professor since August 2011. His research interests include information hiding, multimedia security, digital rights management, network security. He is an honorary member of the Phi Tau Phi Scholastic Honor Society.

DDoS Attack Detection Using Unique Source IP Deviation

Ram Charan Baishya, Nazrul Hoque, and Dhruba Kumar Bhattacharyya
(Corresponding author: Dhruba Kumar Bhattacharyya)

Department of Computer Science and Engineering, Tezpur University
Tezpur University, Napaam, Sonitpur, Assam-784028, India
(Email: dkb@tezu.ernet.in)
(Received Apr. 19, 2016; revised and accepted July 19 & Aug. 29, 2016)

Abstract

In this paper we present a low cost yet robust DDoS detection method to identify all classes of DDoS attacks. Our method attempts to detect DDoS attack by monitoring the deviation of the count of unique source IPs and the count of source IPs whose transmission rate is higher than a given threshold value. Unlike other similar existing methods, our method does not need to maintain a list of source IPs which makes our detection method faster. Another advantage of our method is the ability to detect attack performed by small size bot net. In case of such an attack the packet rate of the attack sources deviate from its mean value significantly and thus we can detect this change. We use a non-parametric change point modeling technique to identify flooding attacks of all types in real time. An other contribution of this work is the development of an attack took referred to as TU-CANNON, to generate different variations of DDoS attack under a controlled test-bed environment.

Keywords: DDoS; DDoS Attack Detection; Low-Rate DDoS Attack; DDoS Attack Tool

1 Introduction

In a distributed denial of service (DDoS) attack, one or more group of compromised users send legitimate traffic to the victim to degrade or even shut down the service of the victim. The goal of such a DDoS attack typically varies from creating simple inconvenience to the user of a website, to incur major financial losses to the on-line service providers. Several such incidents can be found in history [10, 11], which show the great threat of DDoS attack to the Internet service providers as well as the Internet users. Today, our day to day activities are becoming more and more dependent on the Internet, starting from e-shopping to e-banking. Most emergency services are also dependent on the Internet. Hence, the impact of a DDoS attack is becoming more threatening. The availability of large no of DDoS attack tools in the public domain has made it very easy to launch such attack with various levels of intensities, even for unskilled users with malicious intention [2, 21]. Such tools are well equipped with automatic scanning of the vulnerable machines over the Internet, exploitation and deployment of attack code in those machines and then performing the actual attack. A detailed description of DDoS attacks and their classifications can be found in [28].

The design goal of TCP/IP was to deliver packets from source to destination in a fast and accurate way. The payload of the packets are of little concern to TCP/IP. A DDoS attack generates a huge volume of TCP/IP packets from a large number of sources. These attack packets are generally indistinguishable from that of normal traffic packets. Thus when all these attack packets merge at the victims site, they occupy most of the victim's network bandwidth and forces the victim to degrade its service or at worst shut down its services temporarily. Typically a DDoS attack is launched from different sources in a coordinated manner and the attack traffic from individual source is comparatively low which make it difficult to distinguish from normal traffic.

Most defense solutions [13, 24, 25] to DDoS attack detection have been found less effective due to the distributive nature of the attack.

Thus we consider two key challenges for DDoS attack defenders. First, how to detect the attack as close as possible to the sources, so that the traffic from such source can be blocked early. Second, to detect the DDoS attack in the victim site as early as possible so that the victim gets enough time to take appropriate action to such an attack.

One obvious way taken by most researchers [5, 15, 27] is to monitor the volume of traffic that are received by the victim site.

However, such methods are not robust against the bursty nature of Internet traffic. In case of the bursty nature of internet traffic such methods often identify it as an attack. On the other hand, such bursty traffic may ac-

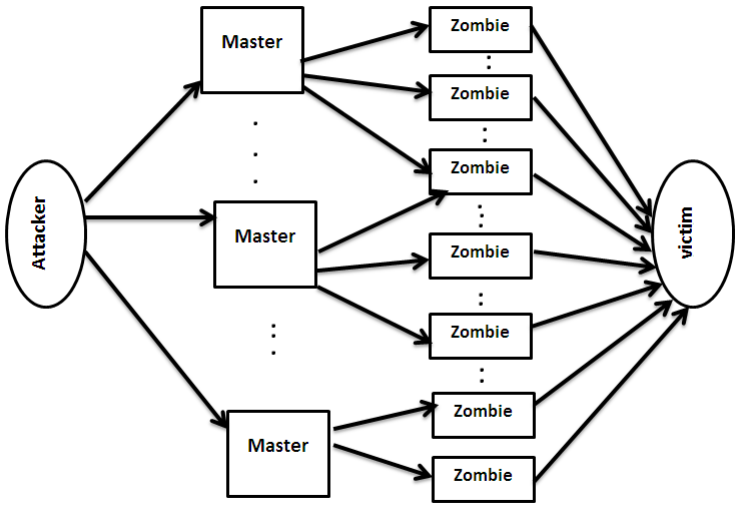


Figure 1: The main elements of a DDoS attack

tually be attack traffic, and a delayed decision may turn out to be very risky for the victim. Since a DDoS attack is highly distributed, the number of source IPs involved in the attack is much larger than the number of source IPs under normal condition. Peng et. al. [22] used the arrival rate of new source IP addresses in the traffic during each observation period. However, this approach needs to maintain a database of trustworthy source IP addresses, which might itself be vulnerable to attack. In this paper, we propose a mechanism that monitors the number of unique source IP addresses during each observation period. Our assumption is that during an attack, the number of source IP will increase abruptly, and by detecting this change we can detect the attack. Also, if the attacker attempts to launch a DDoS attack with less number of sources, the rate of transmission from each source must be high to achieve a bandwidth attack. We monitor the count of sources which transmits above a threshold. Under a less distributive attack the mean value of this count deviates significantly, which indicates the presence of an attack.

Our contribution in this paper is a simple and fast approach to detect DDoS attacks by monitoring the deviation of (i)the count of the unique source IPs from its mean value and (b)the count of the sources transmitting at a high rate. We also introduce an effective tool to launch DDoS attacks, called (TUCANNON) of all types. We establish that our method shows significant detection performance for both benchmark DDoS dataset as well as for our own datasets than most of the previous schemes. To detect changes in our observed features we adopted the non-parametric CUSUM approach and applied it by following the idea of wang et al. [26].

The rest of the paper is organized as follows. Section 2 gives an overview of distributed denial of service attacks. Section 3 mentions some of the related work done and in Section 4 we provide our attack tool called TUCANNON. In Section 5 we present our detection algorithm along with the necessary theory. Section 6 presents the results of performing our algorithm on various network traces. Finally, conclusions are drawn in Section 7.

2 Background of DDoS Attack

The goal of a DoS attack is to overwhelm the victim by occupying its resources like network link,different queues, processing unit etc. As a consequence, the victim's performance degrades, and the legitimate user of the service observes a denial of service from the victim side. Typically, an attacker sends a few malformed packets to the victim. These packets are crafted in a way which is unexpected for the application or a protocol on the victim machine. Processing of such packets forces the victim application or protocol to freeze or to restart. However, in this paper we study the characteristics of flooding attacks. In such an attack the attacker sends a large volume of legitimate traffic to the victim site. As a result, the victims network link might go into congestion. Also the communication protocol may allocate resources to such attack traffic and might run out of resources. In case of a DDoS attack, the attacker generates a high volume of traffic from different sources in a coordinated manner. To perform a DDoS attack, the attacker uses different protocols like Internet Control Message Protocol (ICMP), User Datagram Protocol (UDP)and Transmission Control Protocol (TCP).

The main elements of a DDOS attack are the attacker, the victim and the intermediate network of machines as shown in Figure 1.

Here the attacker is the real source of the attack. The victim could be a single machine or an entire subnetwork. The intermediate network composed of many compromised machines. Generally, the intermediate network

contains masters and zombies. The masters are compromised machines, and are loaded with programs to control multiple attacking machines, called as zombies. The zombies are the compromised machines which send attack traffic to the victim in response to the command received from their masters.

The following steps take place while preparing and conducting a DDoS attack:

1) Discovery and Compromise: In this step the attacker discovers vulnerable machines over the Internet to get access to them. Later on, such systems are loaded with programs which can further detect and compromise other vulnerable machines over the net. Different scanning tools available to discover vulnerable machines such as Nmap, self-propagating tools like the Ramen worm [19] and Code Red [18] are used to discover and compromise victim machines.

2) Communication: Once machines are compromised, the attacker communicates with the masters on various occasions to know what are the agents which are on line, starting time of the attack, the rate of attack traffic, the end time of the attack etc. The agents are configured to communicate with one or more masters. Protocols like TCP, UDP and ICMP are used for these communications. Another mode of communication is by making use of the IRC servers.

3) Actual Attack: In this phase the attacker sends command to the masters which inturn activates the online zombies to carry out the actual attack.

   A detailed description of different types of DDoS attacks and along with different tools to perform such attacks can be found in [28].

3 Related Work

DDoS attack is different from a DoS attack in that a DoS attack can be characterized by high volume of traffic from a single source, however in case of a DDoS attack the volume of traffic seen by the victim is actually generated from a large number of sources over the Internet. Hence the traffic from each source is very dilute. Even if the number of sources are small, use of IP spoofing can create the same effect. Hence detecting the attack based on the source IP seems to be a better option over traffic volume based detection mechanisms [13, 24, 25]. Peng et al. [22] develop a source IP based detection mechanism, SIM (Source Addres Monitoring). It maintains a database of known source IP addresses and updates this database periodically. For each observation period they calculate the percentage of new source IP addresses in the traffic and used non-parametric CUSUM algorithm to detect abrupt change in the percentage of new IP addresses in the network. The working of this method is dependent on the content of the IP database. The attacker can send traffic in a legitimate way and thus get itself an entry in the IP database, which might turnout to be a loophole in the detection mechanism. Also they used a threshold based mechanism to detect high speed sources as attacking source, which may not be the case always. In normal traffic burst from single source can be seen quite often.

In [1] the authors use rank correlation measures to discriminate DDoS attack traffic from legitimate traffic. In [20] the authors propose a DDoS mitigation technique based on correlation pattern suitable for cloud computing environment. In [14] authors propose an entropy based DDoS mitigation technique, which is capable of discriminating flash crowd in VOIP network. In [4] authors discuss a method to prevent DDoS attack which uses IP spoofing to perform the attack.

In this paper we present an effective detection scheme called Violating Source IP Count (VSC) which is based on the observation that under normal condirion the number of unique source IPs over a fixed observation period remains stable. At each interval VSC checks whether the count significantly deviates from its mean value or not. Our detection schema detects such significant change to identify the presence of an attack. Our detection schema also keeps track of the count of such high speed sources. If the attacker uses a small size bot to escape the unique IP count deviation system,the sources has to transmit at high speed. Hence a deviation can be seen in the mean value of the count of high speed sender under an attack. VSC detects such a change and confirms an attack. To evaluate our detection mechanism we perform several experiments on different network traces and are presented in section 5. The evaluation results indicate that VSC has a short detection time and a high accuracy rate. Also since the complexity of this method is very low both in terms of space and time, this method can easily be deployed in a distributed manner in the first mile and intermediate routers to detect the attack in the beginning stage itself.

4 DDoS Attack Generation Tool TUCANNON

Based on the protocol we can classify the basic attack types as follows.

1) TCP Flood: A stream of packets with various flags (SYN, RST, ACK) are sent to the victim machine. The TCP SYN flood works by exhausting the TCP connection queue of the host and thus denying legitimate connection requests. TCP ACK floods can cause disruption at the nodes corresponding to the host addresses of the floods as well. TFN [12] is a popular DDoS tool for this type of attack.

2) ICMP Flood (e.g ping floods): A stream of ICMP packets is sent to the victim host. A variant of the ICMP floods is the Smurf attack in which a spoofed IP packet consisting of an ICMP ECHO_REQUEST is sent to a directed broadcast address. TFN [12] is a popular DDoS tool for this type of attack.

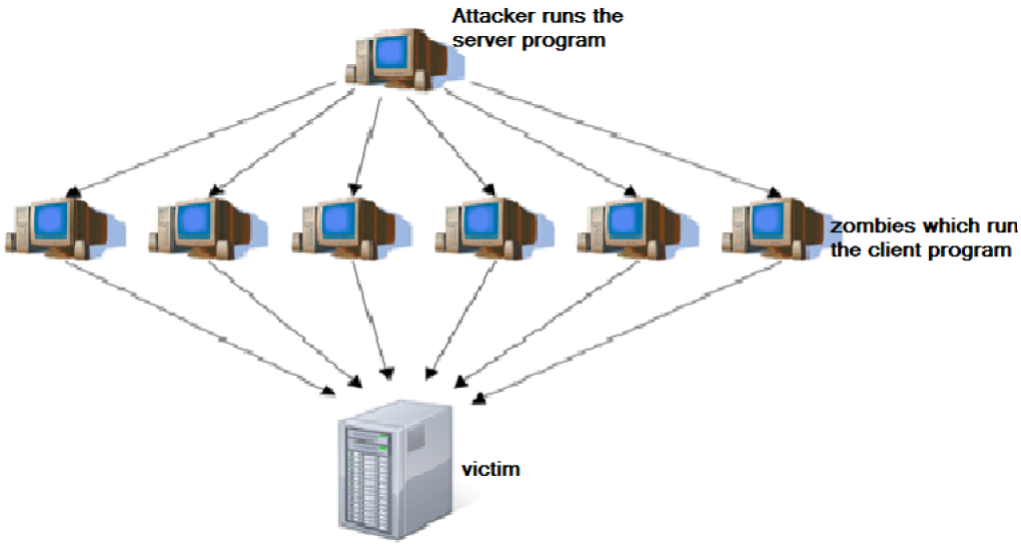


Figure 2: TUCANNON based direct attack strategy

3) UDP Flood: A huge amount of UDP packets are sent to the victim host. Trinoo [12] is a popular DDoS tool that uses UDP floods as one of its attack payloads.

The DDoS attack tools like trinoo, tfn, tfn2k [12] along with a lots of other tools are easily available on the internet. However these tools are not suitable for generating a coordinated DDoS attack in a testbed environment. Hence, as part of our research we developed a DDoS attack generation tool referred as TUCANNON, which can generate all the above mentioned DDoS attacks, also provide a lots of flexibility to adjust the pattern of the attack traffic like constant rate attack, increasing rate attack, pulsing attack and subgroup attack as mentioned in [16].

A pictorial description of our attack tool is shown in Figure 2. The tool comprises of two programs.

1) The first program is used by the attacker to communicate with the bots. This program uses GUI so that the attacker can easily specify various parameters like protocol type, attack pattern type etc. In this paper we will refer to this program as server program.

2) The other program is executed in each bot. This program is responsible for accepting command from server program and launch the attack accordingly. In this paper we will refer to this program as the *client program.*

4.1 Server Program

Using this program we communicate with the machines which are configured as bots in the test-bed. This program is developed with a user interface through which one can easily specify and control different properties of the attack traffic. Such properties are the protocol type (TCP, UDP and ICMP), the attack pattern (constant rate attack, increasing rate attack and pulsing attack) and the type of source IP (actual IP of the machine or randomly generate valid but spoofed IP address), no of threads (where each thread executes one copy of the slave program inside a single bot machine) and range of ports of the victim to send the traffic. When the master starts, it waits for slaves to connect to it. Figure 3 is a snapshot of the GUI of the server program.

The following is a brief description of various components of the interface:

List of Zombies: When the attacker starts the server program, it waits for the client programs to connect to it. As soon as a client program connects to the server, the clients IP address is shown in the left side panel of the interface as shown in Figure 3.

Protocol type: To launch an attack, the attacker has to select the type of protocol by selecting any one of the corresponding radio button.

Source IP Configuration: These options are used to specify whether the attack packet carries the actual source IP or a spoofed one. Also in case of spoof source IP, the attacker can specify the number of different unique spoofed IPs used in the attack. This option allows the attacker to spread the required attack traffic over a specified number of source IP.

No of Threads: The number of machines in our test bed is very limited (around 50). Hence to increase the amount of traffic each client program sends traffic by using multiple threads. The number of threads used by each client can be specified by the attacker through this input. This feature is used by the attacker to control the traffic rate in the attack.

Victim IP: This input field is used by the attacker to specify the IP address of the victim machine.

Low Port and High Port: The attacker can specify

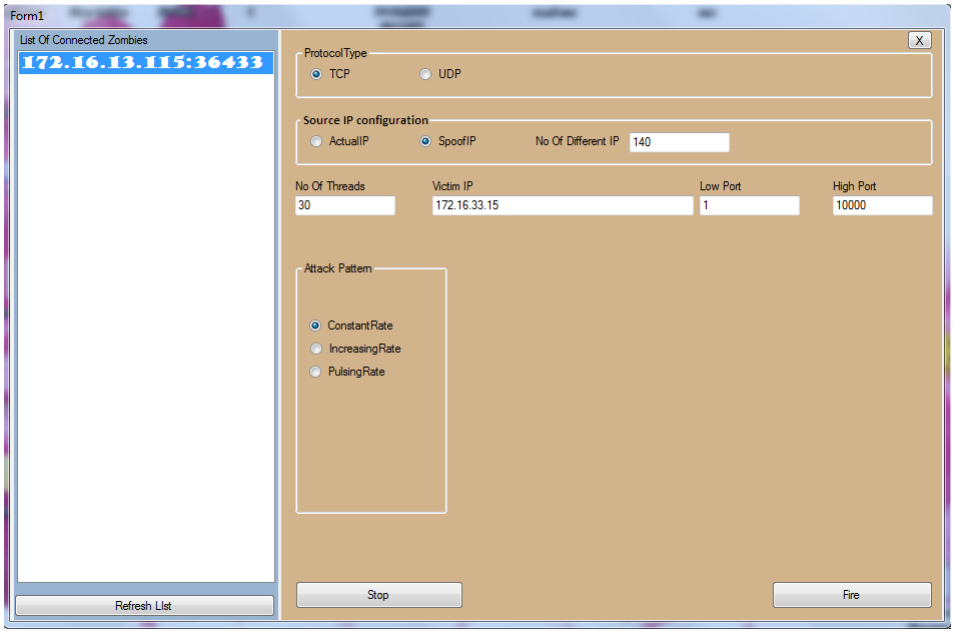


Figure 3: GUI of TUCANNON server program

the range of ports where to send the traffic via these input.

Attack Pattern: As mentioned earlier there can be four different traffic pattern. The Attacker can select the pattern from this list.

Fire: when the attacker clicks on this button, attack command along with the specified input is sent to all clients currently connected to the server.

Stop: The attacker can stop the attack by clicking on this button.

4.2 Client Program

This program is responsible for actually sending the attack traffic as specified by the command sent from the master. When the client program starts it connects to the server whose IP is specified as input to the client program. After connecting to the server it waits for command from the server.

5 Our Solution: Violating Source IP Count (VSC)

For a DDOS attack, the attacker's main goal is to overwhelm the server by sending illegitimate network traffic using different protocols.

One common characteristic of DDoS attack is that the volume of the traffic during an attack is very high. To generate high volume of traffic the attacker either has to use a large botnet consisting of lot of compromised machines or the attacker may send traffic from a small botnet but at very high speed. We are making the assumption that under normal condition the deviation of the number of unique source IP addresses from its mean value is bounded by an upper bound. And also during normal condition the deviation of the number of source IP addresses sending traffic above a threshold from its mean value is also bounded by an upper bound. In this paper we refer to such IP addresses as violating IP.

Based on these two assumptions we present a detection mechanism called violating Source IP count (VSC) to detect a distributed denial of service (DDoS).

The key features of VSC are highlighted below.

1) Researchers have already used source IP addresses as detection feature such as in Peng et al. [22]. However, we use the count of unique source IP addresses, in an observation period, as detection feature. This approach does not require to maintain a database of trustworthy IP addresses for its operation. Thus the memory requirement and speed of this algorithm is comparatively better, which is a key goal for a detection system. This feature makes VSC very suitable to be used in a distributed manner.

2) A DDoS attack may either use a small size bot sending traffic at a high speed or a large size bot consisting of many zombies. VSC monitors changes in both the number of source IP addresses and count of sources transmitting at high speed. Thus VSC traps the attacker from both the directions, hence reducing the scope of the attacker.

5.1 Overview of Violating Source IP Count

VSC attempts to detect the presence of attack by monitoring two features of the traffic, namely the number of

unique source IP addresses and the number of violating source IP addresses. VSC collects the incoming packets during every observation period, say δT and inserts the packets into a binary search tree based on their source IP address. Each node in the tree has a count field that specifies the number of packets from the source IP represented by the source IP field of the node, as illustrated in Figure 4

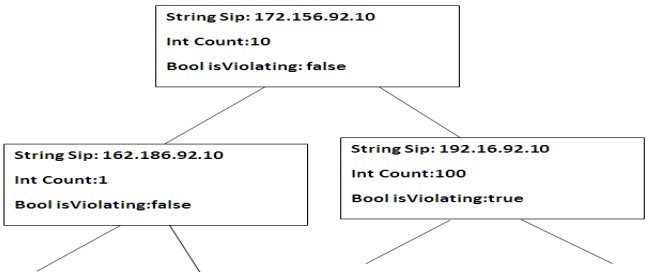


Figure 4: The binary search tree used in the detection engine

If the number of packets for a source IP address is greater than a certain threshold, that IP address is marked as violating IP. This information is used by VSC to detect DDoS attack that uses a small size botnet to carry out the attack. Also, at the end of each observation period the number of nodes in the binary search tree gives the number of unique source IP addresses during the observation period. Thus at the end of each observation period the BST (Binary Search Tree) gives us a) the number of violating source IP addresses V_i, and b) the number of unique source IP addresses X_i in the current observation period t_i.

5.1.1 System Architecture

Figure 5 provides an overview of VSC mechanism. The VSC mechanism consists of three basic components, viz, detection engine, decision engine, and response engine.

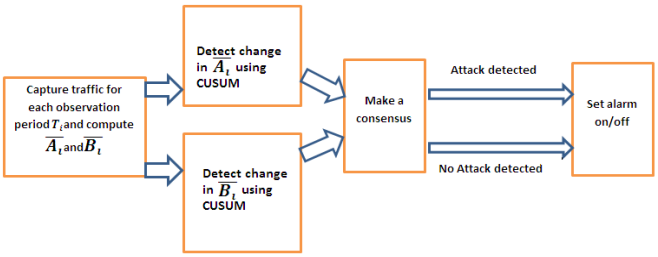


Figure 5: Architecture of VSC

The detection engine processes the incoming traffic to detect any attack. The task of the decision engine is to combine the results from the detection engine and to reach a consensus about the occurrence of an attack. The response engine in turn sets an alarm on or off based on the output of the decision engine.

5.1.2 Placement of the Detection Mechanism

The VSC can be deployed in different locations in a network as shown in Figure 6.

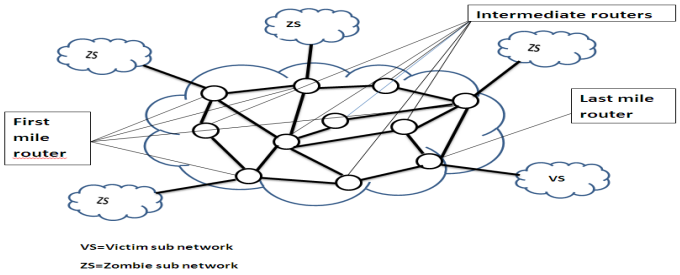


Figure 6: Different possible placement location of VSC

If it is deployed in the first mile router, the detection rate will depend on the size of the bot. If a small bot is used then the number of packets from one IP must be high, which our detection engine can detect by estimating the change in the count of violating source IP count. However, in case of a large bot, the detection rate might decline.

If deployed in intermediate routers, the detection rate increases as it moves towards the victim site. One can thus deploy VSC at different intermediate routers and detect the occurrence of an attack in a distributed manner. Chen et al. [9] describe such an approach in their work.

Another point of deployment is at the victim site, i.e, the last mile outer. Since all the attack traffic aggregate in the last mile router, deviation of $\overline{A}_n$ and/or $\overline{B}_n$ can easily be detected, if there is any.

Among these different detection points the intermediate routers and the last mile router carries the greatest interest from the point of view of the victim. In this paper we present the experimental results performed on the last mile router. Detection of the attack by placing VSC at different intermediate routers in a distributed manner is out of thee scope of this paper.

5.2 Theory Behind VSC

As mentioned above, VSC monitors the number of unique source IP address X_i and number of violating source IP addresses V_i in each observation period to detect the occurrence of an attack in the network. Under the normal condition, the deviation of X_i and V_i from its mean value is less, however under an attack these parameters deviate from their mean largely. Our detection engine thus monitors and detects (if any) such a significant change in these two parameters and confirms as an attack based on some threshold value.

The following section describes the approach we use to detect such change in the above mentioned framework.

5.2.1 The Non-Parametric CUSUM Algorithm:

Let X_n, *where* $n = 0, 1, 2, 3..$ and V_n, *where* $n = 0, 1, 2, 3..$ be the number of unique source IP addresses and the num-

ber of violating source IP address in an observation period t_n. Since X_n and V_n are highly dependent on different attributes of the network (such as size, time of the day, etc) from which they are collected, we first normalize X_n and V_n by the average value of X_n and V_n respectively. Let $\overline{X}_n$ and $\overline{V}_n$ represent the mean value of X_n and V_n. Then $\overline{X}_n$ and $\overline{V}_n$ can be computed as follows $\overline{X}_n$=α*$\overline{X}_n$ − 1+(1-α)*X_n, $\overline{V}_n$=α* $\overline{V}_n$−1+(1-α)*V_n Where α is the memory factor and lies between 0 and 1.

Thus from X_n and V_n we define A_n=$X_n/\overline{X}_n$ B_n=$V_n/\overline{V}_n$ Since A_n and B_n are normalized values, no longer they are dependent on the current network characteristics.

We use the concept of sequential change point detection [10] in our detection algorithm. The goal of the change point detection mechanism is to detect the presence of a change of the mean value in a observed time series data. In our algorithm, it detects change in A_n and B_n. However,accurate estimation of A_n and B_n are challenging task, hence we use non-parametric CUSUM method [24] in our detection algorithm. Non-parametric CUSUM is not model specific and hence suitable for our purpose. The basic idea of using non-parametric CUSUM to detect abrupt change in a time series data is based on the model presented in Wang et al. [10]. The details of non-parametric CUSUM can be found in [24]. Here we demonstrate how to apply non-parametric CUSUM on A_n to detect change. Similar approach is taken in case of B_n.

Under normal condition, the mean of A_n denoted by c ($i.e,\ c = E(A_n)$) is near by 1. We chose a parameter which is the upper bound of c. From A_n we derive another random sequence $\overline{A}_n$ such that $\overline{A}_n = A_n - c$. This transformation will make the mean value of $\overline{A}_n$ negative under normal condition, which is a basic assumption of non-parametric CUSUM algorithm [6]. Now consider h as the lower bound of the amplitude of increase in the mean value of $\overline{A}_n$ during an attack and $h \gg c$. As presented in wang et al. [26] the nonparametric CUSUM algorithm can be written as

$$Y_n = S_n - \min\ S_k, 1 \leq k \leq n.$$

Where $S_k = \sum_{i=1}^{k} \overline{A}_i$, $S_0 = 0$ at the beginning and Y_n is the accumulated positive values of $\overline{A}_i$. Thus if Y_n is very large it is a clear indication of the deviation of the observed value of the random sequence from its mean value.We use a threshold value N which is compared against Y_n at the end of each observation period. If Y_n exceeds N an attack is detected. Thus we can now formally define the detection function as

$$D_N(Y_n) = \begin{cases} 0 & \text{if } Y_n \leq N \\ 1 & \text{otherwise.} \end{cases}$$

Where Íindicates an attack and 0́detects normal traffic.

5.2.2 Parameter Specification

Two key measures of greatest interest for a DDoS attack detection system are given below.

1) False alarm rate, i.e, the number of normal instances reported as attack over a specific period of time.

2) Detection time, i.e, time duration between the starting of an attack and the detection of the attack.

However, both these design goals are mutually conflicting, as expected! To achieve one other one often has to compromise up to extent. in practice (1,4,30 p) CUSUM is considered as optimal in terms of both false alarm rate and detection time. As presented in Brodsky et al. [6]

$$\begin{aligned} \tau_N &= \inf n : D_n(.) = 1 \\ \rho_N &= \frac{(\tau_N - m)^+}{N}, \end{aligned} \tag{1}$$

where τ_N = detection time; ρ_N = normalized detection time after a change occurs. Inf represents $infimum$. n is the time when the attack started.

ρ_N and h can be related by the following equation

$$\rho_N \to \gamma = \frac{1}{h - |c - a|} \tag{2}$$

where $h - |c - a|$ gives the mean of $\overline{A_n}$, after an attack begins. As mentioned in [26] the above equation gives an upper bound of the actual detection time. Thus to achieve our design goals we have to choose optimal values for the parameters a and N. It is clear from Equation (1) and Equation (2) that once we are given a,h and *a detection time period* we can calculate N accordingly.

The parameter a is used to offset A_n to be $\overline{A_n}$, so that $\overline{A_n}$ has a negative mean under normal condition. If a is chosen to be very high the likelihood of getting positive values in the sequence $\overline{A_n}$ is less. In turn the accumulated value i.e, Y_n might not reach the threshold.

The parameter N specifies the threshold for Y_n. If N is chosen to be very high, false alarm rate will be low at a cost of high detection time. On the other hand, a small value of N may increase the false alarm rate.

As mentioned earlier CUSUM algorithm needs a, h and *detection time interval* to be specified and calculates N by using Equations (1) and (2).

Here a is the upper bound estimation of the mean of $\overline{A_n}$. From the definition of $\overline{A_n}$ can safely assume a as 1.1.

The parameter h specifies the amplitude of the minimum increase of the mean value of $\overline{A_n}$ under an attack. By following the same principle as in wang et al. [26] we set $h = 2 * a$.

We used *detection interval* as 3 sec. Assuming $c = 1$, from Equation (1) and (2) we get $N = 6.3$.

6 Performance Evaluation

To evaluate the detection efficiency of our mechanism we validate VSC by using both available benchmark datasets

Table 1: Network traces used in the experiments

Trace	Duration	Created on	Type	Label
TU_170813	1 hr	Aug,2013	Bi-Directional	normal
DARPA	3 week	1999	Bi-Directional	attack
CAIDA2007	1 hr	2007	Uni-Directional	attack

as well as our own testbed network traces. Table 1 describes the traces used in our experiments.

Also to show the effectiveness of VSC under different attack scenario, we embedded simulated attack traffic of various characteristics generated in our testbed into the normal traffic of Table 1. Following sections provide the description of our testbed followed by different experiments and their results.

6.1 Testbed Setup

Our testbed consists of two sub-networks as shown in Figure 7.

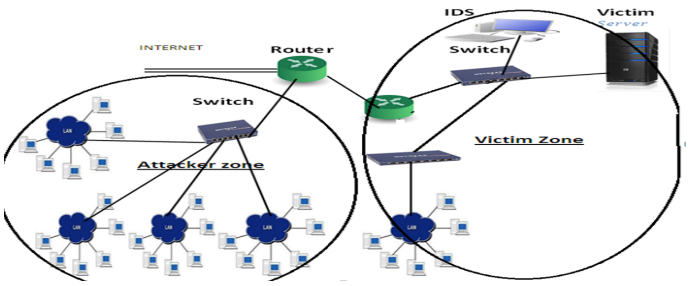
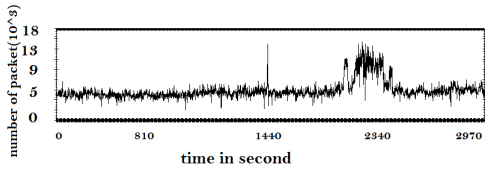


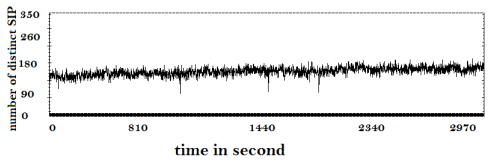
Figure 7: The testbed of our experiments

1) The first sub-network (marked by the left side oval) is used to generate the attack traffic. In this sub-network we installed TU-CANNON as explained earlier.

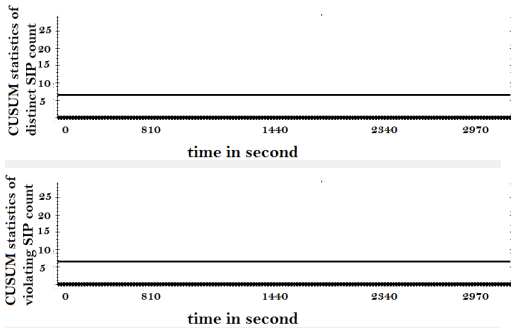
2) The second sub-network (marked by the right side oval) contains the victim server and also the capturing and detection unit (labeled as IDS in the diagram).



(a) packet rate of in normal scenario



(b) unique IP rate in normal scenario



(c) CUSUM statistics under normal condition

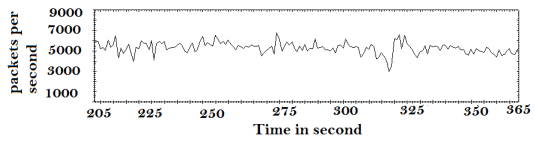
Figure 8: Result of VSC on attack scenario 1

6.2 VSC Under Normal Condition

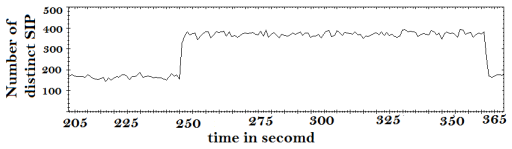
To perform our experiments we used the TU_170813 dataset as normal traffic reference. We are assuming this traffic as attack free traffic. the result of VSC when applied on TU_170813 is shown in Figure 8. We can see that both of our test statistics are much below their threshold value.

6.3 Detection of Low Rate DDoS Attack

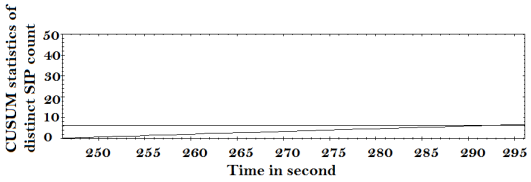
Our detection mechanism detects attack based on the deviation of the number of unique source IP address from its mean value, rather than the volume of the traffic.thus our detection mechanism can detect low rate DDoS attack involving large number of sources. To demonstrate this we embedded a simulated attack of duration 1 minutes into the normal traffic. The attack was performed at a rate of 250 packets/sec, which is much lesser than the usual traffic of the network. Hence such an attack can easily escape a traffic volume based detection mechanism. However as shown in Figure 9, our detection mechanism detects the attack within 23 sec of the starting of the attack. However, during the attack period the traffic volume does not change significantly.



(a) packet rate of in attack scenario

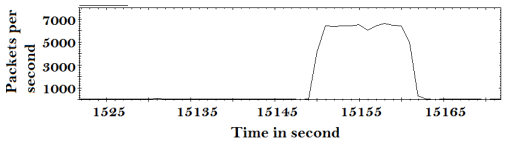


(b) unique IP rate

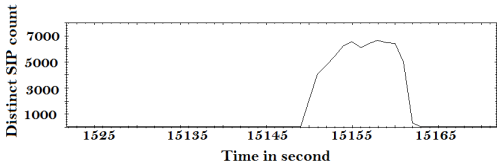


(c) CUSUM detection of attack

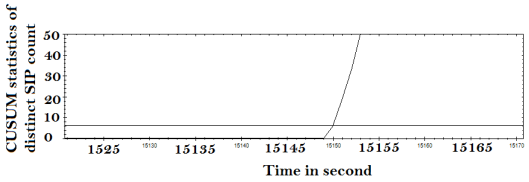
Figure 9: Result of VSC on attack scenario 1



(a) Packet rate of DARPA



(b) Unique IP rate of DARPA



(c) CUSUM detection in DARPA

Figure 10: Result of VSC on DARPA dataset

6.4 Detection of Randomly Spoofed Source IP Attack

We used the DARPA dataset [17] in our experiments to show the effectiveness of VSC in detecting DDoS attack which uses a randomly spoofed source IP addresses.Figure 11 demonstrate the result of this experiment. From the figure we can see that the attack present in the DARPA dataset is easily detectable by VSC in less than 3 seconds.

6.5 Detection of DDoS Attack From A Small Size Bot

The attacker may use a small size bot net (consisting a small number of sources) to carryout the attack. However in that case the speed of the individual source need to be high to end up in an effective DDoS attack. We used a simulated attack consisting of 7 sources, performing an attack at 2000 packets/sec, for a duration of 10 minutes. we used 150 packets/sec as the threshold to mark an IP as violating IP. From Figure 11 it is clear that VSC can detect such an attack in less than 15 seconds, by detecting a change in the violating source IP count.

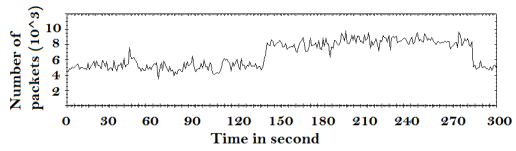
6.6 Detection of DDoS Attack in CAIDA2007 Dataset

CAIDA 2007 [7] is an widely used and benchmark DDoS network trace. We applied VSC on this dataset in our experiment. The result of VSC is shown in Figure 9(a). The length of CAIDA is around 1 hour. The first 30 minutes contains traffic at a low rate, from a small number of sources. However at the begining of the second 30 minitue half, there is an abrupt change in both number of unique source IP address as well as the traffic volume. VSC detects this change with in 8 seconds.

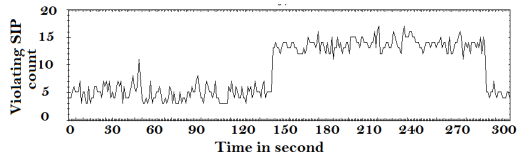
7 Conclusion and Future Work

In this paper we present a robust and low cost method to detect DDoS attack. Our method detects DDoS attack by monitoring the deviation of the count of unique source IP and the count of source IPs whose transmission rate is higher than a threshold value. Another feature of our detection mechanism is that to detect attack performed from small size bot net, i.e, where the transmission rate of each source is very high we keep track of the count of such sources. In case of an attack, this count value deviates from its mean value abruptly, and thus we can detect this change to confirm the attack.

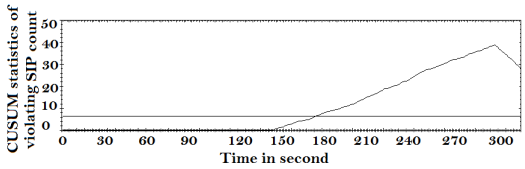
In this paper we present the experimental results of



(a) Packet rate of small size BOT

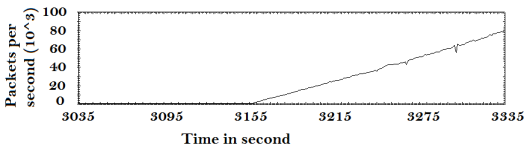


(b) High speed IP rate of small size BOT

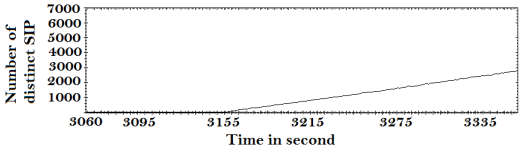


(c) CUSUM detection on small size BOT

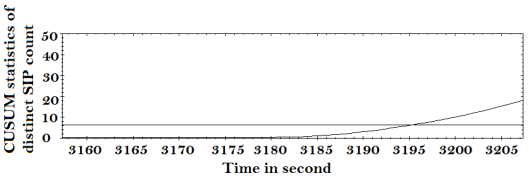
Figure 11: Result of VSC on small size BOT



(a) packet rate of CAIDA



(b) unique IP rate of CAIDA



(c) CUSUM detection on CAIDA

Figure 12: Result of VSC on CAIDA

applying VSC in a last mile router in the context of simulated network traces as well as benchmark datasets like CAIDA and DARPA . Our future research will be on the deployment of the detection mechanism in a distributed manner. A distributive approach will be able to detect the attack as it propagates towards the victim. Thus the victim will get enough time to take necessary precaution.Also a distributive approach will be able to detect the attack near to the source, which is a very desired property of an detection system.

Acknowledgments

We are thankful to MHRD and UGC of Govt of India for funding our work under the schemes (i) Centre of Excellence under FAST and (ii)Special assistance under DRS II.

References

[1] A. Ain, M. H. Bhuyan, D. K. Bhattacharyya, and J. K. Kalita, "Rank Correlation for Low-Rate DDoS Attack Detection: An Empirical Evaluation," *International Journal of Network Security*, vol. 18, no. 3, pp. 474–480, 2016.

[2] A. Anurag, "Network neutrality: Developing business model and evidence based net neutrality regulation," *International Journal of Electronics and Information Engineering*, vol. 3, no. 1, pp. 1–9, 2015.

[3] M. Basseville, and I. V. Nikiforov, *Detection of Abrupt Changes: Theory and Application (vol. 104)*, Englewood Cliffs: Prentice Hall, 1993.

[4] Y. Chen, S. Das, P. Dhar, A. El-Saddik, and A. Nayak, "Detecting and Preventing IP-spoofed Distributed DoS Attacks," *International Journal of Network Security*, vol. 7, no. 1, pp. 69–80, 2008.

[5] R. B. Blazek, H. Kim, B. Rozovskii, and A. Tartakovsky, "A novel approach to detection of denial-of-service attacks via adaptive sequential and batch-sequential change-point detection methods," in *Proceedings of IEEE Systems, Man and Cybernetics Information Assurance Workshop*, pp. 220–226, 2001.

[6] E. Brodsky, and B. S. Darkhovsky, *Nonparametric Methods in Change Point Problems (vol. 243)*, Springer Science & Business Media, 2013.

[7] CAIDA UCSD, *DDoS Attack 2007 Dataset*, 2007. (`http://www.caida.org/data/passive/ddos-20070804_dataset.xml`)

[8] G. Carl, G. Kesidis, R. R. Brooks, and S. Rai, "Denial-of-service attack-detection techniques," *IEEE Internet Computing*, vol. 10, no. 1, pp. 82–89, 2006.

[9] Y. Chen, K. Hwang, and W. S. Ku, "Collaborative detection of DDoS attacks over multiple network domains," *IEEE Transactions on Parallel and*

Distributed Systems, vol. 18, no. 12, pp. 1649–1662, 2007.

[10] CNN, *Immense Network Assault Takes Down Yahoo*, Feb. 2000. (`http://www.cnn.com/2000/TECH/computing/02/08/-yahoo.assault.idg/index.html`)

[11] CNN, *Cyber-attacks Batter Web Heavyweights*, Feb. 2000. (`http://www.cnn.com/2000/TECH/computing/02/09/-cyber.attacks.01/index.html`)

[12] P. J. Criscuolo, *Distributed Denial of Service: Trin00, Tribe Flood Network*, Tribe Flood Network 2000, and Stacheldraht CIAC-2319, California Univ Livermore Radiation Lab., Feb. 2000.

[13] D. Dean, M. Franklin, and A. Stubblefield, "An algebraic approach to IP traceback", *ACM Transactions on Information and System Security*, vol. 5, no. 2, pp. 119–137, 2002

[14] N. Jeyanthi, and N. C. S. N. Iyengar, "An Entropy based approach to detect and distinguish DDoS attacks from flash crowds in VoIP networks," *International Journal of Network Security*, vol. 14, no. 5, pp. 257–269, 2012.

[15] R. Mahajan, S. M. Bellovin, S. Floyd, J. Ioannidis, V. Paxson, and S. Shenker, "Controlling high bandwidth aggregates in the network," *ACM SIGCOMM Computer Communication Review*, vol. 32, no. 3, pp. 62–73, 2002.

[16] J. Mirkovic and P. Reiher, "A taxonomy of DDoS attack and DDoS defense mechanisms", *ACM SIGCOMM Computer Communication Review*, vol. 34, no. 2, pp. 39–53, 2004.

[17] MIT Lincoln Laboratory, *2000 Darpa Intrusion Detection Scenario Specific Data Sets*, 2000.

[18] D. Moore, and C. Shannon, "Code-Red: a case study on the spread and victims of an Internet worm," in *Proceedings of the 2nd ACM SIGCOMM Workshop on Internet Measurment*, pp. 273–284, 2002.

[19] J. Nazario, *Defense and Detection Strategies Against Internet Worms*, Artech House, 2004.

[20] P. Negi, A. Mishra, and B. B. Gupta, "Enhanced CBF packet filtering method to detect DDoS attack in cloud computing environment," arXiv preprint arXiv:1304.7073, 2013.

[21] 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.

[22] T. Peng, C. Leckie, and K. Ramamohanarao, "Proactively detecting distributed denial of service attacks using source IP address monitoring," in *International Conference on Research in Networking*, pp. 771–782, Springer Berlin Heidelberg, 2004.

[23] M. Sachdeva, G. Singh, K. Kumar, and K. Singh, "DDoS incidents and their impact: A review," *International Arab Journal of Information Technology*, vol. 7, no. 1, pp. 14–19, 2010.

[24] S. Savage, D. Wetherall, A. Karlin, and T. Anderson, "Network support for IP traceback", *IEEE/ACM Transactions on Networking*, vol. 9, no. 3, pp. 226–237, 2001.

[25] D. X. Song, and P. Adrian, "Advanced and authenticated marking schemes for IP traceback", in *Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM'01)*, pp. 878–886, 2001.

[26] H. Wang, D. Zhang, and K. G. Shin, "Detecting SYN flooding attacks," in *Proceedings of Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM'02)*, vol. 3, pp. 1530–1539, 2002.

[27] D. K. Yau, J. Lui, F. Liang, and Y. Yam, "Defending against distributed denial-of-service attacks with max-min fair server-centric router throttles," *IEEE/ACM Transactions on Networking*, vol. 13, no. 1, pp. 29–42, 2005.

[28] S. T. Zargar, J. Joshi, and D. Tipper, "A survey of defense mechanisms against distributed denial of service (DDoS) flooding attacks," *IEEE Communications Surveys and Tutorials*, vol. 15, no. 4, pp. 2046–2069, 2013.

Biography

Ram Charan Baishya is a Ph.D. student in the Department of Computer Science and Engineering at Tezpur University.

Nazrul Hoque is a Ph.D. student in the Department of Computer Science and Engineering at Tezpur University.

Dhruba Kr Bhattacharyya received his Ph.D. in Computer Science from Tezpur University in 1999. He is a Professor in the Computer Science & Engineering Department at Tezpur University. His research areas include Data Mining, Network Security and Content based Image Retrieval. Prof. Bhattacharyya has published 220+ research papers in the leading international journals and conference proceedings. In addition, Dr Bhattacharyya has written/ edited 8 books. He is a Programme Committee/ Advisory Body member of several international conferences/workshops.

An Efficient and Provably Secure Certificateless Key Insulated Encryption with Applications to Mobile Internet

Libo He, Chen Yuan, Hu Xiong, and Zhiguang Qin

(Corresponding author: Libo He)

School of Information and Software Engineering & University of Electronic Science and Technology of China
Chengdu, Sichuan, 610054, China
(Email: libowqrs@gmail.com)

(Received Apr. 14, 2016; revised and accepted July 19 & Sept. 26, 2016)

Abstract

Certificateless encryption (CLE) alleviates the heavy certificate management in traditional public key encryption and the key escrow problem in the ID-based encryption simultaneously. Current CLE schemes assumed that the user's secret key is absolutely secure. Unfortunately, this assumption is too strong in case the CLE is deployed in the hostile setting and the leakage of the secret key is inevitable. In this paper, we present a new concept called a certificateless key insulated encryption scheme(CL-KIE). We argue that this is an important cryptographic primitive that can be used to achieve key-escrow free and key-exposure resilience. We also present an efficient CL-KIE scheme based on bilinear pairing. After that, the security of our scheme is proved under the Bilinear Diffie-Hellman assumption in the random oracle model. Further, the potential applications of CL-KIE is also briefly illustrated.

Keywords: Bilinear Pairing; Certificateless Cryptography; Key-insulated

1 Introduction

The public-key cryptography is called asymmetric key encryption, as every user owns a pair of keys: a public key and a private key. In 1978, Ron Rivest, Adi Shamir, and Leonard Adleman [15] published the first practical public key encryption RSA algorithm. The security of this algorithm is relied on practical difficulty of factoring the product of two large prime number. Another widely used public key cryptography Elgama algorithm based on the Diffie-Hellman key exchange was described by Taher Elgamal [10] in 1984.

The public key cryptosystem needs Public Key Infrastructure (PKI) to offer the authentication and validation for the public key [19]. But PKI will encounter a lot of challenges on efficiency and scalability for its complicated structure. In 1984, the Identity-based Encryption was firstly proposed by Shamir [24]. In 2001, Dan Boneh and Matthew K. Franklin [7] proposed a practical identity-based encryption system based on Weil pairing over elliptic curves and finite fields. In IBE, the public key could be any arbitrary characters related to user's identity [13, 22].

The private key is derived from the identity of an entity and the master key only known by Key Generation Center (KGC). So the certificate which is used to authenticate public key will not be necessary. However, the key escrow problem arises that the malicious authority can impersonate any users to get the corresponding private key.

To solve the problem of key escrow in Identity-based Encryption and guarantee the authenticity of public keys without the use of the certificate in public key encryption, the certificateless public key encryption (CL-PKE) has been introduced by Al-Riyami and Paterson [1] in 2003. In CL-PKE, the private key is separated into two parts: one partial private key is still generated in KGC, and the secret key is selected by the user itself. The malicious KGC only can get the partial private key, so it can not impersonate any user to attack the system. Hence, the CL-PKE solves the problems of key escrow. Since then, CL-PKE becomes a research hotspot and several other relevant certificateless encryption schemes [3, 4, 5, 8, 14, 20, 23] have been developed. Until then, current CL-PKE schemes assumed that the user's secret key is absolutely secure. However, a higher security requirement in CL-PKE is needed, for example, the case where an adversary steals the whole private key.

The exposure of private key is a devastating disaster for the cryptosystem. Key-evolving cryptosystem can alleviate the damage of key leakage. Normally, Key-evolving cryptosystem can be categorized into three groups as follows: forward-security [2, 26], key-insulation [9, 11, 16, 18, 21, 25] and intrusion-resilience [12]. In the key-evolving cryptosystem, the lifetime of the system is

Table 1: Functionality comparison

	Public key management-Free	Key-escrow Free	Key-insulation
PKE	×	×	×
IBE	√	×	×
CL-PKE	√	√	×
our scheme	√	√	√

divided into N time periods. For forward-secure scheme, the private key is updated by the user himself during every time period without any interaction with other devices. Even when an adversary compromise the private key at the current time period, the forward-secure scheme can also guarantee the security of the prior time periods. In the key-insulated scheme, a user's private key is updated by communicating with a physically-secure device for every time period. The private key is composed of two parts: one part is generated by the master key and the other is created by the helper key from the physically-secure device. Meanwhile, the public key remains fixed during the whole periods of key updating. By this approach, even an adversary who steals the private key in the present time period cannot get the private key in the former or later period. The private key in the intrusion-resilient scheme also will be updated by interaction with a physically-secure device.

The difference between the key-insulated scheme and the intrusion-resilient scheme is that the intrusion-resilient scheme refreshes the secret keys of the user and physically-secure device many times in one period. So intrusion-resilient scheme remains secure even after many arbitrary compromises of both user and physically-secure device, as long as the compromise are not simultaneous. Among the above three types of key-evolving schemes, the key-insulated scheme and the intrusion-resilient scheme can offer higher security than the forward-secure scheme. Also, the intrusion-resilient scheme gives more security and is less efficient compared with the key-insulated scheme. Therefore, the key-insulated scheme is the trade-off between security and efficiency as shown in Table 1.

1.1 Contribution

In this paper, we resolve the above problem and make the following novel contributions as follows:

- Firstly, we present a concrete paradigm called the certificateless key-insulated encryption scheme (CL-KIE). We integrate key-insulated technique into CL-PKE. Through this method, our new scheme not only can solve the problem of key escrow, but also can achieve the functionality of key-insulation, without heavy public key management.

- We give the formal definition and security model for CL-KIE scheme and the construction of a CL-KIE scheme based on bilinear pairing. We also give the security proof for the CL-KIE scheme under the Bilinear Diffie-Hellman assumption in the random oracle model.

- Finally, our scheme with key updates can give more security fuctionality to solve the problem of key exposure compared with some other CL-PKE schemes, while sacrificing a little on the cost of execution time. This is an attractive advantage which the standard CL-PKE scheme does not possess.

1.2 Organization

The rest of this paper is organized as follows: We formalize the definition and give the security model of CL-KIE schemes in Section 2. Section 3 first gives an introduction to bilinear pairings and the Bilinear Diffie-Hellman Problem, then proposes a construction of the CL-KIE scheme. We prove that our scheme under the Bilinear Diffie-Hellman assumption in the random oracle model and compare our scheme with CL-PKE scheme on efficiency and security capability in Section 4. Further, the potential applications for CL-KIE is discussed in Section 5. At last, we conclude this paper in Section 6.

2 Formal Definition and Security Model

In this section we first formalize the definition of the CL-KIE scheme by cooperating the key-insualted scheme and the CL-PKE scheme. After that, we propose the security model of the CL-KIE scheme.

2.1 Definition of CL-KIE

We denote the CL-KIE scheme, which consists of the following algorithms:

Setup: The algorithm is given a security parameter k, and generates the system parameters $params$, `master-key` and `master-helper-key`. The system parameters include a description of a finite message space $\mathcal{M}$, a description of a finite ciphertext space $\mathcal{C}$ and a randomness space $\mathcal{R}$.

SecretValExtract: The algorithm takes as input $params$ and a identity string $ID_A \in \{0,1\}^*$, and

generates a random $x_A \in Z_q$ as the secret value associated with the entity A.

PartialKeyExtract: The algorithm takes as input *params*, `master-key`, and a identity string ID_A, and return the partial private key D_A corresponding to the entity A.

HelperKeyUpdate: The algorithm takes as input *params*, a time period i, `master-helper-key`, a identity string ID_A, and return the helper key $HK_{A,i}$ for a time period i.

PrivateKeyUpdate: The algorithm takes as input *param*, a time period i, a helper key $HK_{A,i}$, an identity string ID_A, a partial private key D_A and a secret value x_A, and output the private key $S_{A,i}$ for a time period i.

PublicKeyExtract: The algorithm takes as input *params*, a secret value x_A and a identity string ID_A, and output the public key P_A of the entity A.

Encrypt: The algorithm takes as input a time period i, *params*, an identity string ID_A, a public key P_A and a plaintext $M \in \mathcal{M}$. It returns the ciphertext $C \in \mathcal{C}$.

Decrypt: The algorithm takes as input a time period i, *params*, a private key $S_{A,i}$ and a ciphertext C. It returns the corresponding plaintext $M \in \mathcal{M}$.

2.2 Security Model

In this subsection we define the security model for the CL-KIE scheme by Indistinguishability of Encyption Against Adaptive Chosen Ciphertext Attacker (IND-CPA) game which is conducted between a challenger $\mathcal{S}$ and an adversary $\mathcal{A}$. In our scheme, we define two kinds of adversaries $TypeI$ adversary ($\mathcal{A}_1$) and $TypeII$ adversary ($\mathcal{A}_2$): $\mathcal{A}_1$ represents the external attacker who can not access the *master-key* but can replace the public key for an entity with its choice; $\mathcal{A}_2$ represents the malicious KGC who can access the *master-key*. We prohibit $\mathcal{A}_2$ from replacing the public key since $\mathcal{A}_2$ can not seclect the secret value by itself. First we give a list of oracles that a general adversary in our scheme may carry out, then we define chosen ciphertext security of CL-KIE for two kinds of adversaries respectively.

The list of oracles that a general adversary in CL-KIE may carry out is the following:

- **Partial-Private-Key-Queries:** If necessary, $\mathcal{A}$ makes **Partial-Private-Key-Queries** on the identity ID_A, and $\mathcal{S}$ returns the partial private key D_A associated with ID_A to $\mathcal{A}$.
- **Helper-Key-Queries:** $\mathcal{A}$ makes **Helper-Key-Queries** on the identity ID_A at a time period i, and $\mathcal{S}$ returns the helper key $HK_{A,i}$ to $\mathcal{A}$.
- **Secret-Value-Queries:** If necessary, $\mathcal{A}$ makes **Secret-Value-Queries** on the identity ID_A, and $\mathcal{S}$ returns the secret value x_A associated with ID_A to $\mathcal{A}$.
- **Public-Key-Queries:** $\mathcal{A}$ makes **Public-Key-Queries** on the identity ID_A, and $\mathcal{S}$ returns the helper key P_A to $\mathcal{A}$.
- **Public-Key-Replace:** If necessary, $\mathcal{A}$ can repeatedly make **Public-Key-Replace** to set the public key P_A for any value of its choice.
- **Decryption-Queries:** $\mathcal{A}$ makes **Decryption-Queries** for a ciphertext C on the identity ID_A at a time period i. If the recovered redundancy in M is valid, $\mathcal{S}$ returns the associated plaintext M to $\mathcal{A}$.

Semantic security against an adaptive chosen ciphertext for a KI-CLPKE scheme can be defined via the following games between two different Advesaries ($\mathcal{A}_1$ and $\mathcal{A}_2$) and Challenger $\mathcal{S}$:

- **Chosen Plaintext Security for CL-KIE on $\mathcal{A}_1$**

  Setup: $\mathcal{S}$ takes as input a security parameter k and execute the **Setup** algorithm. It returns *params* except *master-key* to $\mathcal{A}_1$.

  Phase 1: $\mathcal{A}_1$ can access a sequence of oracles: **Partial-Private-Key-Queries**, **Helper-Key-Queries**, **Secret-Value-Queries**, **Public-Key-Replace**, **Decryption-Queries**. These queries may be requested adaptively, but restricted by the rule of adversary behavior.

  Challenge: $\mathcal{A}_1$ outputs two equal length plaintext $M_0^*, M_1^* \in \mathcal{M}$ on the challenge identity ID_A^* at a time period i^*. The challenge $\mathcal{S}$ pick a random number $b \in \{0,1\}$ and generate C^* in relation to (i^*, M_b^*, ID^*). C^* is delivered to $\mathcal{A}_1$ as a target challenge.

  Phase 2: $\mathcal{A}_1$ continues to access a sequence of oracles as in Phase 1, and $\mathcal{S}$ responds to these queries as in Phase 1.

  Guess: At the end, $\mathcal{A}_1$ outputs a guess $b' \in \{0,1\}$. The adversary wins the game if $b = b'$. We define $\mathcal{A}_1's$ advantage in this game to be $Adv(\mathcal{A}_1) = 2(Pr[b = b'] - \frac{1}{2})$.

  There are a few restrictions on the $\mathcal{A}_1$ as follows:

  - $\mathcal{A}_1$ are not allowed to extract the partial private key for ID_A^*.
  - In Phase 2, we insist that $\mathcal{A}_1$ cannot make a decryption query on the challenge ciphertext C^* in the relation to the identity ID_A^* and the public key P_A^*.

- **Chosen Plaintext Security for CL-KIE on $\mathcal{A}_2$**

Setup: $\mathcal{S}$ takes as input a security parameter k and execute the *Setup* algorithm. It returns *params* to $\mathcal{A}_2$.

Phase 1: $\mathcal{A}_2$ can access a sequence of oracles: **Helper-Key-Queries**, **Public-Key-Queries**, **Decryption-Queries**. These queries may be requested adaptively, but restricted by the rule of adversary behavior.

Challenge: $\mathcal{A}_2$ outputs two equal length plaintext $M_0^*, M_1^* \in M$ on the challenge identity ID_A^* and a time period i^*. The challenger $\mathcal{S}$ pick a random number $b \in \{0,1\}$, and generate C^* in ralation to (i^*, M_b^*, ID^*). C^* is delivered to $\mathcal{A}_2$ as a target challenge.

Phase 2: $\mathcal{A}_2$ continues to access a sequence of oracles as in Phase 1, and $\mathcal{S}$ responds to these queries as in Phase 1.

Guess: At the end, $\mathcal{A}_2$ outputs a guess $b' \in \{0,1\}$. The adversary wins the game if $b = b'$. We define $\mathcal{A}_2's$ advantage in this game to be $Adv(\mathcal{A}_2) = 2(Pr[b = b'] - \frac{1}{2})$.

There are a few restrictions on the $\mathcal{A}_2$ as follows:

- The **Secret-Value-Queries** is not allowed to access if the public key for entity has been replaced.
- $\mathcal{A}_2$ are not allowed to replace the public key for ID_A^*.
- $\mathcal{A}_2$ are not allowed to extract the secret value for ID_A^*.
- In phase 2, we insist that $\mathcal{A}_2$ cannot make a decryption query on the challenge ciphertext C^* in the relation to the identity ID_A^* and the public key P_A^*.

3 KI-CLPKE Scheme

3.1 Bilinear Pairing and Bilinear Diffie-Hellman (BDH) Problem

Bilinear Pairing

Let $\mathbb{G}_1$ denotes a cyclic additive group of order q for some large prime q, let $\mathbb{G}_2$ be a cyclic multiplicative group of the same order q, We can make use of a bilinear map:$\hat{e} : \mathbb{G}_1 \times \mathbb{G}_1 \to \mathbb{G}_2$ above these two groups which must satisfy the following properties:

- **Bilinearity:** $\hat{e}(aP, bQ) = \hat{e}(P,Q)^{ab}$, where $P, Q \in \mathbb{G}_1$, and $a, b \in \mathbb{Z}_q^*$.
- **Non-Degeneracy:** If P is the generator for $\mathbb{G}_1$, $\hat{e}(P,P)$ is the generator for $\mathbb{G}_2$.
- **Computability:** For $\forall P, Q \in \mathbb{G}_1$, $\hat{e}(P,Q)$ can be computed through an efficient algorithm in a polynomial-time.

Bilinear Diffie-Hellman(BDH) Problem

BDH Problem is for $a, b, c \in \mathbb{Z}_q$, given $P, aP, bP, cP \in \mathbb{G}_1$, to compute abc which satisfies $\hat{e}(P,Q)^{abc} \in \mathbb{G}_2$.

3.2 Construction

Setup: We can randomly select a security parameters $k \in \mathbb{Z}^+$, the Setup algorithm works as follows:

Step 1: Pick two groups $(\mathbb{G}_1, +)$ and $(\mathbb{G}_2, \times)$ of the same prime order q where $|q| = k$. Choose a generator P over $\mathbb{G}_1$ randomly, we can get a bilinear map $\hat{e} : \mathbb{G}_1 \times \mathbb{G}_1 \to \mathbb{G}_2$.

Step 2: Choose a random $s \in \mathbb{Z}_q$ to compute $P_{pub} = sP$, the corresponding s can be regarded as the *master-key*: $M_{mk} = s$;

Choose a random $w \in \mathbb{Z}_q$ to compute $P_{hk} = wP$, the corresponding w can be regarded as the *master-helper-key*: $M_{hk} = w$.

Step 3: For some integer $n > 0$,we can select three cryptographic hash functions:

- $H_1 : \{0,1\}^n \to \mathbb{G}_1$.
- $H_2 : \{0,1\}^n \times \mathbb{Z}^+ \to \mathbb{G}_1$.
- $H_3 : \mathbb{G}_1 \times \mathbb{G}_2 \to \{0,1\}^n$.

The system parameters $params = (\mathbb{G}_1, \mathbb{G}_2, p, \hat{e}, n, P, P_{pub}, P_{hp}, H_1, H_2, H_3)$. The master key $M_{mk} = s$ and the master helper key $M_{hk} = w$.

The message space is $\mathcal{M} = \{0,1\}^n$, the ciphertext space is $\mathcal{C} = \{0,1\}^n \times \{0,1\}^n$, the randomness space is $\mathcal{R} = \{0,1\}^n$.

SecretValExtract $(params, ID_A)$**:** For a given identity ID_A and *params*, the algorithm outputs a random $x_A \in Z_q$ as the secret value for entity A.

PartialKeyExtrat $(params, M_{mk}, ID_A)$**:** For a given identity $ID_A \in \{0,1\}^*$ of entity A, *params* and M_{mk}, the algorithm computes $D_A = sH_1(ID_A)$.

HelperKeyUpdate $(i, ID_A, M_{hk}, params)$**:** Given a identity string ID_A and a time period $i \in \{0, \ldots, n-1\}$, the helper generates a helper key $HK_{A,i}$ which can help the private key to be updated at the time period $i \in \{0, \ldots, n-1\}$:

$$HK_{A,i} = wH_2(ID_A, i)$$

PrivateKeyExtract $(i, ID_A, HK_{A,i}, params, D_A, x_A)$**:** Given a identity ID_A, At a time period $i \in \{0, \ldots, n-1\}$, the private key is generated as:

$$\begin{aligned} S_{A,i} &= x_A H_1(ID_A) + D_A + HK_{A,i} \\ &= x_A H_1(ID_A) + sH_1(ID_A) + wH_2(ID_A, i) \end{aligned}$$

the value $S_{A,i-1}$ will be deleted subsequently.

PublicKeyExtract $(params, x_A, ID_A)$**:** Given $params$ and x_A, the algorithm outputs $P_A = \langle X_A, Y_A \rangle = \langle x_A P, x_A sP \rangle$.

Encrypt $(i, params, ID_A, P_A, M)$**:** At a time period $i \in \{0, \ldots, n-1\}$, to encrypt a plaintext $M \in \{0,1\}^n$, the algorithm does:

1) Check the equality $\hat{e}(X_A, sP) = \hat{e}(Y_A, P)$ holds. If not, output $\perp$ and abort encryption.
2) Select a random $r \in \mathbb{Z}_q$, $U = rP$.
3) Compute $\xi = \hat{e}(X_A, rH_1(ID_A))\hat{e}(P_{pub}, rH_1(ID_A))\hat{e}(P_{hk}, rH_2(ID_A, i))$.
4) Output the ciphertext: $C = \langle i, U, M \oplus H_3(U, \xi) \rangle$.

Decrypt$(i, params, S_{A,i}, C)$**:** Received the ciphertext $C = \langle i, U, V \rangle$. at the time period $i \in \{0, \ldots, n-1\}$, the algorithm performs the following steps with the Private key $S_{A,i}$:

1) Compute $\xi' = \hat{e}(U, S_{A,i})$.
2) Compute $M' = V \oplus H_3(U, \xi')$.
3) If the recovered redundancy in M is valid, then accept M' the plaintext.

4 Analysis

4.1 Security Proof

Theorem 1. *Let hash functions H_1, H_2, H_3 be random oracles. For TypeI adversary in polynomial time, suppose further that there is no* IND-CPA *adversary $\mathcal{A}_1$ that has non-negligible advangtage against the* KI-CLPKE *scheme. Then the* KI-CLPKE *is* IND-CPA *secure.*

Proof. We first deal with the *TypeI* adversary A_1. For the first type adversary A_1 is external attacker who can not get the *master-key*, Given a BDH problem (P, aP, bP, cP), we can construct a challenger $\mathcal{S}$ to compute $\hat{e}(P,P)^{abc}$ by making use of A_1 as an adversary. Now, we begin to propose the concrete proof.

Setup: Firstly, challenger $\mathcal{S}$ sets $P_{pub} = aP$ and selects $params = (\mathbb{G}_1, \mathbb{G}_2, p, \hat{e}, n, P, P_{pub}, P_{hp})$ then sends $params$ to adversary $\mathcal{A}_1$.

Phase 1: H_1 **queries:** $\mathcal{S}$ keeps a list H_1^{list} of tuples $\langle ID_j, u_j \rangle$ which is initially empty. When $\mathcal{A}_1$ issues a query on ID_i, $\mathcal{S}$ responds as follows:

- If ID_i is on H_1^{list} in a tuple $\langle ID_i, u_i \rangle$, then $\mathcal{S}$ responds with u_i. If $ID_i = ID^*$, then $\mathcal{S}$ set $H_1(ID^*) = bP$.
- Otherwise, $\mathcal{S}$ selects a random integer $u_i \in Z_p$ and stores $\langle ID_i, u_i \rangle$ into the tuple list. $\mathcal{S}$ responds with u_i.

H_2 **queries:** $\mathcal{S}$ keeps a list H_2^{list} of tuples $\langle ID_j, u_j, w_j \rangle$ which is initially empty. When $\mathcal{A}_1$ issues a query on ID_i and u_i, $\mathcal{S}$ responds as follows:

- If ID_i and u_i is on H_2^{list} in a tuple $\langle ID_i, u_i, w_j \rangle$, then β_I responds with w_i.
- Otherwise, $\mathcal{S}$ selects a random integer $w_i \in Z_p$ and stores $\langle ID_i, u_i, w_i \rangle$ into the tuple list. $\mathcal{S}$ responds with w_i.

H_3 **queries:** $\mathcal{S}$ keeps a list H_3^{list} of tuples $\langle u_j, w_j, Str_j \rangle$ which is initially empty. When $\mathcal{A}_1$ issues a query on u_i and w_i, $\mathcal{S}$ responds as follows:

- If u_i and w_i is on H_3^{list} in a tuple $\langle u_i, w_i, Str_i \rangle$, then $\mathcal{S}$ responds with Str_i.
- Otherwise, $\mathcal{S}$ selects a random integer $Str_i \in \{0,1\}^n$ and stores $\langle u_i, w_i, Str_i \rangle$ into the tuple list. $\mathcal{S}$ responds with Str_i.

Partial-Private-Key-Queries: $\mathcal{S}$ keeps a list PP^{list} of tuples $\langle ID_j, D_{A,j} \rangle$. On receiving a query **Partial-Private-Key-Queries**(ID_i), $\mathcal{S}$ responds to the query as follows:

- If $ID_i = ID^*$, $\mathcal{S}$ aborts.
- Else, if ID_i is on the list in the tuple $\langle ID_i, D_{A,i} \rangle$, then $\mathcal{S}$ responds with $D_{A,i}$
- Else, $\mathcal{S}$ first searches H_1^{list} for the tuple with ID_i. If no such tuple is found then $H_1(ID_i)$ is queried. Then $\mathcal{S}$ compute $D_{A,i} = sH_1(ID_i)$ and output $D_{A,i}$ as the answer.

Helper-Key-Queries: $\mathcal{S}$ keeps a list HK^{list} of tuples $\langle ID_j, j, HK_{A,j} \rangle$. On receiving a query **Helper-Key-Queries**$(ID_i,\ i)$, $\mathcal{S}$ responds to the query as follows:

- If $ID_i = ID^*$, $\mathcal{S}$ aborts.
- Else, if ID_i and the time period i are on the list in the tuple $\langle ID_i, i, HK_{A,i} \rangle$, then $\mathcal{S}$ responds with $HK_{A,i}$.
- Else, $\mathcal{S}$ first searches H_2^{list} for the tuple with ID_i and the time period i. If no such tuple is found then $H_2(ID_i, i)$ is queried. Then $\mathcal{S}$ compute $HK_{A,i} = wH_2(ID_i, i)$ and then output $HK_{A,i}$ as the answer.

Secret-Value-Queries: $\mathcal{S}$ keeps a list SV^{list} of tuples $\langle ID_j, x_{A,j} \rangle$. On receiving a query **Secret-Value-Queries**(ID_i), $\mathcal{S}$ responds to the query as follows:

- If $ID_i = ID^*$, $\mathcal{S}$ aborts.
- Else, if ID_i and $x_{A,i}$ is on SV^{list} in a tuple $\langle ID_i, x_{A,i} \rangle$, then $\mathcal{S}$ responds with $x_{A,i}$.
- Eles, $\mathcal{S}$ selects a random integer $x_{A,i} \in Z_q$ and stores $\langle ID_i, x_{A,i} \rangle$ into the tuple list. $\mathcal{S}$ responds with $x_{A,i}$.

Public-Key-Queries: $\mathcal{S}$ keeps a list PK^{list} of tuples $\langle ID_j, P_{A,j}\rangle$. On receiving a query **Public-Key-Queries**(ID_i), $\mathcal{S}$ responds to the query as follows:

- If ID_i is on the list in the tuple $\langle ID_i, P_{A,i}\rangle$. Then $\mathcal{S}$ responds with $P_{A,i}$.
- Otherwise $\mathcal{S}$ first searches S^{list} for the tuple with ID_i. If no such tuple is found then Secret-Value-Queries(ID_i) is queried. Then $\mathcal{S}$ compute $X_A = x_{A,i}P, Y_A = x_{A,i}sP$ and output $P_A = \langle X_A, Y_A\rangle$ as the answer.

Public-Key-Replace: Assume a query that is to replace the public key for ID_i with value $\langle X_i', Y_i'\rangle$. If $\hat{e}(X_i', P_0) = \hat{e}(Y_i', P)$, then $P_A'(\langle X_A', Y_A'\rangle)$ is a valid public key. $\mathcal{S}$ replace the public key with new values $\langle X_i', Y_i'\rangle$.

Decryption-Queries: On receiving a query **Decryption-Queries**(ID_i, C_i) where $C_i = (i, U_i, V_i)$, $\mathcal{S}$ responds to the query as follows:

- If $ID_i = ID^*$, $\mathcal{S}$ aborts.
- Else, $\mathcal{S}$ derives the private key $S_{A,i} = x_A H_1(ID_A) + sH_1(ID_A) + wH_2(ID_A, i)$, then compute $\xi_i' = \hat{e}(U, S_{A,i})$.
- Else,$\mathcal{S}$ first searches H_3^{list} for the tuple with (U_i, ξ_i'). If no such tuple is found then $H_3(U_i, \xi_i')$ is queried. Then $\mathcal{S}$ compute $M' = V \oplus H_3(U_i, \xi_i')$, and output M' as the answer.

Challenge phase: $\mathcal{A}_1$ outputs two equal length plaintext $M_0^*, M_1^* \in \mathcal{M}$ on the challenge identity ID_A^* at a time period i^*. The challenge $\mathcal{S}$ picks a random number $b \in \{0,1\}$, sets $U^* = cP$, and generates C^* in relation to (i^*, M_b^*, ID^*). C^* is delivered to $\mathcal{A}_1$ as a target challenge.

Phase 2: $\mathcal{A}_1$ continues to access a sequence of oracles as in Phase 1, and $\mathcal{S}$ responds to these queries as in Phase 1.

Guess: At the end, $\mathcal{A}_1$ outputs a guess $b' \in \{0,1\}$. The adversary wins the game if $b = b'$. We define $\mathcal{A}_1's$ advantage in this game to be $Adv(\mathcal{A}_1) = 2(Pr[b = b'] - \frac{1}{2})$.

When the challenge games begin, $\mathcal{S}$ sets $P_{pub} = aP$ as an instance of BDH problem and simulates hash functions as random oracles. During the simulation, $\mathcal{S}$ needs to guess every bit in target plaintext M_1^* with a time period i^*. $\mathcal{S}$ will set $H_1(ID_A^*) = bP$, $H_2(ID_A^*, i^*) = (h^*{}_{,i^*}P)$, $V^* = H_3(U^*, \xi^*) = H_3(cP, \xi^*)$. After that, $\mathcal{S}$ returns a simulated ciphertext $C^* = (i^*, U^*, V^*)$, which implies the parameter ξ^* is defined as:

$$\begin{aligned}\xi^* =& \hat{e}(X_A, rH_1(ID_A^*))\hat{e}(P_{pub}, rH_1(ID_A^*))\\ & \hat{e}(P_{hk}, rH_2(ID_A^*, i^*))\\ =& \hat{e}(x_A rP, bP)\hat{e}(bP, acP)\hat{e}(wP, r(h^*{}_{,i^*}P))\\ =& \hat{e}(P,P)^{abc}\hat{e}(aP, cP)^{x_A}\hat{e}(wP, (h^*{}_{,i^*})cP)\end{aligned}$$

Above all, $\mathcal{S}$ can get the solution for BDH problem, i.e. $\hat{e}(P,P)^{abc} = \xi^*(\hat{e}(aP, cP)^{-x_A}\hat{e}(wP, (h^*{}_{,i^*})cP))^{-1}$. Thus we have proved the security of the scheme for the $TypeI$ adversary through this reduction.

□

Theorem 2. *Let hash functions H_1, H_2, H_3 be random oracles. For TypeII adversary in polynomial time, suppose further that there is no* IND-CPA *adversary $\mathcal{A}_2$ that has non-negligible advangtage against the* KI-CLPKE *scheme. Then the* KI-CLPKE *is* IND-CPA *secure.*

Proof. We secondly deal with the $TypeII$ adversary A_2. For the $TypeII$ adversary is a malicious KGC attacker who can get the *master-key*, Given a BDH problem (P, aP, bP, cP), we can construct a challenger $\mathcal{S}$ to compute $\hat{e}(P,P)^{a,b,c}$ by making use of $\mathcal{A}_2$ as an adversary. Now, we begin to propose the concrete proof.

Setup: Firstly, challenger $\mathcal{S}$ selects $params = (G_1, G_2, p, \hat{e}, n, P, P_{pub}, P_{hp})$, then sends $params$ to adversary $\mathcal{A}_2$, where P_{hp} is set as aP.

Phase 1: H_1 **queries:** $\mathcal{S}$ keeps a list H_1^{list} of tuples $\langle ID_j, u_j\rangle$ which is initially empty. When $\mathcal{A}_2$ issues a query on ID_i, $\mathcal{S}$ responds as follows:

- If ID_i is on H_1^{list} in a tuple $\langle ID_i, u_i\rangle$, then $\mathcal{S}$ responds with u_i. If $ID_i = ID^*$, then $\mathcal{S}$ set $H_1(ID^*) = bP$.
- Otherwise, $\mathcal{S}$ selects a random integer $u_i \in Z_p$ and stores $\langle ID_i, u_i\rangle$ into the tuple list. $\mathcal{S}$ responds with u_i.

H_2 **queries:** $\mathcal{S}$ keeps a list H_2^{list} of tuples $\langle ID_j, u_j, w_j\rangle$ which is initially empty. When $\mathcal{A}_2$ issues a query on ID_i and u_i, $\mathcal{S}$ responds as follows:

- If ID_i and u_i is on H_2^{list} in a tuple $\langle ID_i, u_i, w_j\rangle$, then $\mathcal{S}$ responds with w_i.
- Otherwise, $\mathcal{S}$ selects a random integer $w_i \in Z_p$ and stores $\langle ID_i, u_i, w_i\rangle$ into the tuple list. $\mathcal{S}$ responds with w_i.

H_3 **queries:** $\mathcal{S}$ keeps a list H_3^{list} of tuples $\langle u_j, w_j, Str_j\rangle$ which is initially empty. When $\mathcal{A}_2$ issues a query on u_i and w_i, $\mathcal{S}$ responds as follows:

- If u_i and w_i is on H_3^{list} in a tuple $\langle u_i, w_i, Str_i\rangle$, then $\mathcal{S}$ responds with Str_i.

Table 2: Performance comparison

	CL-PKE [1]	**CL-PKE1** [6]	**CL-PKE2** [6]	**Our Scheme**
PartialKeyExtract	M	M	$2M$	$3M$
PublicKeyExtract	$2M$	M	M	$2M$
Encrypt	$M+P+E$	$2M+P$	$4M+E$	$4M+3P$
Decrypt	P	$M+P$	$M+2P$	P
Key-Insulation	$\times$	$\times$	$\times$	$\surd$

- Otherwise, $\mathcal{S}$ selects a random integer $Str_i \in \{0,1\}^n$ and stores $\langle u_i, w_i, Str_i\rangle$ into the tuple list. $\mathcal{S}$ responds with Str_i.

Helper-Key-Queries: $\mathcal{S}$ keeps a list HK^{list} of tuples $\langle ID_j, j, HK_{A,j}\rangle$. On receiving a query **Helper-Key-Queries**$(ID_i,\ i)$, $\mathcal{S}$ responds to the query as follows:

- If $ID_i = ID^*$, $\mathcal{S}$ aborts.
- Else, if ID_i and the time period i are on the list in the tuple $\langle ID_i, i, HK_{A,i}\rangle$, then $\mathcal{S}$ responds with $HK_{A,i}$.
- Else, $\mathcal{S}$ first searches H_2^{list} for the tuple with ID_i and the time period i. If no such tuple is found then $H_2(ID_i, i)$ is queried. Then $\mathcal{S}$ compute $HK_{A,i} = wH_2(ID_i, i)$ and then output $HK_{A,i}$ as the answer.

Public-Key-Queries: $\mathcal{S}$ keeps a list PK^{list} of tuples $\langle ID_j, P_{A,j}\rangle$ where $P_{A,j} = \langle X_A, Y_A\rangle$. $\mathcal{S}$ sets $X_A = aP$. On receiving a query **Public-Key-Queries**(ID_i), $\mathcal{S}$ responds to the query as follows:

- If ID_i is on the list in the tuple $\langle ID_i, P_{A,i}\rangle$. Then $\mathcal{S}$ responds with $P_{A,i}$.
- Otherwise $\mathcal{S}$ first searches S^{list} for the tuple with ID_i. If no such tuple is found then Secret-Value-Queries(ID_i) is queried. Then $\mathcal{S}$ compute $X_A = x_{A,i}P, Y_A = x_{A,i}sP$ and output $P_A = \langle X_A, Y_A\rangle$ as the answer.

Decryption-Queries: On receiving a query **Decryption-Queries**$(ID_i,\ C_i)$ where $C_i = (i, U_i, V_i)$, $\mathcal{S}$ responds to the query as follows:

- If $ID_i = ID^*$, $\mathcal{S}$ aborts.
- Else, $\mathcal{S}$ derives the private key $S_{A,i} = x_A H_1(ID_A) + sH_1(ID_A) + wH_2(ID_A, i)$, then compute $\xi_i' = \hat{e}(U, S_{A,i})$.
- Else,$\mathcal{S}$ first searches H_3^{list} for the tuple with (U_i, ξ_i'). If no such tuple is found then $H_3(U_i, \xi_i')$ is queried. Then $\mathcal{S}$ compute $M' = V \oplus H_3(U_i, \xi_i')$, and output M' as the answer.

Challenge phase: $\mathcal{A}_{II}$ outputs two equal length plaintext $M_0^*, M_1^* \in \mathcal{M}$ on the challenge identity ID_A^* at a time period i^*. The challenge $\mathcal{S}$ pick a random number $b \in \{0,1\}$, sets $U^* = cP$, and generate C^* in relation to (i^*, M_b^*, ID^*). C^* is delivered to $\mathcal{A}_2$ as a target challenge.

Phase 2: $\mathcal{A}_2$ continues to access a sequence of oracles as in Phase 1, and $\mathcal{S}$ responds to these queries as in Phase 1.

Guess: At the end, $\mathcal{A}_2$ outputs a guess $b' \in \{0,1\}$. The adversary wins the game if $b = b'$. We define $\mathcal{A}_2's$ advantage in this game to be $Adv(\mathcal{A}_2) = 2(Pr[b = b'] - \frac{1}{2})$.

When the games begin, $\mathcal{S}$ set $X_A = aP$ as an instance of BDH problem and simulate hash functions as random oracles. During the simulation, $\mathcal{S}$ need to guess every bit in target plaintext M_2^* with a time period i^*. $\mathcal{S}$ will set $H_1(ID_A^*) = bP$, $H_2(ID_A^*, i^*) = (h^*{}_{,i^*}P)$, $V^* = H_3(U^*, \xi^*) = H_3(cP, \xi^*)$. In the challenge phase, $\mathcal{S}$ returned a simulated ciphertext $C^* = (i^*, U^*, V^*)$, which implies the parameter ξ^* is defined as:

$$\begin{aligned}\xi^* =& \hat{e}(X_A, rH_1(ID_A^*))\hat{e}(P_{pub}, rH_1(ID_A^*))\\ & \hat{e}(P_{hk}, rH_2(ID_A^*, i^*))\\ =& \hat{e}(aP, bcP)\hat{e}(bP, cP)^s\hat{e}(wP, r(h^*{}_{,i^*}P))\\ =& \hat{e}(P,P)^{abc}\hat{e}(bP, cP)^s\hat{e}(wP, (h^*{}_{,i^*})cP)\end{aligned}$$

Above all, $\mathcal{S}$ can get the solution for BDH problem, i.e. $\hat{e}(P,P)^{abc} = \xi^*(\hat{e}(bP, cP)^{-s}\hat{e}(wP, (h^*{}_{,i^*})cP))^{-1}$. Thus we have proved the security of the scheme for the $TypeII$ adversary through this reduction.

□

4.2 Performance Comparison

We compare the major computational cost of our scheme with CL-PKE proposed by Al-Riyami and Paterson [1], CL-PKE1 and CL-PKE2 proposed by Cheng *et al.* [6] in Table 2. We assume both schemes are implemented on $|\ \mathbb{G}_1\ | = 160$ bits, $|\ \mathbb{G}_2\ | = 1024$ bits, $|\ p\ | = 160$ bits and hash value $= 160$ bits. We denote by M the point multiplication in $\mathbb{G}_1$, E the exponentiation in $\mathbb{G}_2$ and P the pairing computation. The other computations are trivial so we omitted them.

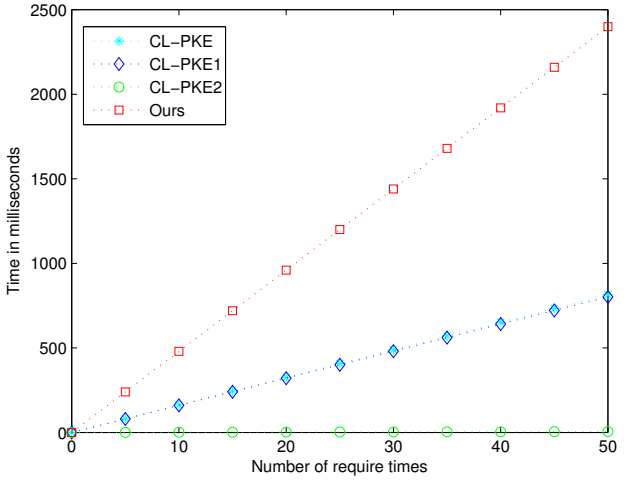


Figure 1: The comparison of the encryption computational cost

The simulation results are shown in the Figures 1, 2, 3. This experiment is executed on common desktop with 3.20GHz CPU Intel i5-4460 by using PBC Library [17]. Obviously, the ciphertext size and decryption computational cost of our scheme are comparable with the counterparts in the existing CL-PKE schemes. Indeed, our scheme is less efficient on execution time compared with CL-PKE in encryption phase, it is acceptable since the desirable key insulation function in our scheme as shown in Table 2. The additional composition of the private key in our scheme can be updated periodically, so our scheme provides extra security capability that can alleviate the problem of private key leakage. Therefore, this is a trade-off between efficiency and security capability.

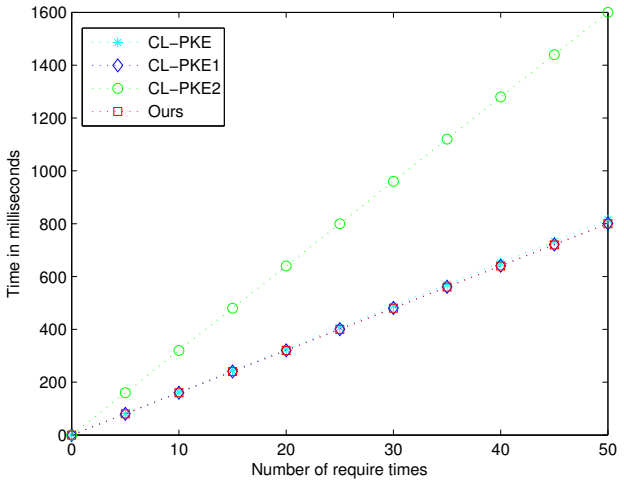


Figure 2: The comparison of the decryption computational cost

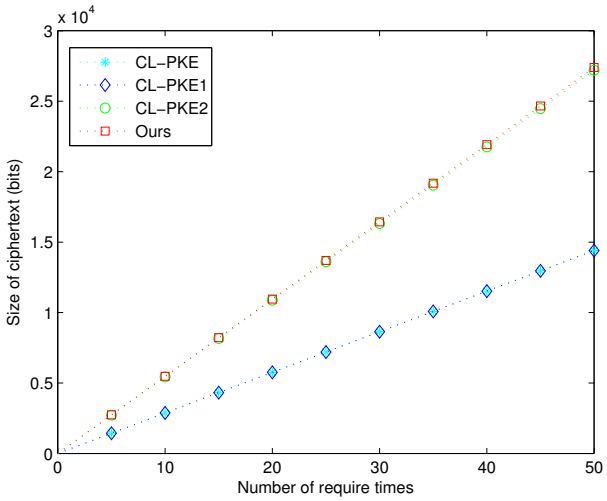


Figure 3: The comparison of the ciphertext size

5 Potential Applications

In view of the desirable merits, namely free from certificate authority and key escrow problem, and mitigating consequence of key exposure, the certificateless key insulated encryption system can be applied to a range of practical environments which are troubled by the private key exposure problem.

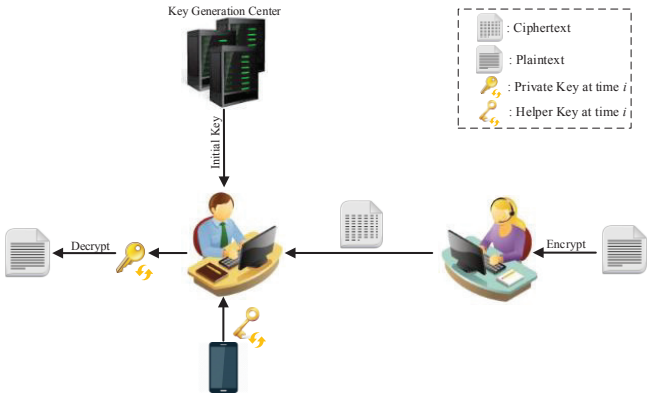


Figure 4: The system model of CL-KIE cryptosystem

In an office situation, for instance, it is common that an officer Bob might leave his seat without logging out email account. A malicious colleague could sneak into his seat and check his private mails from Alice easily even if these mails were encrypted. In contrast, the consequence of this case can be mitigated by adopting CL-KIE scheme as shown in Figure 4. To decrypt an encrypted mail, Bob should generate the latest private key at time period i with the help of Helper Key at time i that can only be produced by the helper device (his smart phone). Without

the helper, obviously, the malicious colleague can hardly decrypt these mails even if he had the access to Bob's desktop as the existing private key stored in Bob's computer has expired. Besides, other privacy-sensitive environments, such as cloud data sharing system and personal information system, are also the potentially applications for CL-KIE.

6 Conclusion

In this paper, we proposed the CL-KIE scheme by integrating the key-insulated security notion into the CL-PKE scheme in order to solve the private key exposure problem. We formalized the definition of CL-KIE scheme and proposed a concrete construction of the CL-KIE scheme. Moreover, the IND-CCA2 security proof of our scheme under BDH problem in the random oracle model was proposed. After that, we compared our scheme with three CL-PKE scheme on efficiency and security. Our scheme with key updated periodically can achieve key-escrow and key-exposure resilience which CL-PKE does not possess, while sacrifcing a little on the cost of computing time. Besides, we further extended the CL-KIE into the potential environments for the future practical application.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant 61003230, Grant 61370026, No. 61133016 and Grant 61272527. The authors gratefully acknowledge the anonymous reviewers for their valuable comments.

References

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

[2] R. Anderson, *Two Remarks on Public Key Cryptology*, 1997. (`http://www.cl.cam.ac.uk/users/rja14`)

[3] J. Baek, R. Safavi-Naini, and W. Susilo, "Certificateless public key encryption without pairing," in *Information Security*, pp. 134–148, 2005.

[4] S. K. Balakrishnan and V. P. J. Raj, "Practical implementation of a secure email system using certificateless cryptography and domain name system," *International Journal of Network Security*, vol. 18, no. 1, pp. 99–107, 2016.

[5] L. Benoit and J. J. Quisquater, "On constructing certificateless cryptosystems from identity based encryption," in *Public Key Cryptography (PKC'06)*, pp. 474–490, 2006.

[6] Z. Cheng, L. Chen, L. Ling, and R. Comley, "General and efficient certificateless public key encryption constructions," in *International Conference on Pairing-Based Cryptography*, pp. 83–107, 2007.

[7] B. Dan and F. Matt, "Identity-based encryption from the weil pairing," in *Advances in Cryptology (CRYPTO'01)*, pp. 213–229, 2001.

[8] A. W. Dent, B. Libert, and K. G. Paterson, "Certificateless encryption schemes strongly secure in the standard model," in *Public Key Cryptography (PKC'08)*, pp. 344–359, 2008.

[9] Y. Dodis, J. Katz, S. Xu, and M. Yung, "Key-insulated public key cryptosystems," in *International Conference on the Theory and Application of Cryptographic Techniques*, pp. 65–82, 2002.

[10] T. ElGamal, "A public key cryptosystem and a signature scheme based on discrete logarithms," in *Advances In Cryptology*, pp. 10–18, 1984.

[11] H. Goichiro, H. Yumiko, and I. Hideki, "Parallel key-insulated public key encryption," in *Public Key Cryptography (PKC'06)*, pp. 105–122, 2006.

[12] G. Itkis and L. Reyzin, "Sibir: Signer-base intrusion-resilient signatures," in *Advances in Cryptology (CRYPTO'02)*, pp. 101–116, 2002.

[13] C. Lin, Y. Li, K. Lv, and C. Chang, "Ciphertext-auditable identity-based encryption," *International Journal of Network Security*, vol. 17, no. 1, pp. 23–28, 2015.

[14] J. Liu, M. Au, and W. Susilo, "Self-generated-certificate public key cryptography and certificateless signature/encryption scheme in the standard model," in *Proceedings of the 2nd ACM Symposium on Information, Computer and Communications Security*, pp. 273–283, 2007.

[15] R. LRivest, S. Adi, and A. Len, "A method for obtaining digital signatures and public-key cryptosystems," *Communications of the ACM*, vol. 21, no. 2, pp. 120–126, 1978.

[16] B. Mihir and P. Adriana, "Protecting against key-exposure: Strongly key-insulated encryption with optimal threshold," *Applicable Algebra in Engineering, Communication and Computing*, vol. 16, no. 6, pp. 379–396, 2006.

[17] PBC Library, *The Pairing-based Cryptography Library*, Mar. 27, 2013. (`https://crypto.stanford.edu/pbc/`)

[18] W. Qiu, Y. Zhou, B. Zhu, Y. Zheng, M. Wen, and Z. Gong, "Key-insulated encryption based key predistribution scheme for wsn," in *Advances in Information Security and Assurance*, pp. 200–209, 2009.

[19] G. A. V. R. C. Rao, P. V. Lakshmi, and N. R. Shankar, "A new modular multiplication method in public key cryptosystem," *International Journal of Network Security*, vol. 15, no. 1, pp. 23–27, 2013.

[20] Y. Sun and H. Li, "Short-ciphertext and bdh-based cca2 secure certificateless encryption," *Science China Information Sciences*, vol. 53, no. 10, pp. 2005–2015, 2010.

[21] Y. Wang, X. Liu, L. Liang, W. Feng, and G. Yang, "Mitigating key escrow in attribute-based encryption," *International Journal of Network Security*, vol. 17, no. 1, pp. 94–102, Jan. 2015.

[22] Z. Wang and W. Chen, "An id-based online/offline signature scheme without random oracles for wireless sensor networks," *Personal and Ubiquitous Computing*, vol. 17, no. 5, pp. 837–841, 2013.

[23] W. Yang, F. Zhang, and L. Shen, "Efficient certificateless encryption with standing attacks from malicious kgc without using random oracles," *Security and Communication Networks*, vol. 7, no. 2, pp. 445–454, 2014.

[24] C. Youngblood, "An introduction to identity-based cryptography," *CSEP 590TU*, Mar. 2005.

[25] H. Yumiko, H. Goichiro, S. Junji, and I. Hideki, "Unconditionally secure key insulated cryptosystems: Models, bounds and constructions," in *Information and Communications Security*, pp. 85–96, 2002.

[26] X. Zhang and C. Xu, "A practical forward-secure public-key encryption scheme with untrusted update," *International Journal of Network Security*, vol. 17, no. 5, pp. 619–628, 2015.

Biography

Libo He is currently pursuing her Ph.D degree in the School of Information and Software Engineering, UESTC. Her research interests include cryptographic protocol and network security.

Chen Yuan received his B.S. degree in the School of Computer Science and Technology, Shandong University of Finance and Economics in Jun 2009. He is currently pursuing his M.S. degree in the School of Information and Software Engineering, UESTC. His research interests include cryptographic authentication protocols and network security.

Hu Xiong is an associate professor at University of Electronic Science and Technology of China (UESTC). He received his Ph.D degree from UESTC in 2009. His research interests include: cryptographic protocol, and network security.

Zhiguang Qin is the dean and professor in the School of Computer Science and Engineering, UESTC. He received his PH.D. degree from UESTC in 1996. His research interests include: information security and computer network.

Analysis of One Scheme for Enabling Cloud Storage Auditing with Verifiable Outsourcing of Key Updates

Zhengjun Cao[1], Lihua Liu[2], Olivier Markowitch[3]
(Corresponding author: Zhengjun Cao)
Department of Mathematics, Shanghai University[1]
No.99, Shangda Road, Shanghai, China
(Email: caozhj@shu.edu.cn)
Department of Mathematics, Shanghai Maritime University[2]
No.1550, Haigang Ave, Pudong New District, Shanghai, China
Computer Sciences Department, Université Libre de Bruxelles[3]
Boulevard du Triomphe CP 212, 1050 Bruxelles, Belgique
(Received July 25, 2016; revised and accepted Nov. 15 & Dec. 25, 2016)

Abstract

Cloud computing supports a paradigm shift from local to network-centric computing and enables customers with limited computational resources to outsource large-scale computational tasks to the cloud, such as linear equations and linear programming. Recently, Yu et al. [IEEE TIFS, 11(6), 2016, 1362-1375] have proposed a scheme for cloud storage auditing with verifiable outsourcing of key updates. In this note, we remark that Yu et al.'s scheme has two inherent weaknesses: 1) it does not truly mitigate the client's computational burden for key updates; 2) it does not ensure confidentiality since the files uploaded to the cloud by the client are eventually not encrypted at all.

Keywords: Cloud Computing; Cloud Storage Auditing; Confidentiality; Third-party Auditor

1 Introduction

Cloud computing making use of the tremendous resources of computing and storage systems via the Internet, supports a paradigm shift from local to network-centric computing [17, 19], and benefits scientific and engineering applications, such as data mining and many other computational and data-intensive activities. It enables customers with limited computational resources to outsource large-scale computational tasks to the cloud, including linear equations [21, 23], linear programming [9, 20, 22], bilinear pairing [1, 7, 8, 11], matrix inversion computation [14] and matrix multiplication computation [13]. Many researchers have studied the new computing paradigm and proposed a lot of schemes [2, 6, 10, 12, 15, 16]. But some kinds of flaws in some outsourcing schemes [3, 4, 5] were found for security, efficiency or other reasons.

Key-exposure problem is a special one related to key management. A primary observation on the problem in the scenario of cloud storage auditing is that once the client's secret key for storage auditing is exposed to the cloud, the cloud is able to easily hide the data loss incidents for maintaining its reputation. More seriously, the cloud maybe discard the client's data rarely accessed for saving the storage space [24].

Very recently, Yu et al. [25] have proposed a scheme for cloud storage auditing with verifiable outsourcing of key updates. The scheme involves three entities, the client, the cloud infrastructure and a third-party auditor (TPA). The client uploads his files to the cloud. The TPA audits the integrity of the files stored in the cloud and regularly updates the encrypted secret keys of the client to prevent any secret key exposition. Upon receiving the returned encrypted secret key, the client decrypts it to get the real secret key.

In this note we would like to stress that in Yu et al.'s scheme the recovered secret key is only used to generate authenticators for the files instead of protecting these files. No secret key associated to a symmetric key encryption is dedicated to the protection of the files. Besides, the scheme does not truly mitigate the client's computational burden for key updates. In view of these weaknesses, we would like to point out that Yu et al.'s scheme could not be practically implemented.

2 Review of Yu et al.'s Scheme

The scheme [25] uses the following notations. G_1, G_2 are two multiplicative groups with some prime order q. g is a generator of G_1. $\hat{e} : G_1 \times G_1 \to G_2$ is a bilinear pairing. $H_1 : G_1 \to G_1, H_2 : \{0,1\}^* \times G_1 \to \mathbb{Z}_q^*$ and $H_3 : \{0,1\}^* \times G_1 \to G_1$ are three cryptographic hash functions. T is the total periods number of the whole lifetime for the files stored in the cloud. $w_1 \cdots w_t$ is the binary string of the node w^j associated with period j. $w^j|_k (k \leq t)$ is the k-prefix of w^j. w^{j_0}, w^{j_1} are the left child node and the right child node of w^j, respectively. $w^j|_{\bar{k}}$ is the sibling node of $w^j|_k$. PK is the public key which is unchanged in the whole lifetime. ES_{w^j} is the encrypted node secret key. R_{w^j} is the verification value which is used to verify the validity of authenticators. ESK_j is the client's encrypted secret key in period j. X_j is the set composed by the key pairs. Ω_j is a set composed by the verification values. (ES, R) is the key pair of the root node. F is a file which the client wants to store in cloud. $m_i (i = 1, \cdots, n)$ are n blocks of file F. DK is the decryption key to recover the encrypted secret key for cloud storage auditing.

It involves three entities, some clients, a third-party auditor (TPA), and the cloud. The client has a secret key associated to a signature $SSig$ for ensuring the integrity of not only the file identifier *name* but also the time period j. The protocol consists of the following phases:

SysSetup: Given a security parameter k and the total time period T, the client picks a generator u of G_1, $\rho, \tau \in \mathbb{Z}_q^*$ and computes $R = g^\rho, G = g^\tau, ES = H_1(R)^{\rho-\tau}$. Set $PK = (R, G, u)$. Set $X_0 = \{(ES, R)\}, \Omega_0 = \emptyset$ (where $\emptyset$ is null set), $DK = \tau$ and keep it himself. The client sends $ESK_0 = (X_0, \Omega_0)$ to the TPA.

EkeyUpdate: Input ESK_j, the time period j, and the public key PK. The TPA parses $ESK_j = (X_i, \Omega_j)$ where X_j is organized as a stack which consists of (ES_{w^j}, R_{w^j}) and the key pairs of the right siblings of the nodes on the path from the root to w^j. The top element of the stack is (ES_{w^j}, R_{w^j}). Firstly, pop (ES_{w^j}, R_{w^j}) off the stack. Then do as follows:

- If w^j is an internal node ($w^{j+1} = w^{j_0}$ in this case), select $\rho_{w^{j_0}}, \rho_{w^{j_1}} \in \mathbb{Z}_q^*$. And then compute $R_{w^{j_0}} = g^{\rho_{w^{j_0}}}, R_{w^{j_1}} = g^{\rho_{w^{j_1}}}$, $h_{w^{j_0}} = H_2(w^{j_0}, R_{w^{j_0}}), h_{w^{j_1}} = H_2(w^{j_1}, R_{w^{j_1}})$, $ES_{w^{j_0}} = ES_{w^j} \cdot H_1(R)^{\rho_{w^{j_0}} h_{w^{j_0}}}$, $ES_{w^{j_1}} = ES_{w^j} \cdot H_1(R)^{\rho_{w^{j_1}} h_{w^{j_1}}}$. Push $(ES_{w^{j_1}}, R_{w^{j_1}})$ and $(ES_{w^{j_0}}, R_{w^{j_0}})$ onto the stack orderly. Let X_{j+1} denote the current stack and define $\Omega_{j+1} = \Omega_j \bigcup \{R_{w^{j_0}}\}$.
- If w^j is a leaf, define X_{j+1} with the current stack. If $w_t = 0$ (the node w_{j+1} is the right sibling node of w^j in this case), then set $\Omega_{j+1} = \Omega_j \bigcup \{R_{w^{j+1}}\} - \{R_{w^j}\}$ ($R_{w^{j+1}}$ can be read from the new top $(ES_{w^{j+1}}, R_{w^{j+1}})$ of the stack). If $w_t = 1$ ($w^{j+1} = w''1$ in this case, where w'' is the longest string such that $w''0$ is a prefix of w^j), then set $\Omega_{j+1} = \Omega_j \bigcup \{R_{w^{j+1}}\} - \{R_{w''0}, R_{w''01}, \cdots, R_{w_t}\}$ ($R_{w^{j+1}}$ can be read from the new top $(ES_{w^{j+1}}, R_{w^{j+1}})$ of the stack).
- Erase key pair (ES_{w^j}, R_{w^j}), and return $ESK_{j+1} = (X_{j+1}, \Omega_{j+1})$.

VerESK: Input a client's encrypted secret key $ESK_j = (X_j, \Omega_j)$, the current period j and the public key PK. The client checks that

$$\hat{e}(g, ES_{w^j}) = \hat{e}\left(R/G \cdot \prod_{m=1}^{t} R_{w^{j_1}}^{h_{w^{j_1}}}, H_1(R)\right),$$

where $h_{w^j} = H_2(w^j, R_{w^j})$.

DecESK: The client computes $S_{w^j} = ES_{w^j} \cdot H_1(R)^\tau$. The real secret key is set as $SK_j = (X'_j, \Omega_j)$, where X'_j is the same stack as X_j except that the top element in X'_j is (S_{w^j}, R_{w^j}) instead of (ES_{w^j}, R_{w^j}) in X_j.

AuthGen: For a file $F = \{m_1, \cdots, m_n\}$ and the current period j, the client proceeds as follows.

- Parse $SK_j = (X'_j, \Omega_j)$ and read the top element (S_{w^j}, R_{w^j}) from the stack X'_j. Select $r \in \mathbb{Z}_q^*$ and compute $U = g^r$,

$$\sigma_i = H_3(name \| i \| j, U)^r \cdot S_{w^j} \cdot u^{r m_i}$$

$(i = 1, \cdots, n)$, where the *name* is chosen randomly from $\mathbb{Z}_q^*$ as the identifier of the file F. Generate a file tag for F and j using the signature $SSig$ in order to ensure the integrity of *name* and j. Denote the set of authenticators in time period j with $\Phi = (j, U, \{\sigma_i\}_{1 \leq i \leq n}, \Omega_j)$.
- Send the file F and the set of authenticators along with the file tag to cloud.

ProofGen: Input a file F, a set of authenticators $\Phi = (j, U, \{\sigma_i\}_{1 \leq i \leq n}, \Omega_j)$, a time period j, a challenge $Chal = \{(i, v_i)\}_{i \in I}$ (where $I = \{s_1, \cdots, s_c\}$ is a c-element subset of set $[1, n]$ and $v_i \in \mathbb{Z}_q$) and the public key PK. The cloud calculates an aggregated authenticator $\Phi = (j, U, \sigma, \Omega_j)$, where $\sigma = \Pi_{i \in I} \sigma_i^{v_i}$. It also computes $\mu = \sum_{i \in I} v_i m_i$. It then sends $P = (j, U, \sigma, \mu, \Omega_j)$ along with the file tag as the response proof of storage correctness to the TPA.

ProofVerify: Input a proof P, a challenge $Chal$, a time period j and the public key PK. The TPA parses $\Omega_j = (R_{w^j|_1}, \cdots, R_{w^j|_t})$. He then verifies the integrity of *name* and j by checking the file tag. After that, the client verifies whether the following equation holds:

$$\begin{aligned} \hat{e}(g, \sigma) &= \hat{e}(R \cdot \Pi_{m=1}^{t} R_{w^j|_m}^{h_{w^j|_m}}, H_1(R)^{\sum_{i \in I} v_i}) \\ &\quad \cdot \hat{e}(U, u^\mu \cdot \Pi_{i \in I} H_3(name \| i \| j, U)^{v_i}), \end{aligned}$$

Table 1: Yu et al.'s outsourcing scheme

Client		TPA		Cloud
–*SysSetup*. Set $PK=(R,G,u)$, $X_0=\{(ES,R)\}, \Omega_0=\emptyset$. $DK=\tau, ESK_0=(X_0,\Omega_0)$.	$\xrightarrow{ESK_0}$ $\xleftarrow{ESK_1}$ $\vdots$ $\xleftarrow{ESK_j}$	–EkeyUpdate. Update it as $ESK_1=(X_1,\Omega_1)$.		
–*VerESK*. Check $\hat{e}(g,ES_{\omega^j}) \stackrel{?}{=} \hat{e}(R/G\cdot\prod_{m=1}^{t} R_{\omega_{j_1}}{}^{h_{\omega^{j_1}}}, H_1(R))$.				
–*DecESK*. Compute $S_{\omega^j}=ES_{\omega^j}\cdot H_1(R)^{\tau}$, $SK_j=(X'_j,\Omega_j)$.				
–*AuthGen*. For a file $F=\{m_1,\cdots,m_n\}$, compute $U=g^r$, $\sigma_i=H_3(name\|i\|j,U)^r\cdot S_{\omega^j}\cdot u^{rm_i}$ $i=1,\cdots,n$. Generate *tag* for F,j using the signature $SSig$.				
Set $\Phi=(j,U,\{\sigma_i\}_{1\le i\le n},\Omega_j)$.	– – –	upload $_{F,\ \Phi,\ tag}$ – – – – – – – – – – – – –	→	Store F, Φ, *tag*.
		–*Challenge*. Set $Chal=\{(i,v_i)\}_{i\in I}$.	$\xrightarrow{Chal}$	–*ProofGen*. $\sigma=\prod_{i\in I}\sigma_i^{v_i}$, $\mu=\sum_{i\in I} v_i m_i$.
		–*ProofVerify*. Verify the integrity of *name* and j by checking the *tag*. Check $\hat{e}(g,\sigma)\stackrel{?}{=}\hat{e}\left(R\cdot\prod_{m=1}^{t}R_{\omega^j\|m}^{h_{\omega^j\|m}}, H_1(R)^{\sum_{i\in I}v_i}\right)\cdot\hat{e}\left(U,u^{\mu}\cdot\prod_{i\in I}H_3(name\|i\|j,U)^{v_i}\right)$	$\xleftarrow{P}$	$P=\{j,U,\sigma,\mu,\Omega_j\}$.

where $h_{w^j}=H_2(w^j,R_{w^j})$. If it holds, returns "True", otherwise returns "False".

We refer to the following Table 1 for a brief description of Yu et al.'s scheme [25].

3 Analysis of Yu et al.'s Scheme

The scheme [25] aims to deal with the key exposure problem. They proposed the paradigm of cloud storage auditing which enables a client to outsource the burden of key updates to the third party auditor. But we find the scheme has two inherent flaws.

1) *The scheme does not truly mitigate the client's computational burden for key updates.* Concretely, in the time period j, the client's main computational task is to calculate

$$\begin{aligned} U &= g^r,\ S_{\omega^j}=ES_{\omega^j}\cdot H_1(R)^{\tau},\\ \sigma_i &= H_3(name\|i\|j,U)^r\cdot S_{\omega^j}\cdot u^{rm_i},\\ &\qquad i=1,\cdots,n\\ \hat{e}(g,ES_{\omega^j}) &\stackrel{?}{=} \hat{e}(R/G\cdot\prod_{m=1}^{t}R_{\omega_{j_1}}{}^{h_{\omega^{j_1}}},H_1(R)). \end{aligned}$$

while the TPA's main computational task is to calculate

$$R_{\omega^{j_0}}=g^{\rho_{\omega^{j_0}}},\ ES_{\omega^{j_0}}=ES_{\omega^j}\cdot H_1(R)^{\rho_{\omega^{j_0}}}h_{\omega^{j_0}},$$

$$R_{\omega^{j_1}}=g^{\rho_{\omega^{j_1}}},\ ES_{\omega^{j_1}}=ES_{\omega^j}\cdot H_1(R)^{\rho_{\omega^{j_1}}}h_{\omega^{j_1}},$$

$$\hat{e}(g,\sigma) \stackrel{?}{=} \hat{e}\left(R\cdot\prod_{m=1}^{t}R_{\omega^j|m}^{h_{\omega^j|m}},H_1(R)^{\sum_{i\in I}v_i}\right)\cdot\hat{e}\left(U,u^{\mu}\cdot\prod_{i\in I}H_3(name\|i\|j,U)^{v_i}\right).$$

It is easy to find that the computational task for the client is almost equal to that for the TPA. That means the client's computational burden is not truly alleviated.

2) *The scheme does not ensure any confidentiality since the files uploaded by the client to the cloud are in fact not encrypted at all.* In the phase *AuthGen*, it is specified that the client has to parse the file F as $\{m_1,\cdots,m_n\}$ and generate the authenticator Φ by invoking these m_i. The client then sends $\{F,\ \Phi,\ tag\}$ to the cloud. Clearly, the authors have forgotten to assign a symmetric key encryption to protect the file F. In practice, it is conventional to encrypt files using any symmetric key encryption, such as AES. Due to computing overhead [18], a public key encryption is usually just used to establish the secure channel needed to exchange the secret key of a symmetric-key system.

3) We would like to stress that if the uploaded file F is viewed as an encrypted file, then the scheme has to assign three groups of secret keys, SK_1 for encrypting the file F, SK_2 *for generating the authenticator* Φ *for* F, and SK_3 for signing *name* and the timestamp j (it specifies that "the client has held a secret key for a signature $SSig$, which is used to ensure the integrity not only of the file identifier *name* but also of the time period j"). Obviously, both SK_1 and SK_3 held by the client could also be exposed (as SK_2), but the authors [25] have not realized the main danger. The proposed method only solves the problem of updating SK_2. Therefore, *the proposed scheme does not solve correctly the client's key exposure problem.*

4) We think it seems impossible to revise the scheme such that it could simultaneously update the involved secret keys SK_1, SK_2 and SK_3, because the client has to retrieve the file F stored previously on the cloud in order to update the symmetric key SK_1. In such case, it is totally unnecessary to introduce the key SK_2 for generating the authenticator Φ for the file F. Frankly speaking, this is an inherent flaw in the proposed model by Yu et al.

4 Conclusion

We show that there are two flaws in Yu et al.'s scheme for cloud storage auditing with verifiable outsourcing of key updates, and remark that the scheme cannot be practically implemented. We would like to stress that the proposed paradigm is somewhat artificial because the client is able to update his secret keys solely by himself.

Acknowledgments

The work is supported by the National Natural Science Foundation of China (61303200, 61411146001), the Brussels Region (Innoviris) and the SeCloud Project. We are grateful to the reviewers for their valuable suggestions.

References

[1] S. Canard, J. Devigne, and O. Sanders, "Delegating a pairing can be both secure and efficient," in *Proceedings of Applied Cryptography and Network Security (ACNS'14)*, pp. 549–565, Lausanne, Switzerland, June 2014.

[2] S. Canard et al., "Toward generic method for server-aided cryptography," in *Proceedings of Information and Communications Security (ICICS'13)*, pp. 373–392, Beijing, China, November 2013.

[3] Z. J. Cao and L. H. Liu, "Comment on 'harnessing the cloud for securely outsourcing large-scale systems of linear equations'," *IEEE Transactions on Parallel Distributed Systems*, vol. 27, no. 5, pp. 1551–1552, 2016.

[4] Z. J. Cao and L. H. Liu, "A note on two schemes for secure outsourcing of linear programming," *International Journal of Network Security*, vol. 19, no. 2, pp. 323–326, 2017.

[5] Z. J. Cao, L. H. Liu, and O. Markowitch, "On two kinds of flaws in some server-aided verification schemes," *International Journal of Network Security*, vol. 18, no. 6, pp. 1054–1059, 2016.

[6] F. Chen, T. Xiang, and Y. Y. Yang, "Privacy-preserving and verifiable protocols for scientific computation outsourcing to the cloud," *Journal of Parallel Distributed Computing*, vol. 74, pp. 2141–2151, 2014.

[7] X. F. Chen et al., "Efficient algorithms for secure outsourcing of bilinear pairings," *Theoretical Computer Science*, no. 562, pp. 112–121, 2015.

[8] B. Chevallier-Mames et al., "Secure delegation of elliptic-curve pairing," in *Proceedings of Smart Card Research and Advanced Application, 9th IFIP WG 8.8/11.2 International Conference (CARDIS'10)*, pp. 24–35, Passau, Germany, Apr. 2010.

[9] J. Dreier and F. Kerschbaum, "Practical privacy-preserving multiparty linear programming based on problem transformation," in *Proceedings of IEEE International Conference on Privacy, Security, Risk, and Trust, and IEEE International Conference on Social Computing (PASSAT/SocialCom'11)*, pp. 916–924, Boston, MA, USA, Oct. 2011.

[10] M. Girault and D. Lefranc, "Server-aided verification: Theory and practice," in *Proceedings of Advances in Cryptology (ASIACRYPT'05)*, pp. 605–623, Chennai, India, Dec. 2005.

[11] S. Hohenberger and A. Lysyanskaya, "How to securely outsource cryptographic computations," in

Proceedings of Theory of Cryptography (TCC'05), pp. 264–282, Cambridge, MA, USA, Feb. 2005.

[12] W. F. Hsien, C. 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.

[13] X. Y. Lei, X. F. Liao, T. W. Huang, and F. Heriniaina, "Achieving security, robust cheating resistance, and high-efficiency for outsourcing large matrix multiplication computation to a malicious cloud," *Information Sciences*, vol. 280, pp. 205–217, 2014.

[14] X. Y. Lei, X. F. Liao, T. W. Huang, H. Q. Li, and C. Q. Hu, "Outsourcing large matrix inversion computation to a public cloud," *IEEE Transactions on Cloud Computing*, vol. 1, no. 1, pp. 78–87, 2013.

[15] C. W. Liu, W. F. Hsien, C. 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.

[16] J. Liu, C. K. Chu, and J. Y. Zhou, "Identity-based server-aided decryption," in *Proceedings of Information Security and Privacy (ACISP'11)*, pp. 337–352, Melbourne, Australia, July 2011.

[17] D. Marinescu, *Cloud Computing Theory and Practice*. USA: Elsevier, 2013.

[18] A. Menezes, P. van Oorschot, and S. Vanstone, *Handbook of Applied Cryptography*, New York, U.S.A.: CRC,Taylor & Francis, 1996.

[19] 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.

[20] H. X. Nie, X. F. Chen, J. Li, J. Liu, and W. J. Lou, "Efficient and verifiable algorithm for secure outsourcing of large-scale linear programming," in *Proceedings of 28th IEEE International Conference on Advanced Information Networking and Applications (AINA'14)*, pp. 591–596, Victoria, BC, Canada, May 2014.

[21] S. Salinas, C. Q. Luo, X. H. Chen, and P. Li, "Efficient secure outsourcing of large-scale linear systems of equations," in *Proceedings of 2015 IEEE Conference on Computer Communications (INFOCOM'15)*, pp. 1035–1043, Hong Kong, China, Apr. 2015.

[22] C. Wang, K. Ren, and J. Wang, "Secure optimization computation outsourcing in cloud computing: A case study of linear programming," *IEEE Transactions on Computers*, vol. 65, no. 1, pp. 216–229, 2016.

[23] C. Wang, K. Ren, J. Wang, and Q. Wang, "Harnessing the cloud for securely outsourcing large-scale systems of linear equations," *IEEE Transactions on Parallel Distributed Systems*, vol. 24, no. 6, pp. 1172–1181, 2013.

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

[25] J. Yu, K. Ren, and C. Wang, "Enabling cloud storage auditing with verifiable outsourcing of key updates," *IEEE Transactions on Information Forensics Security*, vol. 11, no. 6, pp. 1362–1375, 2016.

Biography

Zhengjun Cao is an associate professor with the Department of Mathematics at Shanghai University. He received his Ph.D. degree in applied mathematics from Academy of Mathematics and Systems Science, Chinese Academy of Sciences. His research interests include applied cryptography, discrete algorithms and quantum computation.

Lihua Liu is an associate professor with the Department of Mathematics at Shanghai Maritime University. She received her Ph.D. degree in applied mathematics from Shanghai Jiao Tong University. Her research interests include combinatorics and cryptography.

Olivier Markowitch is an associate professor with the Computer Sciences Department at the Universite Libre de Bruxelles. He is also information security advisor of his University. He is working on the design and analysis of two-party and multi-party cryptographic protocols as well as on the design and analysis of digital signature schemes.

Towards an Optimum Authentication Service Allocation and Availability in VANETs

Safi Ibrahim[1], Mohamed Hamdy[1], and Eman Shaaban[2]
(Corresponding author: Mohamed Hamdy)
Information Systems Department, Ain Shams University[1]
Computer Systems Department, Ain Shams University[2]
Al Khalifa Al Maamoun st., Abbassia , 11566 Cairo, Egypt
(Email: m.hamdy@cis.asu.edu.eg)
(Received May 28, 2016; revised and accepted Aug. 9 & Sept. 5, 2016)

Abstract

Authentication as a security key issue is required for securing the inter-vehicle communication. Mostly, authentication schemes that depend on the network infrastructure Road Side Unit (RSU) had been proven to have low computation and communication overhead. RSUs may become unavailable due to congestion or failure conditions. Replicating authentication service offered by RSUs to trusted vehicles in their ranges may present novel alternative to support the availability of such services. A complement to our previous replication protocol Central Push-based Replication Protocol (CPRP) is presented in this work. Two new replica allocation techniques are proposed to spawn new replicas in order to improve the authentication service availability that are offered from RSUs. The optimality (Correctness) of these techniques has been evaluated with different ratios of RSU failures in different realistic scenarios. The results showed that both techniques improve CPRP in increasing the authentication services availability.

Keywords: Allocation; Authentication; Availability; Road Side Unit; VANET

1 Introduction

Exchanging life critical messages in VANETs requires authentication. Many security researches consider authentication as the most important security issue for exchanging such messages. Authentication is needed to prove that the sender is the actual owner of the message and to avoid impersonation attack. Some Studies show that authentication services offered by Road Side units (RSU) have less computation and communications calculations than others [20, 21]. RSUs are located at certain positions on the road network similar to access points in traditional wireless networks to provide the necessary infrastructure support for network setup and communications.

RSUs may become unavailable due to congestion, physical damage or because of their high deployment and maintenance cost. Unavailability of RSUs will cause absence of vehicles communications at those areas. Replicating authentication services from RSUs to vehicles may offers an alternative to support such service availability when they are unavailable.

Replication is a classical approach for increasing service or data availability in wired or wireless networks. Pushing Authentication service from RSUs to vehicles may raises two questions : when to replicate the service and how to select the vehicle which will hold the replica.

Replication cost can be defined as the number of active replicas in the network at certain time. However, replica Allocation correctness measures how the distribution of replicas is optimum.

This work is a complement of our work that was published in [9]. The proposed Central push-based Replica Protocol (CPRP) was introduced to increase the availability of authentication service offered from the RSUs.

CPRP has three basic mechanisms: replica allocation, replica activation and replica deallocation. Its basic idea is to push the authentication service by RSU to replicate it to the most central vehicle in its communication range which is the nearest to the RSU. This replica still inactive until it finds that no RSU in its range. It becomes active replica and replaces the RSU in authenticating vehicles in this area.

As mentioned, the communication range of vehicles is quarter that of RSUs, thus vehicle can't cover as many vehicles when it becomes active. Another replicas should be spawned to assist the replica in authenticating vehicles.

In this work, two methods are proposed for generating new replicas. The first is the Epidemic-based scheme which doesn't depend on network analysis. It generates new replicas by pushing replicas epidemically. The second is topology-based scheme which depends on analyzing the network and compute node degrees, then push the subreplicas to vehicles with high degrees. In general,

the topology analysis exposes the network to more security threats. The service provider vehicle should reveal the real identity for all analyzed vehicles to avoid malicious ones and to prevent impersonation attacks. Both methods are explained and evaluated in the next sections. Both techniques are compared in terms of enhancement the performance of CPRP on increasing the availability of the authentication service offered from RSU. They are also compared in terms of their replication degree and correctness of replica allocation.

The rest of the paper is organized as follows: Section 2 presents the literature review. In Section 3, basic assumptions for the two proposed schemes: Epidemic-based and Topology-based are stated. Section 4 introduces the proposed epidemic based scheme. The topology-based scheme is presented in Section 5. Simulation Configuration is introduced in Section 6. Section 7 presents comparison evaluation for the protocols in terms of authentication service availability improvement, Increase in replication degree and replica allocation correctness. Section 8 presents conclusion of work.

2 Related Work

In this section the literature review is introduced in terms of three points. On the one hand, authentication has an increasing interest from research in VANETs. It represents the key function of communication in any network. On the other hand, recently, researches on increasing the availability of network functionalities provide promising opportunities to achieve better security function. Data and service replication have been introduced in this work as an important mean for preserving the availability. Finally, Optimality (Correctness) of replica allocation schemes are introduced.

2.1 Authentication

Public Key Infrastructure (PKI) scheme is implemented in [16, 17, 19]. This scheme requires that each vehicle should be preloaded with a large number of public and private key pairs and the corresponding public key certificates. PKI requires large storage space in vehicles. Authority takes long time for tracking misbehaving vehicles due to long revocation list. Updating certificate revocation lists in vehicles consumes long time.

Many researches had tackled how to overcome the problems triggered by using PKI. In [21], (RAISE) rsu-aided message authentication scheme is proposed. RAISE is a symmetric key authentication scheme. In RAISE, RSUs assist vehicles in authenticating messages. Each message is attached with a short keyed-hash message authentication code (HMAC) which is generated by the vehicle, and the RSU in the range. RSU sends notice of authenticity to each vehicle. With the short HMAC attached to the message, the verification of message authenticity can be performed in a fast and efficient way. Although this technique outperforms PKI, it is not scalable. If two vehicles are not in the same range of the same RSU they can't communicate. Scalability problem was solved in [20]. Vehicle generates symmetric secret key with the first RSU it pass by. Then it uses this symmetric key to generate session keys with RSUs that are controlled by the same CA.

In [8], an anonymous ring signature scheme is introduced. This scheme outperforms the above schemes in which it offers low storage requirements and fast message authentication. It also doesn't depend on RSUs in authenticating vehicles.

In [3], they proposed an efficient message authentication scheme which is not vulnerable to impersonation attack based on elliptic curve cryptography.

2.2 Data and Service Replication

There are several studies that address service replication. Service replication is classified into two major classes interms of spawning new replicas. The first class doesn't depend on network topology analysis to make the replication decisions. This type is network transparent. Replication decisions occur at application layers and no information is required from lower network layers. [4] and [12] have largely been based on schemes that epidemically push the service on all available nodes. Using all nodes as a service holder is wasteful and unnecessary.

In [5], (SDP) Service Distribution Protocol For MANET is proposed. In SDP, the replication decision is based on service popularity which can be gained from client interest in the services. The service is replicated by the client with the highest interest in the service. This approach can achieve high service availability and correct service distribution.

Second class requires network topology. In [11] RegRes (Region Resident Service approach) is proposed. Each service determine its desired service carriers density within region. The RegRes runs on the carries to estimate the current density of service carriers. Then it applies a spawn policies to decide when and which node to spawn as new carriers. It account for variable node density, variable node mobility, replication cost and carriers that fail or leave the region.

In [1] V-PADA Vehicle Platoon Aware Data Access, a service replication solution. The concept lying behind this approach, is that vehicles move in platoons and follow the leader of the platoon. A vehicle which has a service to share with other vehicles in the same platoon, can predict platoon splits. If a vehicle leaves a platoon, it transfers its services or data to other vehicles to be able to access it. Each node has four states to be transferred between them, initials, Join, Quasi split and split.

In [2], Scalable data lookup and replication protocol for MANET (SCALAR) is proposed. SCALAR depends on constructing a connected dominating set based on a network graph. This set forms a virtual backbone upon which data or service replication takes place. SCALAR

had solved the scalability and data accessibility of services and data in large networks. However, unneeded replicas may ba generated. SCALAR overloads the network with the dominating set computational overhead and recovery.

2.3 Optimality of Replica Allocation

Service replication protocols overload the network with additional computation overhead.

The optimality or correctness of replica allocation can reflect the optimal service distribution all over the network. In [5], they proposed a measure of correctness to mobile ad hoc network (MANET), which is a relation between the number of available active replicas inside a given partition and its size. The two proposed methods are *Linear Correctness Ratio* and *Rational Correctness Ratio*. They assumed that the correctness ratio is bounded between 0% and 100%. In Linear Correctness, If the partition has no replica, the ratio of correctness will be 0%. Else, if there is one or two replicas in the same partition, correctness ratio will be 100%. Otherwise, it is linearly inversely proportional to the number of replicas.

However, in rational correctness ratio; it is more sensitive to the number of active replicas inside a partition.

3 Basic Assumptions

In this section, some basic assumptions are stated as follows:

- As an extension to our previous work [9], the concrete authentication scheme that is applied in this work is Symmetric key Scheme of VANETs, because of its low computation and communication overhead compared to Public Key Infrastructure [21].
- Only the replica that is pushed by the RSU has the privilege to spawn new replicas and is termed Replica.
- Replicas spawned from the vehicle replica are considered Followers and haven't the privilege to spawn new replicas.
- Communication range of the replica is quarter than that of the RSU.
- If one of followers loose the connection of the Replica because it is out of its range it become invalid item If the Replica become inactive or hibernated because it enters RSU range, it hibernates all its Followers.

4 Epidemic Based Scheme

Epidemic based scheme is one of the proposed techniques to improve the performance of CPRP by epidemically spawn new replicas.

Figure 1 explains the mechanism of generating new replicas using this scheme. After the replica that is pushed by RSU becomes active because it is in uncovered area. It pushes the replica epidemically to the two farthest vehicles with in its communication range. one is in its movement direction. The other, is in the opposite movement direction. The two generated replicas becomes followers to the main replica.

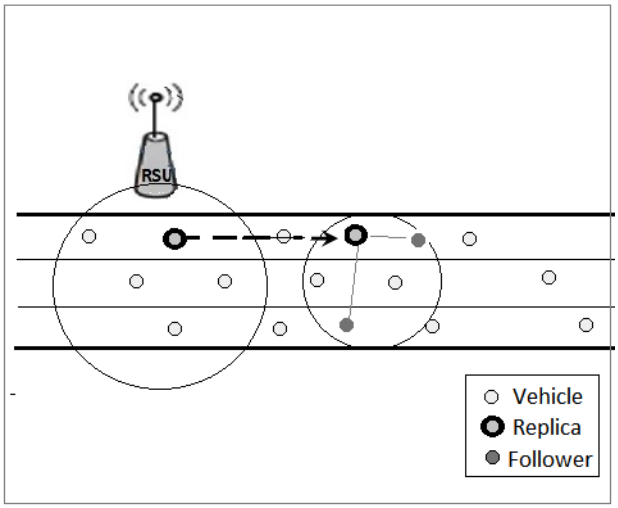


Figure 1: Spawning new replicas by epidemic-based approach

At each second, the main replica checks if it has two replicas within its communication range. If it looses one or both, it reassigns new replicas.

Algorithm 1, summarizes the process of the epidemic based scheme as follows:

Algorithm 1 Enhance replica allocation protocol by epidemically analysis

```
1: Begin
2: for EACH SECOND do
3:    for all Active Vrep do
4:       if Vrep doesn't find Followers in its range then
5:          Vrep calculates Dist with its neighbors
6:          Vrep chooses the two farthest neighbors to
            push SUB-REPLICA
7:          The first in its direction
8:          The other in opposite direction
9:          Vrep assign SUB-REPs as followers replica to
            it
10:      end if
11:   end for
12: end for
```

5 Topology Based Scheme

Topology based scheme is the second proposed scheme to enhance CPRP performance. This scheme depends on topology analysis of the network. After the replica that is pushed by RSU becomes active because it is in uncovered area.

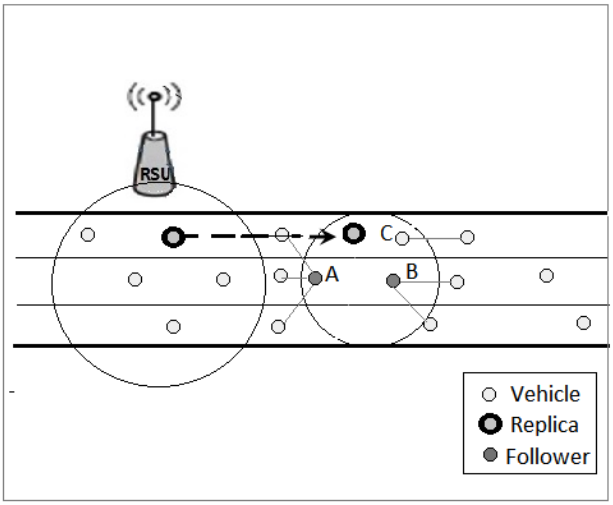


Figure 2: Spawning new replicas by topology analysis based approach

It sends inquiry to its neighbors about their degrees. Then, calculates the average degree received from all neighbors. Finally, it pushes its replica to nodes which have degrees more than the average degree.

Figure 2 explains the toplogy based scheme as follows: when the replica is out of the RSU range it becomes active. It sends to its neighbors vehicles inquiry about their neighbor number. Vehicle A replied with three neighbors, vehicle B replied with two neighbors and vehicle C replied with one neighbor. Replica vehicle calculate the average number of vehicles as $(3+2+1)/3 = 2$. Then, it makes decision to choose vehicles A and B which have two or more neighbors to push its replica.

Algorithm 2 summarizes the mechanism of the topology-based scheme as follows:

Algorithm 2 Enhance replica allocation by topology analysis

1: Begin
2: **for** EACH SECOND **do**
3: **for all** Active $Vrep$ **do**
4: $Vrep$ sends inquery to its neighbors about their DEG
5: $Vrep$ calculate TOT-DEG from all neighbors
6: $Vrep$ calculate AVG-DEG
7: $Vrep$ push SUB-REP to neighbor V which has DEG greater than AVG-DEG
8: $Vrep$ assign SUB-REPs as followers replica to it
9: **end for**
10: **end for**

6 Simulation Configuration

In simulating this work, we use SUMO [18][1] (Simulation of Urban Mobility) to model the movement of vehicles. For network simulation we use OPNET [13, 15][2]. The vehicles movement are generated using car following model that explained in details in [14].

In this simulation we use a frame work proposed in [10] to simulate VANETs SUMO with OPNET. With SUMO we generate network and route files, then simulate all vehicles positions at each second. These positions are written in a dump file. We use the trace exporter for OPNET which is implemented and explained to generate one xml topology file with trajectory for each vehicle.To generate vehicles movement we use randomTrip.py module which assigns a trip for each vehicle randomly.

For simplicity we use bidirectional highway with length 5km, we used 400m as the RSU range and 100m as the vehicles range. RSUs are placed at 350m distance between each other. The average number of simulation's run is three. Table 1 summarizes the simulation configurations.

Table 1: Table of configuration parameters

Parameter	Value
Way	bidirectional highway
No. of lanes	two
Way Length	5 KM
RSU range	400 m
Vehicle range	100 m
RSU separated dis.	350 m
Hand off area	50 m
Mobility model	CAR FOLLOWING MODEL
Vehicle movements	random trips

7 Evaluation and Discussion

In this section, evaluation and comparison of the proposed techniques: epidemic based and topology based technique are done in terms of availability improvement, replication degree and replica allocation correctness.

As a complement of our previous work, simulation is done with the same configuration. 24 scenarios are generated for the simulation and are divided into two groups. We assume four RSUs failure ratios at: 30%, 50%, 70%, 90%. For the first group we have 12 scenarios generated as follows: For each RSU failure ratio, three scenarios are generated; with high network density, moderate network density and low network density. The same has done for the second group. For each RSU failure ratio, also three scenarios are generated but with high vehicles speed, moderate vehicles speed and low vehicles speed.

[1] A microscopic traffic vehicle simulator `http://sourceforge.net/projects/sumo/`

[2] Version 17.1, licensed to NTI (National Telecommunication Institute)

7.1 Availability Improvement

Authentication service availability is computed by accumulating and averaging the availability of all vehicles during the network life time.

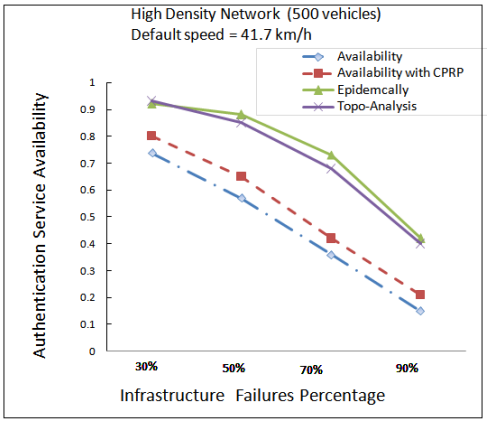
$$AuthAvail = \frac{1}{N}\sum_{i=1}^{N} A(v_i) \quad (1)$$

where $AuthAvail$ is the total Authentication Service Availability, N is the total number of vehicles, $A(v_i)$ is the vehicle i availability and is measured as follows:

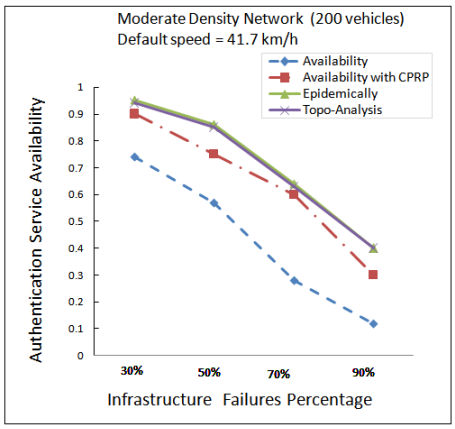
$$A(v_i) = \begin{cases} 1 & \mathrm{v}_i \in RSU_i or Replica \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

First, when applying these techniques using the different network densities shown in Figure 3. We can observe that they converge and almost have the same performance in the three different densities. This is because both techniques spawn new replicas to improve the authentication service availability. The epidemic-based generates replicas surround the original, while topology-based selects the suitable vehicle by topology analysis. They improve the performance of CPRP in high density with highest difference about 2%. It is also observed that in high density network, the epidemic method slightly outperforms the topology based.

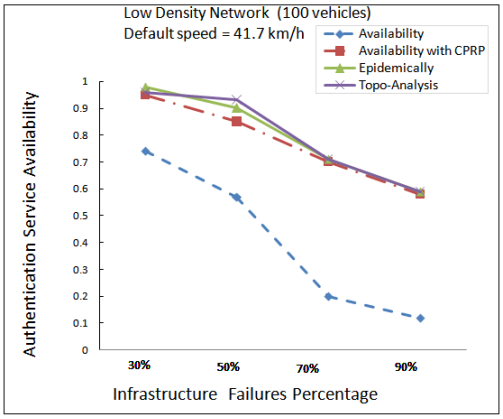
Next, when applying both techniques on different speed networks as shown in Figure 4. It can be observed that they converge also in their performance to be almost the same. They improve the performance of CPRP in low speed networks with highest difference about 1.3%. It is observed that the epidemic based slightly outperforms the topology based in moderate speed. However, in low speed density the topology based outperforms epidemical based.



(a) High Density



(b) Moderate Density



(c) Low Density

Figure 3: Authentication service availability improvement using two techniques vs. different network densities

7.2 Increase in Replication Degree

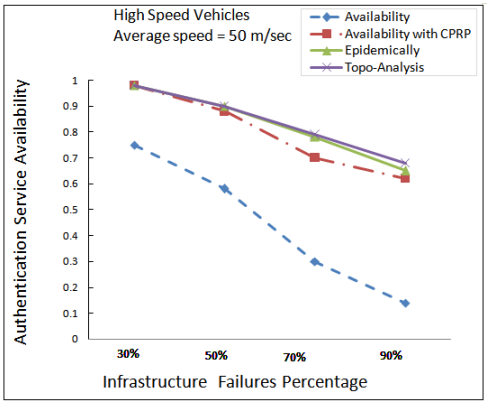
In this subsection, the results of using the two proposed techniques (Epidemic & Topology) to improve the performance of CPRP on increasing the replication degree during the network lifetime are displayed.

$$ReplicationDegree\% = \frac{1}{N}\sum_{i=1}^{N} V_i rep \cdot 100 \quad (3)$$

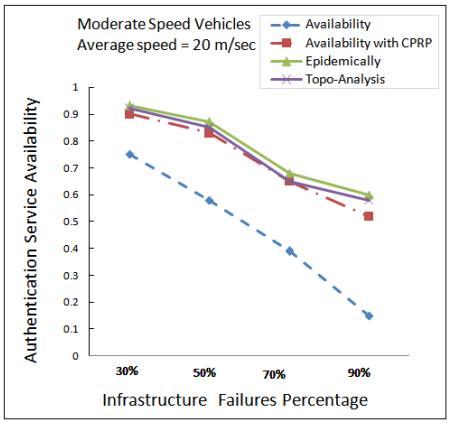
where N is the total number of vehicles and $V_i rep$ represents vehicle i that holds a replica.

Figures 5, 6 shows these effects on different network densities and different network speeds respectively. It is observed that applying both techniques begin with a value at 30% infrastructure failure then increases at 50% of infrastructure failure and finally decrease at 70% and 90%. The explanation of this is at low infrastructure failures spawning new replicas is low because the infrastructure existence hibernates the active replicas and prevents spawning new ones. From 70% infrastructure failure the situation changed because the main source to generate original replicas is the infrastructure. These replicas then can spawn new replicas.

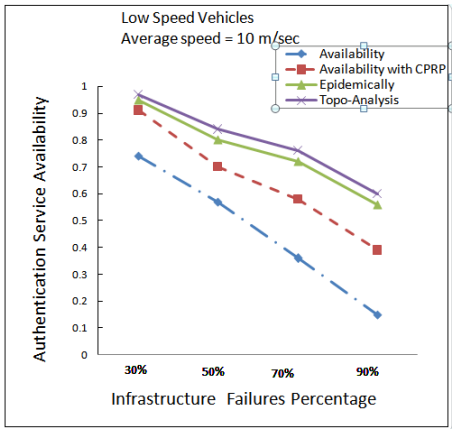
First, for different network densities as shown in Figure 5. It is observed that applying both techniques, add



(a) High Speed



(b) Moderate Speed



(c) Low Speed

Figure 4: Authentication service availability improvement using two techniques vs. different vehicles speeds

additional replication efforts which cause additional waste of resources. Topology based technique has better performance in all scenarios and almost the same but with few decrease in high density network. The average replication degree increases after applying topology based by about 1% in high density network, 1.5% in moderate density network and by 1.75% in low density networks. On the other hand, applying epidemic based adds extra additional waste of resources in all densities scenarios. Average replication degree increases after applying epidemic by about 3% in high density network, 4.6% in moderate density network and by 7.6% in low density network. So, both topology based and epidemic based have low replication degree in high density network then it increases by decreasing the density. The interpretation is due to definition of replication degree which is the percentage of number of replicas over total nodes. So, the replicas ratio to total total nodes increases by decreasing the network density. Topology based outperforms epidemic based by 2% in high density, 3.1% in moderate density and by 5.85% in low density.

Second, for different network speed as shown in Figure 6. Also, applying both techniques add additional replication efforts which cause additional waste of recourse. Topology based outperforms the epidemic based in all scenarios and almost the same in all scenarios with few improvements in low speed networks. The average replication degree increases after applying topology based by about 1.3% in low speed network, 1.7% in moderate speed network and by 2.3% in high speed networks. On the other hand, applying epidemic based adds extra additional waste of resources in different speeds scenarios. Average replication degree increases after applying epidemic by about 2.5% in low speed network, 3.7% in moderate speed network and by 7.6% in high speed networks. So, both topology based and epidemic based have low replication degree in low speed network then it increases by increasing the speed. The interpretation of this is due to the increase of vehicles speeds allows for increasing of spawning new replicas when replicas reached an area not covered with infrastructure. So, the replicas ratio to total total nodes increases by increasing the network speed. Topology based outperforms epidemic based by 1.2% in low speed, 2% in moderate speed and by 5.3% in low speed.

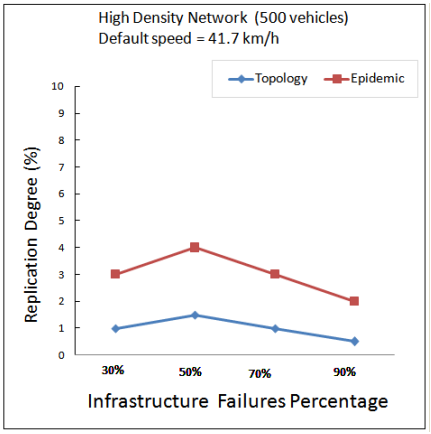
7.3 Replica Allocation Correctness

The replica allocation correctness means how the distribution of the generated replicas is optimal. In [6, 7], they proposed four different allocation correctness methods. In this work we propose a new replica allocation correctness for the optimal distribution of services in VANET.

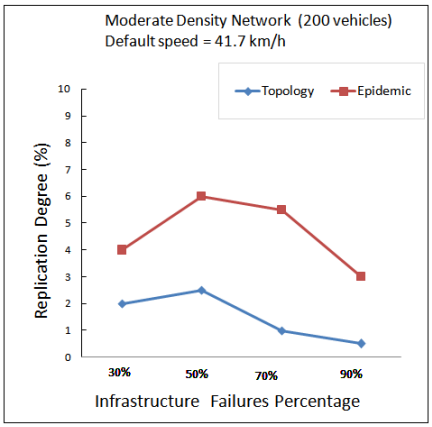
Replica Allocation Correctness for the whole network is computed by accumulating and averaging the correctness values for all vehicles During the network lifetime as shown in Equation (4).

$$RAC = \frac{1}{N}\sum_{i=1}^{N} C_v(v_i) \tag{4}$$

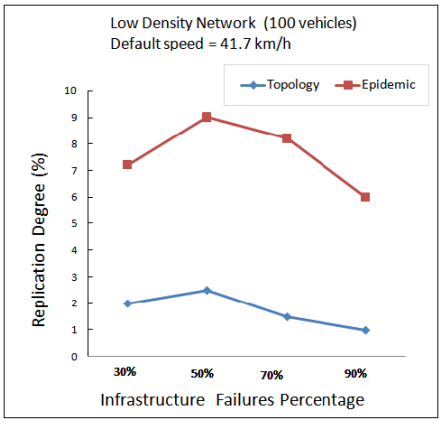
where RAC is the Total replica Allocation for the whole network, N is the total number of vehicles, $C_v(v_i)$ is the



(a) High Density

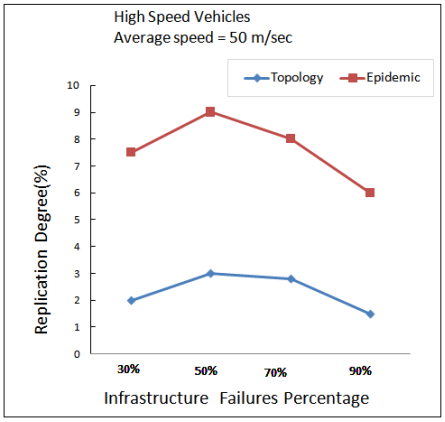


(b) Moderate Density

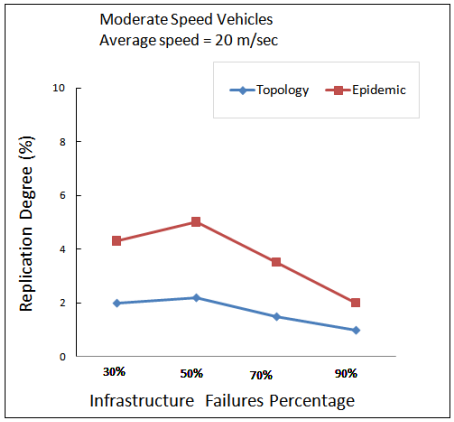


(c) Low Density

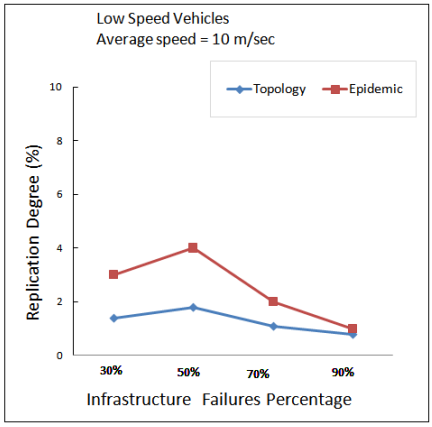
Figure 5: Increase of replicas percentage vs. different network densities



(a) High Speed



(b) Moderate Speed



(c) Low Speed

Figure 6: Increase of replicas percentage vs. different vehicles speeds

vehicle's i correctness value and is measured as illustrated in Equation (5):

$$C_v(v_i) = \begin{cases} 0 & r = 0 \\ 1 & r = 1 \\ \frac{n-r}{n-1} & r > 1 \end{cases} \quad (5)$$

where r represents the sum of replicas that exist in the range of the vehicle v_i, and n represents the number of neighbors of the vehicle v_i.

The explanation of how to compute replica correctness value at each vehicle is as follows: if the vehicle is not in range with any RSU and have no replica in its range its correctness will be zero. But, if it has 1 replica in its range, its correctness is one. Otherwise, it is computed related to its neighbor as shown in the equation. This ratio adapts to congestion conditions. Figure 7 illustrates an example of VANET in ad hoc mode in the area that

lack infrastructure. Gray nodes represent replicas. By using the Equation (5), vehicle A has no replica in its range its correctness is zero.Vehicle C has replicas on all of its neighbors, so its correctness is zero. Vehicle E has 1 replica so, its correctness is 1. Vehicles that have more than one replica in its range their correctness is computed related to their number of neighbors. If the number of neighbors is large it has greater correctness. Vehicle B has 2 replicas and 3 neighbors so its correctness is 1/2. Vehicle D has 5 neighbors, and has 2 replicas. Its correctness value is 3/4.

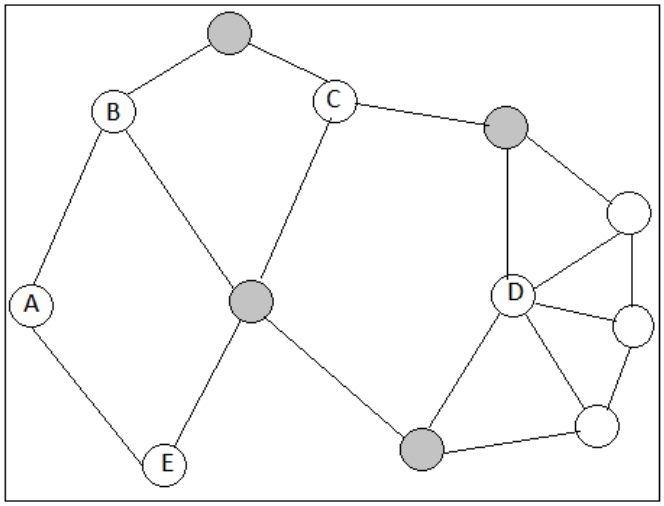
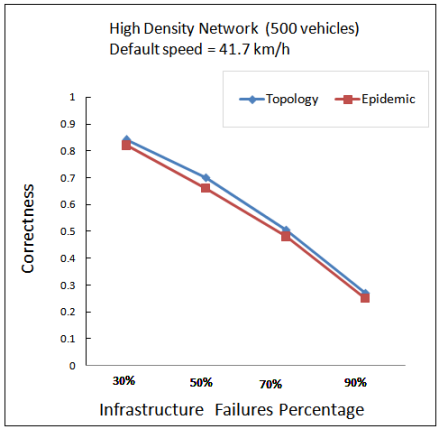


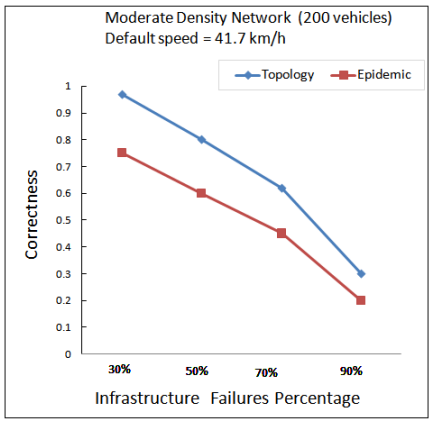
Figure 7: Example of replica allocation correctness

Figure 8 illustrates the replica allocation correctness values for both techniques using different network densities and different infrastructure failure ratios. It is observed that topology based schemes outperforms the epidemic based in almost all scenarios. Topology based scheme has the best performance with the moderate network density with average correctness about 67% for all infrastructure failure ratios. Then it achieves an average of 61% for low density network, and about 58% in High density networks. The interpretation of this is due to the topology scheme which depends on the analysis of network. When it is applied for high density networks there still vehicles not covered by infrastructure or one of the replicas especially with high infrastructure failure ratios. On the other hand when it is applied with low network densities; there may be additional unused replicas that decrease the correctness of the scheme. So, the optimum or best value gained from applying this scheme on different network densities is with the medium density network.

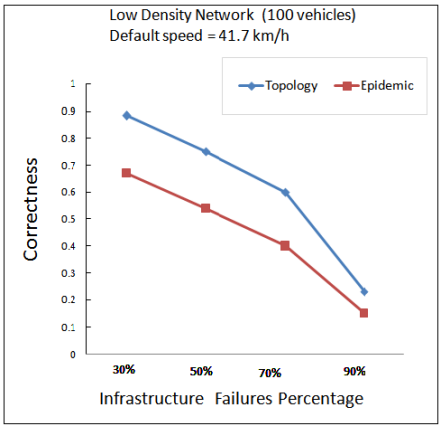
When the epidemic based scheme is applied with different network densities and with different infrastructure failure ratios, it shows the best average with high density network with about 55%. Then about 50% average correctness for medium density network, and 44% average correctness value for low density networks. This is interpreted as the epidemic scheme may generate additional useless replicas as density of the network decreases. So, it has its best correctness value with High Density Networks.



(a) High Density



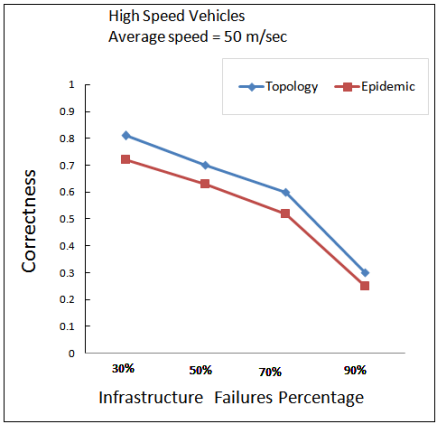
(b) Moderate Density



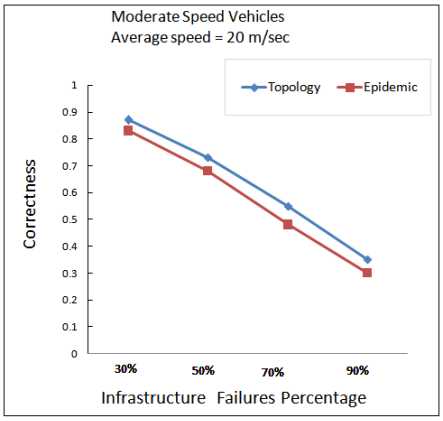
(c) Low Density

Figure 8: Correctness of replica allocation VS. different network densities

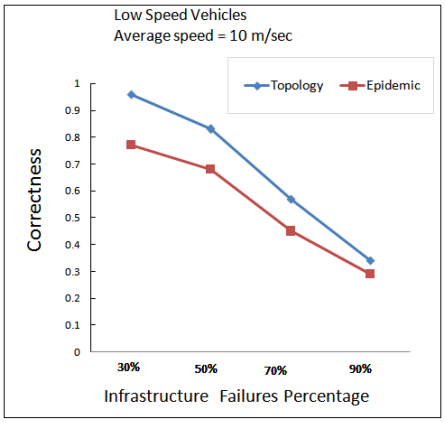
Second, Figure 9 shows the replica allocation correctness values for both techniques using different network



(a) High Speed



(b) Moderate Speed



(c) Low Speed

Figure 9: Correctness of replica allocation VS. different Vehicles speeds

speeds and different infrastructure failure ratios. It is observed that topology based schemes outperforms the epidemic based in almost all scenarios. Topology based scheme has the best performance with the low speed network with average correctness about 67% for all infrastructure failure ratios. Then it achieves an average of 62% for moderate speed network, and about 60% in High speed networks. The interpretation of this is due to the topology scheme which depends on the analysis of network. When it is applied for low speed networks; the changing in replicas locations occur slowly which results in more stable availability. On the other hand, when it is applied with high speed networks the changing of replica allocation happens rapidly. So, applying the topology based scheme to improve the availability of CPRP has its best average correctness value with low speed networks, then with moderate speed then with high speed networks.

When the epidemic based scheme is applied with different network speeds and with different infrastructure failure ratios, it shows the best average correctness with moderate speed network with about 57%. Then about 55% average correctness for low speed network, and 53% average correctness value for high speed networks. This is interpreted as the epidemic scheme which depends on generating replicas epidemically generates more replicas in high speed networks. On the other hand, it generates insufficient replicas with the low speed networks. So, it has its best average correctness performance with moderate speed networks.

8 Conclusions

In this work, two schemes are proposed to improve the performance of CPRP protocol in terms of increasing the availability of authentication service offered by RSU in cases RSU unavailability. The first scheme is epidemic-based which spawns new replicas epidemically and doesn't depend on network analysis. The second scheme is topology-based scheme which depends on analyzing the network and compute node degrees, then pushes the new replicas to vehicles with high degrees.

This work is a complement to our previous work. Simulation is done with the same configuration. 24 scenarios are generated for the simulation and are divided into two groups. Four RSUs failure ratios are assumed at: 30%, 50%, 70% and 90%.

For the first group, 12 scenarios are generated as follows: for each RSU failure ratio, three scenarios are generated; with high network density, moderate network density and low network density. The same has done for the second group. For each RSU failure ratio, also three scenarios are generated but with high vehicles speed, moderate and low vehicles speed.

The performance of the two proposed schemes are compared and evaluated in terms of improve the availability of authentication service offered by RSU, increase in replication degree and optimality (correctness) of replica allocation.

Although, the topology-based scheme exposes the network to more security threats because it depends on network analysis which requires revealing the real identity of more vehicles, it shows better performance in all scenarios.

Table 2: Performance evaluation of both epidemic-based and topology-based with different network densities

Network	High Density	Moderate Density	Low Density
Improve Availability with Epidemic	21.8%	7.6%	2.5%
Improve Availability with Topology	19.4%	6.8%	2.7%
Replication Degree with Epidemic	3%	4.6%	7.6%
Replication Degree with Topology	1%	1.5%	1.75%
Correctness with Epidemic	55%	50%	44%
Correctness with Topology	58%	67%	61%

Table 3: Performance evaluation of both epidemic-based and topology-based with different speed networks

Network	High Speed	Moderate Speed	Low Speed
Improve Availability with Epidemic	3.25%	4.5%	11.25%
Improve Availability with Topology	4.25%	2.5%	14.7%
Replication Degree with Epidemic	7.6%	3.7%	2.5%
Replication Degree with Topology	1.3%	1.5%	1.7%
Correctness with Epidemic	53%	57%	55%
Correctness with Topology	60%	62%	67%

First, the performance of the two schemes are evaluated for the first group (Different Network Densities) as shown in Table 2.

- In terms of increasing the availability of authentication service: When applying the two techniques, they achieve about the same performance. But, they improve the performance of CPRP in high density with highest difference.
- In terms of increase in replication degree, topology-based scheme has better performance in all scenarios and is almost the same. Highest value of average replication degree by applying the topology-based scheme is with low density network. By applying epidemic-based scheme, the average replication degree has its greatest value also with low density network.
- In terms of correctness of replica allocation, topology-based scheme outperforms the epidemic-based in almost all scenarios. Topology-based has the highest average correctness with moderate density network. By applying epidemic based scheme, it shows the best average correctness performance with high density network.

Second, the performance of the two schemes are evaluated for the second group (Different speed networks) as shown in Table 3.

- In terms of increasing the availability of authentication service: When applying the two techniques, they converge to be almost the same. The improve the performance of low speed networks with highest difference. It is observed that the epidemic based scheme outperforms the topology based in moderate speed with few difference. But, in low speed network the topology based outperforms the epidemic based.
- In terms of increase in replication degree, topology-based outperforms the epidemic-based in all scenarios with few improvements in low speed networks. Average replication degree has its greatest value with high speed networks. By applying epidemic-based scheme. the average replication degree has its highest value also with high speed networks.
- In terms of correctness of replica allocation, topology-based scheme outperforms the epidemic-based in almost all scenarios. Topology-based has the highest average correctness with low speed network. By applying epidemic based scheme, it shows the best average correctness performance with moderate speed network.

References

[1] P. AGITH, "Enhancement of vehicular ad-hoc networks using vehicle platoon aware data access," *International Journal of Modern Engineering Research*, vol. 2, no. 2, pp. 273–277, 2012.

[2] E. Atsan and O. Ozkasap, "Scalar: Scalable data lookup and replication protocol for mobile ad hoc networks," *Computer Networks*, vol. 57, no. 17, pp. 3654–3672, 2013.

[3] M. Bayat, M. Barmshoory, M. Rahimi, and M. R. Aref, "A secure authentication scheme for vanets with batch verification," *Wireless Networks*, vol. 21, no. 5, pp. 1733–1743, 2015.

[4] M. Caliskan, D. Graupner, and M. Mauve, "Decentralized discovery of free parking places," in *Proceedings of the Third International Workshop on Vehicular Ad Hoc Networks (VANET'06)*, pp. 30–39, Los Angeles, CA, USA, Sept. 2006.

[5] M. El-Eliemy, *Service replication in wireless mobile ad hoc networks*, PhD thesis, Friedrich Schiller University of Jena, 2010.

[6] M. Hamdy and B. König-Ries, "An extended analysis of an interest-based service distribution protocol for mobile ad hoc networks," in *Proceedings of the International Conference on Wireless Information Networks and Systems (WINSYS'08)*, pp. 203–210, Porto, Portugal, July 2008.

[7] M. Hamdy and B. König-Ries, *Service Availability, Success Ratio, Prevalence, Replica Allocation Correctness, Replication Degree, and Effects of Different Replication/Hibernation Behavior Effects of the Service Distribution Protocol for Mobile Ad Hoc Networks*, Germany: University, 2008.

[8] Y. Huang, S. Zeng, and X. Liu, "Privacy-preserving communication for VANETs with conditionally anonymous ring signature," *International Journal of Network Security*, vol. 17, no. 2, pp. 135–141, 2015.

[9] S. Ibrahim and M. Hamdy, "Enabling less-infrastructure road communication networks for vanets," in *Proceedings of The fourth4th International Conference on Computer Science and Network Technology (ICCSNT'15)*, pp. 1–7, Harbin, China, Dec. 2015.

[10] F. Kaisser, C. Gransart, and M. Berbineau, "Simulations of VANET scenarios with OPNET and SUMO," in *4th International Workshop on Communication Technologies for Vehicles*, pp. 103–112, Vilnius, Lithuania, Apr. 2012.

[11] E. Koukoumidis, L. Peh, and M. Martonosi, "RegReS: Adaptively maintaining a target density of regional services in opportunistic vehicular networks," in *Ninth Annual IEEE International Conference on Pervasive Computing and Communications (PerCom'11)*, pp. 120–127, Seattle, WA, USA, Mar. 2011.

[12] U. Lee, E. Magistretti, M. Gerla, P. Bellavista, and A. Corradi, "Dissemination and harvesting of urban data using vehicular sensing platforms," *IEEE Transaction on Vehicular Technology*, vol. 58, no. 2, pp. 882–901, 2009.

[13] Z. Lu and H. YANG, *Unlocking the power of Opnet Modeler*, NewYork, USA: Cambridge University Press, 2012.

[14] T. V. Mathew, *Car Following Models*, ch. 14, pp. 1–8, Indian Institute of Technology, Bombay, 2014.

[15] OPNET, *Simulator*, Apr. 10, 2017. (`http://www.opnet.com/`)

[16] M. Raya and J. Hubaux, "Securing vehicular ad hoc networks," *Journal of Computer Security*, vol. 15, no. 1, pp. 39–68, 2007.

[17] M. Raya, D. Jungels, P. Papadimitratos, I. Aad, and J. Hubaux, "Certificate revocation in vehicular networks," Technical. Report, Laboratory for Computer Communications and Applications (LCA) School of Computer and Communication Sciences, Jan. 2006.

[18] SUMO, *Simulation of Urban MObility*, Apr. 10, 2017. (`http://sumo.sourceforge.net/`)

[19] A. Wasef and X. Shen, "EDR: efficient decentralized revocation protocol for vehicular ad hoc networks," *IEEE Transaction on Vehicular Technology*, vol. 58, no. 9, pp. 5214–5224, 2009.

[20] H. Wu and W. Hsieh, "RSU-based message authentication for vehicular ad-hoc networks," *Multimedia Tools Applications*, vol. 66, no. 2, pp. 215–227, 2013.

[21] C. Zhang, X. Lin, R. Lu, and P. Ho, "RAISE: an efficient RSU-aided message authentication scheme in vehicular communication networks," in *Proceedings of IEEE International Conference on Communications (ICC'08)*, pp. 1451–1457, Beijing, China, May 2008.

Biography

Safi Ibrahim. Received her BSc, MSc, in Information Technology from the faculty of Computers and Information Sciences, Cairo University, Cairo, Egypt. She is currently a Lecturer Assistant in the Department of Information Technology, Egyptian E-Learning University EELU, Cairo, Egypt. Her research interests include ad hoc networks, vehicular communication and Network Security. She has published about four publications in conference proceedings concerning these research areas. She teaches courses on Programming, Wireless Netwoks and Network Security.

Dr. Mohamed Hamdy is an associate professor in the Information Systems Department, the Faculty of Computer and information Sciences at Ain Shams University since 2016. Early 2011, he got his PhD in Mobile Ad Hoc Networks from Friedrich Schiller University Jena, Germany, then he has been back to his home Ain Shams University in Egypt and was working as an assistant professor. His research interest is in general in Wireless Networks, Mobile Data Management. He has engaged in several research projects in MANETs and Data Management.

Dr. Eman Shaaban Received her BSc, MSc, and PhD in computer engineering from Ain-Shams university, Cairo, Egypt. She is currently an associate professor in the Department of computer systems, faculty of computer and information sciences, Ain-Shams university, Cairo, Egypt. She teaches undergraduates courses on data communication, computer architectures and embedded systems, in addition to teaching graduate courses on wireless communications and sensor networks. Her research interests include ad hoc networks, wireless sensor networks, and vehicular communication. She has published over 25 technical papers in peer-reviewed journals and major conference proceeding concerning these research areas.

Certificateless Hybrid Signcryption Scheme with Known Session-Specific Temporary Information Security

Ming Luo, Yuwei Wan, and Donghua Huang
(Corresponding author: Ming Luo)
School of Software, Nanchang University, Nanchang, China
999, Xuefu Avenue, Jiangxi Sheng, Nanchang Shi, Xinjian Xian, China (Email: lmhappy21@163.com)
(Received Sept. 23, 2016; revised and accepted Jan 10 & Feb. 1, 2017)

Abstract

The hybrid signcryption scheme based on certificateless public key cryptography avoids the complexity of certificate management existing in the traditional public key cryptography and the inherent key escrow problem existing in identity-based public key cryptography. The certificateless hybrid signcryption scheme combined with certificateless signcryption key encapsulation mechanism and data encapsulation mechanism can dispose the messages with arbitrary length while conventional certificateless signcryption schemes cannot. Meanwhile, almost all the proposed certificateless hybrid signcryption schemes cannot survive against the known session-specific temporary information security (KSSTIS) attack. In this paper we propose an efficient certificateless hybrid signcryption scheme, and formally prove its security in random oracle model under the assumption of Diffie-Hellman mathematical hard problems. Compared with the previous schemes, our scheme has the advantage of lower computational cost by reducing the amount of bilinear pairing computation. Moreover, our scheme achieves KSSTIS attribute.

Keywords: Certificateless; Hybrid Signcryption; Random Oracle Model; KSSTIS

1 Introduction

Signcryption is a cryptographic primitive which performs both the functions of signature and encryption in one logical step. With lower computational and communication cost, signcryption promotes the development of public key cryptography. Traditional public key cryptography, identity-based public key cryptography (IBC) and certificateless public key cryptography are three important stages of public key cryptography. For a long period of time, many signcryption schemes using conventional public key infrastructure (PKI) have been proposed, which binds user's identity and public key with a certificate. But the certificate management is a particularly prominent issue. In order to solve this problem and reduce the burden on traditional PKI, Identity-based public key cryptography was proposed, and a number of related signcryption schemes [7, 8] have been proposed in recent years. For IBC, the public key is computed with the binary string of users identity, thus IBC does not need the certificate used in PKI. However, the private key of IBC is generated by a private key generator (PKG). In this situation, private key escrow becomes an inherent problem in IBC. The PKG can forge or decrypt any ciphertext.

The notion of certificateless public key cryptography (CLC) was presented by Al-Ryiami and Paterson [2], which solves the certificate management problem of the traditional PKI and the inherent key escrow problem of IBC. For CLC, the private key is divided into two parts, one part is selected by users themselves and the other is generated by a key generation center (KGC). In 2008, Barbosa and Farshim [3] firstly proposed a certificateless signcryption scheme and its security notions. Recently, many signcryption schemes [6,13] using certificateless cryptography have been proposed.

The notion of hybrid encryption was presented by Abe et al. [1], and then Dent proposed the notion of hybrid signcryption [4]. Hybrid signcryption includes two parts. One part is a key encapsulation mechanism (CLSC-KEM) and the other part is a data encapsulation mechanism (DEM).In recent years, some hybrid signcryption schemes have been proposed for various network applications [9, 11]. Li et al. [5] proposed the first certificateless hybrid signcryption(CLHSC) scheme. The scheme consists of a tag key encapsulation mechanism (tag-KEM) and a data encapsulation mechanism (DEM), and their scheme makes up for the lack of authentication security in Dent's scheme [4]. At the signcryption stage, a symmetric key is generated by the key encapsulation mechanism, and then outputs the signcryption data. At the decryption stage, after obtaining the symmetric key by decapsulating the signcryption data, the ciphertext will

be decrypted. Later, Selvi et al. [10] pointed out that Li's scheme may be existentially forgeable and proposed an improved scheme. Recently, Yin and Liang [12] pointed out almost all certificateless signcryption schemes that have been proposed in the literature cannot effectively against the public-key-replacement attacks, and they proposed an enhanced scheme to fill this security gaps.

However, we find these certificateless hybrid signcryption schemes above cannot survive against known session-specific temporary information security (KSSTIS) attack. To compensate for this security flaw, this paper proposes a new hybrid signcryption scheme based on certificateless cryptography and proves that the scheme meets the confidentiality and unforgeability in random oracle model, also our scheme can against the public-key-replacement attacks. Compared with the schemes above, our scheme achieves KSSTIS security attributes and has less bilinear pairing computation.

2 Preliminaries

Let $\mathbb{G}_1$ and $\mathbb{G}_2$ be a cyclic additive and multiplicative group respectively, whose prime order is a large prime number q. P is a generator of the group $\mathbb{G}_1$. If a map $\hat{e} : \mathbb{G}_1 \times \mathbb{G}_1 \to \mathbb{G}_2$ satisfies the following properties, we call it bilinear pairing.

1) Bilinearity: for all $a, b \in \mathbb{Z}_q^*$, there is $\hat{e}(aP, bP) = \hat{e}(P, P)^{ab}$.

2) Computability: for all $P, N \in \mathbb{G}_1$, there is an efficient algorithm to compute $\hat{e}(P, N)$.

3) Non-degeneracy: there exists $P \in \mathbb{G}_1$, such that $\hat{e}(P, P) \neq 1_{G2}$.

We can construct bilinear pairing $\hat{e}$ using the modified Tate pairing and Weil pairing of elliptic curve over a finite field. The security of our scheme relies on the following hard problems.

Definition 1. *Computational Diffie-Hellman(CDH) problem: For two integers $a, b \in \mathbb{Z}_q^*$ and a generator P of $\mathbb{G}_1$, given the tuple (P, aP, bP) to compute abP is hard.*

Definition 2. *Computational Bilinear Diffie-Hellman (CBDH) problem: For three integers $a, b, c \in \mathbb{Z}_q^*$ and a generator P of $\mathbb{G}_1$, given the tuple (P, aP, bP, cP) to compute $\hat{e}(P, P)^{abc}$ is hard.*

3 Certificateless Hybrid Signcryption Scheme

In this section, the certificateless hybrid signcryption scheme is described in details. Our scheme includes the following algorithms:

Setup: On input of a security parameter k, KGC picks a bilinear pairing $\hat{e} : \mathbb{G}_1 \times \mathbb{G}_1 \to \mathbb{G}_2$ and three security cryptographic hash functions $H_1 : \{0,1\}^* \times \mathbb{G}_1 \to \mathbb{G}_1$, $H_2 : \{0,1\}^* \times (\mathbb{G}_1)^4 \times \mathbb{G}_2 \to \{0,1\}^n$ and $H_3 : \{0,1\}^* \times (\mathbb{G}_1)^4 \to \{0,1\}^n$. Then the KGC randomly chooses a master key $s \in \mathbb{Z}_q^*$ and computes the master public key $P_{pub} = sP$. The KGC keeps the master key s and publishes the system parameters $params =< G_1, G_2, \hat{e}, q, P, P_{pub}, H_1, H_2, H_3 >$.

GUK (Generate user key): On input of an identity ID and the system parameters $params$, a user randomly choose $x_{ID} \in \mathbb{Z}_q^*$ as his secret key, and then computes his public key $PK_{ID} = x_{ID}P$.

EPPK (Extract partial private key): On input of an identity ID and the system parameters $params$, KGC computes $Q_{ID} = H_1(ID||PK_{ID})$, and then computes the partial private key $D_{ID} = sQ_{ID}$.

GSK (Generate symmetric key): On input of sender's identity ID_s, public key PK_s, and private key (x_s, D_s), receiver's identity ID_r and public key PK_r. Randomly choose $x, y \in \mathbb{Z}_q^*$, the sender does the following steps.

1) Compute $U = xP$, $T = \hat{e}(D_s, Q_r)$.
2) Compute session key $K_{AB} = H_2(ID_r, T, U, xPK_r, x_sPK_r, PK_r)$.
3) Obtain internal state information $\bar{W} = (x, y, U, x_s, D_s, ID_s, PK_s, ID_r, PK_r)$.
   Output $(K_{AB}, \bar{W})$.

Encapsulation: On input of a tag τ and internal state information $\bar{W}$. The algorithm works as the following steps.

1) Compute $w = y(D_s + x_sPK_r)$.
2) Compute $h = H_3(\tau, U, w, PK_s, Pk_r)$.
3) Compute $v = 1/(y(x + h))$.
   Output $\delta = (U, w, v)$.

Decapsulation: On input of signcryption δ, a tag τ, the sender's identity ID_s, public key PK_s, and the receiver's identity ID_r, public key PK_r, private key (x_r, D_r). The receiver does the following steps.

1) Compute $h = H_3(\tau, U, w, PK_s, PK_r)$.
2) Check if $\hat{e}(vw, U + hP) \stackrel{?}{=} \hat{e}(Q_s, P_{pub})\hat{e}(PK_s, PK_r)$. If it is correct, go on and do the following computations. Otherwise stop and return $\perp$.
3) Compute $T = \hat{e}(D_r, Q_s)$.
4) Compute session key $K_{AB} = H_2(ID_r, T, U, x_rU, x_rPK_s, PK_r)$.

4 Security Analysis

In this section, we use some mathematical hard problems to analyze the confidentiality and unforgeability security of the scheme in the random oracle model, then

we show that our scheme can survive against known session-specific temporary information security (KSSTIS) attacks.

4.1 Consistency

Our scheme satisfies the consistency.

$$\begin{aligned}&\hat{e}(vw, U + hP)\\ &= \hat{e}(y(D_s + x_s PK_r)/y(x + h), xP + hP)\\ &= \hat{e}((D_s + x_s PK_r), P)\\ &= \hat{e}(D_s, P)\hat{e}(x_s PK_r, P)\\ &= \hat{e}(Q_s, P_{pub})\hat{e}(PK_s, PK_r).\end{aligned}$$

4.2 Confidentiality

Theorem 1. *Assuming that CBDH is hard to solve in random oracle model, the scheme is secure against any IND-CLHSC-CCA2-I adversary A_I attack.*

Proof. Assuming that the challenger C receives an CBDH challenge tuple (P, aP, bP, cP), where P is a generator of cyclic additive $\mathbb{G}_1$. And the goal for C is to compute the answer of $\hat{e}(P, P)^{abc}$. The challenger C sends the system parameters *params* to A_I, and sets $P_{pub} = aP$. C maintains several lists $L_1, L_2, L_3, L_u, L_e, L_d$ and answers the following queries. Among these lists, L_1, L_2, L_3 simulate H_1, H_2, H_3 oracle respectively,L_u is used to track GUK query, L_e is used to track Encapsulation query, L_d is used to track Decapsulation query.

H_1 query: C selects two random numbers $i, j \in \{1, 2, \cdots, q_1\}$, where q_1 is the number of H_1 queries. At the n-th query:

1) if $ID_n = ID_i$, C answers $Q_i = bp$, and adds the tuple $(ID_i, \perp, bP)$ into list L_1.
2) if $ID_n = ID_j$, C answers $Q_j = cP$, and adds the tuple $(ID_j, \perp, cP)$ into list L_1.
3) if $ID_n \notin \{ID_i, ID_j\}$, C randomly chooses $w \in \mathbb{Z}_q^*$, answers $Q_n = wP$, and then returns it and adds the tuple (ID_n, w, Q_n) into list L_1.

H_2 query: C checks if there exists a tuple $(ID_r, T, U, xPK_r, x_s PK_r, PK_r, h_2)$ in the list L_2. If the tuple is found, C returns h_2. Otherwise, C randomly chooses $h_2 \in \{0, 1\}^n$, and then returns it and adds the tuple $(ID_r, T, U, xPK_r, x_s PK_r, PK_r, h_2)$ into list L_2.

H_3 query: C checks if there exists a tuple $(\tau, U, w, PK_s, PK_r, h_3)$ in the list L_3. If the tuple is found, C returns h_3. Otherwise, C randomly chooses $h_3 \in \mathbb{Z}_q^*$, and then returns it and adds the tuple$(\tau, U, w, PK_s, PK_r, h_3)$ into list L_3.

***GUK* query:** A_I picks an identity ID_n, C randomly chooses $x_n \in \mathbb{Z}_q^*$, and then answers $PK_n = x_n P$, adds the tuple(ID_n, x_n, PK_n) into list L_u.

***EPPK* query:** A_I picks an identity ID_n. Assuming that the identity ID_n has made H_1 query before, if $ID_n \in \{ID_i, ID_j\}$, stops the challenge. Otherwise, C searches the corresponding tuple (ID_n, w, Q_n) in the list L_1, returns $D_n = wP_{pub}$ and answers D_n.

Corruption query: A_I picks an identity ID_n. Assuming that the identity ID_n has made GUK query before, C searches the corresponding tuple in the list L_1, and answers x_n.

RPK query: A_I picks a new tuple (ID_n, PK_n), C updates the list L_u and replaces with $(ID_n, \perp, PK_n)$.

GSK query: A_I picks a tuple (ID_s, PK_s, ID_r, PK_r).

1) If $ID_s \notin \{ID_i, ID_j\}$, C randomly chooses $x, y \in \mathbb{Z}_q^*$, computes U and T. And then C runs the symmetric key generation algorithm and answers K_{AB}, updates and stores the internal state information.
2) If $ID_s \in \{ID_i, ID_j\}$, C stops simulation.

Encapsulation query: A_I produces a tag τ, at the same time, C checks if there exists an internal state information $\bar{W}$. If it is found, C performs the following steps. Otherwise, C stops the simulation and returns a $\perp$.

1) If $ID_s \notin \{ID_i, ID_j\}$, C computes $w = y(D_s + x_s PK_r)$ with the internal state information, and then computes $h = H_3(\tau, U, w, PK_s, PK_r)$and $v = 1/y(x + h)$. Finally, C answers the signcryption $\delta = (U, w, v)$ to A_I.
2) If $ID_s \in \{ID_i, ID_j\}$, C stops simulation.

Decapsulation query: A_I picks the tag τ, signcryption $\delta = (U, w, v)$, the sender's identity ID_s and the receiver's identity ID_r. C does the following processing:

1) If $ID_r \notin \{ID_i, ID_j\}$, firstly C computes $h = H_3(\tau, U, w, PK_s, PK_r)$, and then checks if $\hat{e}(vw, U + hP) \stackrel{?}{=} \hat{e}(Q_s, P_{pub})\hat{e}(PK_s, PK_r)$. If it is failure, C stops simulation and returns $\perp$. Otherwise, C computes $T = \hat{e}(D_r, Q_s)$, and then computes the session key $K_{AB} = H_2(ID_r, T, U, x_r U, x_r PK_s, PK_r)$.
2) If $ID_r \in \{ID_i, ID_j\}$, C stops simulation.

Challenge: A_I can stop the phase 1 queries whenever he wants, and then produces two challenge identities $\{ID_A, ID_B\}$, which $ID_A \neq ID_B$. if $\{ID_A, ID_B\} \notin \{ID_i, ID_j\}$, C stops simulation. Otherwise C randomly chooses $x, y \in \mathbb{Z}_q^*$, and sets $T^* = \eta$ (η as a candidate answer for CBDH problem), and then computes $U^* = xP$, $K_1 = H_2(ID_B, T^*, U^*, xPK_B, x_A PK_B, PK_B)$. C randomly chooses a number $K_0 \in \{0, 1\}^n$ and a bit $d \in \{0, 1\}$, sends K_d to A_I. A_I chooses a tagτ^*

and sends it to C, C picks $w^* \in \mathbb{G}_1$, computes $h^* = H_3(\tau^*, U^*, w^*, PK_A, PK_B)$, $v^* = 1/y(x + h^*)$. Finally, C sends the signcryption $\delta^* = (U^*, w^*, v^*)$ to A_I.

A_I makes the queries of Phase 2 just like he made in the first phase. At last A_I produces a bit $d' \in \{0, 1\}$ as a guess to d. Only when A_I uses the tuple $(ID_B, T^*, U^*, xPK_B, x_APK_B, PK_B)$to make H_2 query, he can check the correctness of the signcryption $\delta^* = (U^*, w^*, v^*)$, and if $d' = d$, C outputs T as a solution of the CBDH since the candidate answer $K_{AB} = H_2(ID_B, T^*, U^*, x_BU^*, x_BPK_A, PK_B)$ for CBDH problem is in the list L_2, where $T^* = \eta = \hat{e}(D_B, Q_A) = \hat{e}(acP, bP) = \hat{e}(P, P)^{abc}$. If $d' \neq d$, C fails and outputs F.

Thus, if the adversary A_I wants to break the signcryption algorithm, he must solve the CBDH with non-negligible advantage first. What he can do is to extract information from the signcryption messages, then uses some polynomial-time algorithm to solve the CBDH problem. But we all know that this algorithm does not exist so far. Therefore, when attacked by an IND-CLHSC-CCA2 adversary A_I, the proposed CLHSC scheme can maintain a safe state. □

Theorem 2. *Assuming that CDH is hard to solve in random oracle model, the scheme is secure against any IND-CLHSC-CCA2-II adversary A_{II} attack.*

Proof. Assuming that the challenger C receives an CDH challenge tuple (P, aP, bP), where P is a generator of cyclic additive $\mathbb{G}_1$. And the goal for C is to compute the answer of abP. C randomly chooses a number $s \in \mathbb{Z}_q^*$ as the master secret key, sets $P_{pub} = sP$, and sends the system parameters *params* and s to A_{II}. C maintains several lists $L_1, L_2, L_3, L_u, L_e, L_d$ and answers the following queries. Among these lists, L_1, L_2, L_3 simulate H_1, H_2, H_3 oracle respectively, L_u is used to track GUK query, L_e is used to track Encapsulation query, L_d is used to track Decapsulation query.

H_1 query: A_{II} randomly picks an identity ID_i, and sends it to C. C randomly chooses $w \in \mathbb{Z}_q^*$, computes $Q_n = wP$, and then returns it and adds the tuple (ID_n, w, Q_n) into list L_1.

H_2 query: The same as Theorem 1.

H_3 query: The same as Theorem 1.

GUK query: C selects a random number $i \in \{0, 1, \cdots, q_u\}$, where q_u is the number of GUK queries. At the n-th query:

1) If $ID_n \neq ID_i$, C randomly chooses $x_n \in \mathbb{Z}_q^*$ as the secret value, computes the public key $PK_n = x_nP$, and then adds the tuple (ID_n, x_n, PK_n) into list L_u and answers PK_n.
2) If $ID_n = ID_i$, C sets $PK_i = bP$, adds the tuple $(ID_i, \perp, bP)$ into list L_u.

Corruption query: A_{II} picks an identity ID_n. Assuming that the identity ID_n has made GUK query before, if $ID_n = ID_i$, C stops simulation. Otherwise, C searches the corresponding tuple in the list L_u and answers x_n.

GSK query: A_{II} picks a tuple (ID_s, PK_s, ID_r, PK_r).

1) If $ID_s \neq ID_i$, C randomly chooses $x, y \in \mathbb{Z}_q^*$, computes U and T. And then C runs the symmetric key generation algorithm and answers K_{AB}, updates and stores the internal state information.
2) If $ID_s = ID_i$, C stops simulation.

Encapsulation query: A_{II} produces a tag τ, and at the same time, C checks if there exists an internal state information $\bar{W}$. If it is found, perform the following steps. Otherwise, C stops the simulation and returns $\perp$.

1) If $ID_s \neq ID_i$, C computes $w = y(D_s + x_sPK_r)$ with the internal state information, and then computes $h = H_3(\tau, U, w, PK_s, PK_r)$ and $v = 1/y(x + h)$. Finally, C answers the signcryption $\delta = (U, w, v)$ to A_{II}.
2) If $ID_s = ID_i$, C stops simulation.

Decapsulation query: A_{II} picks the tag τ, signcryption $\delta = (U, w, v)$, the sender's identity ID_s and the receiver's identity ID_r. C does the following processing:

1) If $ID_r \neq ID_i$, firstly C computes $h = H_3(\tau, U, w, PK_s, PK_r)$, and then checks if $\hat{e}(vw, U + hP) \stackrel{?}{=} \hat{e}(Q_s, P_{pub})\hat{e}(PK_s, PK_r)$. If it is failure, C stops simulation and returns $\perp$. Otherwise, C computes $T = \hat{e}(D_r, Q_s)$, and then computes the session key $K_{AB} = H_2(ID_r, T, U, x_rU, x_rPK_s, PK_r)$.
2) If $ID_r = ID_i$, C stops simulation.

Challenge: A_{II} can stop the phase 1 queries whenever he wants, and produces two challenge identities $\{ID_A, ID_B\}$, which $ID_A \neq ID_B$. If $ID_B \neq ID_i$, C stops simulation. Otherwise C sets $U^* = aP$, randomly chooses $y \in \mathbb{Z}_q^*$, and then computes $T^* = \hat{e}(D_A, Q_B)$, $K_1 = H_2(ID_B, T^*, U^*, \eta, x_APK_B, PK_B)$ (η as a candidate answer for CDH problem). C randomly chooses a number $K_0 \in \{0, 1\}^n$ and a bit $d \in \{0, 1\}$, sends K_d to A_{II}. A_{II} chooses a tag τ^* and sends it to C, C picks $v^* \in \mathbb{Z}_q^*$, computes $w^* = y(D_A + x_APK_B)$, $h^* = H_3(\tau^*, U^*, w^*, PK_A, PK_B)$. Finally, C sends the signcryption $\delta^* = (U^*, w^*, v^*)$ to A_{II}.

A_{II} makes the queries of Phase 2 just like he made in the first phase. At last A_{II} produces a bit $d' \in \{0, 1\}$ as a guess to d. Only when A_{II} uses the tuple $(ID_B, T^*, U^*, \eta, x_APK_B, PK_B)$ to make H_2 query,

he can check the correctness of the signcryption $\delta^* = (U^*, w^*, v^*)$, and if $d' = d$, C outputs T as a solution of the CDH since the candidate answer $K_{AB} = H_2(ID_B, T^*, U^*, \eta, x_A PK_B, PK_B)$ for CDH problem is in the list L_2, where $\eta = x_B U^* = baP = abP$. If $d' \neq d$, C fails and outputs F.

Thus, if the adversary A_{II} wants to break the signcryption algorithm, he must solve the CDH with non-negligible advantage first. What he can do is to extract information from the signcryption messages, then use some polynomial-time algorithm to solve the CDH problem. This algorithm does not exist yet. Therefore, when attacked by an IND-CLHSC-CCA2 adversary A_{II}, the proposed CLHSC scheme can maintain a safe state. □

4.3 Unforgeability

Theorem 3. *Assuming that CDH is hard to solve in random oracle model, our scheme is secure against any sUF-CLHSC-CMA-I adversary A_I attack.*

Proof. Assuming that the challenger C receives an CDH challenge tuple (P, aP, bP), where P is a generator of cyclic additive $\mathbb{G}_1$. And the goal for C is to compute the answer of abP. The challenger C sends the system parameters *params* to A_I, and set $P_{pub} = aP$. C maintains several lists $L_1, L_2, L_3, L_u, L_e, L_d$ and answers the following queries. Among these lists, L_1, L_2, L_3 simulate H_1, H_2, H_3 oracle respectively, L_u is used to track GUK query,L_e is used to track Encapsulation query, L_d is used to track Decapsulation query.

H_1 **query:** C selects a random number $i \in \{1, 2, \cdots, q_1\}$,where q_1 is the number of H_1 queries. At the n-th query:

1) If $ID_n = ID_i$, C answers $Q_i = bP$, and adds the tuple $(ID_i, \perp, bP)$ into list L_1.
2) If $ID_n \neq ID_i$, C randomly chooses $w \in Z_q^*$, answers $Q_n = wP$, and then returns it and adds the tuple (ID_n, w, Q_n) into list L_1.

H_2 **query:** The same as Theorem 1.

H_3 **query:** The same as Theorem 1.

GUK **query:** The same as Theorem 1.

EPPK **query:** A_I picks an identity ID_n. Assuming that the identity ID_n has made H_1 query before, if $ID_n = ID_i$, stops the challenge. Otherwise, C searches the corresponding tuple (ID_n, w, Q_n) in the list L_1, returns $D_n = wP_{pub}$ and answers D_n.

Corruption query: The same as Theorem 1.

RPK query: The same as Theorem 1.

GSK query: The same as Theorem 2.

Encapsulation query: The same as Theorem 2.

Decapsulation query: The same to Theorem 2.

Eventually, A_I produces a valid forgery quaternion $(\tau^*, \delta^*, ID_A, ID_B)$. C checks if $ID_A \neq ID_i$. If it is the case, C aborts. Otherwise, with the help of GUK oracle, C can obtain ID_A's public key PK_A and ID_B's public key PK_B, respectively. After that C uses tuple $(\tau^*, U^*, w^*, PK_A, PK_B)$ to make H_3 query and obtains h^* from list L_3. Then C does the following verification:

$$\begin{aligned} \hat{e}(v^* w^*, U^* + h^* P) &= \hat{e}(Q_A, P_{pub})\hat{e}(PK_A, PK_B) \\ \hat{e}(w^*/y, P) &= \hat{e}(bP, aP)\hat{e}(x_A P, PK_B) \\ \hat{e}(abP, P) &= \hat{e}(P, (w^*/y) - x_A PK_B). \end{aligned}$$

At last, C can compute $abP = (w^*/y) - x_A PK_B$.

If verification is right, C returns 1, otherwise 0.

So, if there exists a special adversary A_I who can forge a valid encapsulation message by learning something about the signcryption, that means there is an algorithm which can solve CDH problem with non-negligible advantage. However, this cannot happen. In other words, there is no adversary who can forge in this way. Thus, the scheme is secure against any sUF-CLHSC-CMA-I adversary A_I attack. □

Theorem 4. *Assuming that CDH is hard to solve in random oracle model, the scheme is secure against any IND-CLHSC-CCA2-II adversary A_{II} attack.*

Proof. Assuming that the challenger C receives an CDH challenge tuple (P, aP, bP), where P is a generator of cyclic additive $\mathbb{G}_1$. And the goal for C is to compute the answer of abP. C randomly chooses a number $s \in \mathbb{Z}_q^*$ as the master secret key, sets $P_{pub} = sP$, and sends the system parameters *params* and s to A_{II}. C maintains several lists, $L_1, L_2, L_3, L_u, L_e, L_d$ and answers the following queries. Among these lists, L_1, L_2, L_3 simulate H_1, H_2, H_3 oracle respectively, L_u is used to track GUK query, L_e is used to track Encapsulation query, L_d is used to track Decapsulation query.

H_1 **query:** The same as Theorem 2.

H_2 **query:** The same as Theorem 1.

H_3 **query:** The same as Theorem 1.

GUK **query:** C selects two random numbers $i, j \in \{1, 2, \cdots, q_u\}$, where q_u is the number of GUK queries. At the n-th query:

1) If $ID_n = ID_i$, C answers $PK_i = aP$, and adds the tuple $(ID_i, \perp, aP)$ into list L_u.
2) If $ID_n = ID_j$, C answers $PK_j = bP$, and adds the tuple $(ID_j, \perp, bP)$ into list L_u.
3) If $ID_n \notin \{ID_i, ID_j\}$, C randomly chooses $x_n \in \mathbb{Z}_q^*$ as the secret key and computes $PK_n = x_n P$, and then answers it and adds the tuple (ID_n, x_n, PK_n) into list L_u.

Corruption query: A_{II} picks an identity ID_n. Assuming that the identity ID_n has been made GUK query before, if $ID_n \in \{ID_i, ID_j\}$, C stops simulation. Otherwise, C searches the corresponding tuple in the list L_u and answers x_n.

GSK query: The same as Theorem 1.

Encapsulation query: The same as Theorem 1.

Decapsulation query: The same as Theorem 1.

Eventually, A_{II} produces a valid forgery quaternion $(\tau^*, \delta^*, ID_A, ID_B)$. C checks if $\{ID_A, ID_B\} \notin \{ID_i, ID_j\}$ and $ID_A \neq ID_B$. If it is the case, C aborts. Otherwise, with the help of GUK oracle, C can obtain $ID_A's$ public key PK_A and $ID_B's$ public key PK_B respectively. After that C uses tuple $(\tau^*, U^*, w^*, PK_A, PK_B)$ to make H_3 query and obtains h^* from list L_3. Then do the following verification:

$$\hat{e}(v^*w^*, U^* + h^*P) = \hat{e}(Q_A, P_{pub})\hat{e}(PK_A, PK_B)$$
$$\hat{e}(w^*/y, P) = \hat{e}(D_A, P)\hat{e}(aP, bP)$$
$$\hat{e}(abP, P) = \hat{e}(P, (w^*/y) - D_A).$$

At last, C can compute $abP = (w^*/y) - D_A$.

If verification is right, C returns 1, otherwise 0.

So, if there exists a special adversary A_{II} who can forge a valid encapsulation message by learning something about the signcryption, that means there is an algorithm which can solve CDH problem with non-negligible advantage. This is impossible. In other words, there is no adversary who can forge in this way. Thus, the scheme is secure against any sUF-CLHSC-CMA-II adversary A_{II} attack. □

4.4 Known Session-specific Temporary Information Security

Assuming that at the $j-th$ communication, ephemeral key x_j and signcryption $\delta_j = (U_j, w_j, v_j)$ is leaked. For adversary A_I, he can not obtain the related information about private key (D_s, x_s) or (D_r, x_r). A_I cannot compute $T_j = \hat{e}(D_s, Q_r)$ or $T_j = \hat{e}(D_r, Q_s)$ under the assumption of CBDH problem and cannot compute x_sPK_r or x_rPK_s under the assumption of CDH problem. All above problems will lead to the result that it is hard to obtain the value of session key $K_{AB} = H_2(ID_r, T, U, x_jPK_r, x_sPK_r, PK_r)$ for A_I. For adversary A_{II}, in the scheme, A_{II} can obtain the partial private key D_s or D_r, and then he can compute $T_j = \hat{e}(D_s, Q_r)$ or $T_j = \hat{e}(D_r, Q_s)$. But A_{II} cannot compute x_sPK_r or x_rPK_s without x_s or x_r under the assumption of CDH problem. This leads to the result that it is hard to compute $K_{AB} = H_2(ID_r, T, U, x_jPK_r, x_sPK_r, PK_r)$. Hence, our scheme can survive against Known session-specific temporary information security (KSSTIS) attack. But in Li's scheme [5], when the adversary obtains the ephemeral key r_j of $j-th$ communication, he can obtain $T = \hat{e}(P_{pub}, Q_{ID_r})^{r_j}$ easily. And then it is easy for the adversary to obtain the session key $K_{AB} = H_2(U, T, r_jPK_{ID_r}, ID_r, PK_{ID_r})$. The same situation happens in Yin's scheme [12]. When the adversary obtains the ephemeral key r_{1-j}, r_{2-j} of $j-th$ communication, he can obtain $R_1 = r_{1-j}P, R_2 = r_{2-j}P, U = r_{1-j}PK_R$ and $V = \hat{e}(r_{2-j}Q_R, P_{pub})$ easily. And the session key $K = H_2(ID_S, ID_R, R_1, R_2, U, V)$ can be easily obtained.

5 Performance Analysis

In this section, we will compare the scheme with Li's scheme and Yin's scheme from two aspects: the security and the efficiency of Encapsulation(include GSK phase) and Decapsulation phase in the table 1. We assume that all the three schemes use the same parameters $< G_1, G_2, \hat{e}, q >$. In the column of "Security", "KISSTIS" refers to known session-specific temporary information security. "Y" and "N" denote that whether satisfy this security property. In the column of "Computation Cost", the notations "Encapsulation" and "Decapsulation" refer to the computation of Encapsulation and Decapsulation, respectively. Note that offline computation is not included in "Computation Cost". And here, three operations will be involved. MUL, EXP and PAI refer to the number of point scalar multiplications, exponentiations and bilinear pairing computations, respectively.

Table 1: Comparison of efficiency

Scheme	Security	Computation Cost	
	KISSTIS	Encapsulation	Decapsulation
Li [5]	N	4MUL+EXP	MUL+4PAI
Yin [12]	N	5MUL+EXP	4MUL+3PAI
Ours	Y	3MUL	3MUL+PAI

Through Table 1 we can see that our scheme only needs 3 point scalar multiplications at the Encapsulation step, which is more efficient than the other two schemes. And at the Decapsulation stage, our scheme needs three point scalar multiplications and one bilinear pairing computation. The computation cost of bilinear pairing computation is the most expensive in the scheme based on bilinear pairing. Although Li's scheme only needs one point scalar multiplication, the number of bilinear pairing computations is far more than our scheme. Hence, our scheme is the most efficient. And from the security aspect, our scheme achieves the known session-specific temporary information security, which Li's and Yin's schemes can not satisfy.

6 Conclusion

In this paper, a secure CLHSC scheme is proposed from bilinear pairing in random oracle model. In addition, the scheme is highly efficient with only one bilinear pairing operation. In terms of security, we solve the flaw that most of the hybrid signcryption schemes cannot survive against known session-specific temporary information security attack. Considering any length of plaintext can be handled by hybrid signcryption and the efficiency of our scheme, our scheme can be applied to the high security requirements of communication networks and bandwidth-constrained communication environments, such as ad hoc net, 4G communication and so on.

Acknowledgment

We would like to thank the anonymous reviewers for their valuable comments and suggestions. This work is supported by the National Natural Science Foundation of China under grant [no. 61662046 and 61601215] and the Science and Technology project of Jiangxi Province of China under grant no. 20142BBE50019.

References

[1] M. Abe, R. Gennaro, K. Kurosawa, and V. Shoup, "Tag-kem/dem: A new framework for hybrid encryption and a new analysis of kurosawa-desmedt kem," in *Advances in Cryptology (EUROCRYPT'05)*, pp. 128–146, 2005.

[2] S. S. Al-Riyami and K. G. Paterson, "Certificateless public key cryptography," in *International Conference on the Theory and Application of Cryptology and Information Security*, pp. 452–473, 2003.

[3] M. Barbosa and P. Farshim, "Certificateless signcryption," in *Proceedings of the 2008 ACM symposium on Information, Computer and Communications Security*, pp. 369–372, 2008.

[4] A. W. Dent, "Hybrid signcryption schemes with outsider security," in *International Conference on Information Security*, pp. 203–217, 2005.

[5] F. G. Li, M. Shirase, and T. Takagi, "Certificateless hybrid signcryption," *Mathematical and Computer Modelling*, vol. 57, no. 3, pp. 324–343, 2013.

[6] M. Luo, S. Q. Wang, and J. Hu, "A more efficient and secure broadcast signcryption scheme using certificateless public-key cryptography for resource-constrained networks," *Journal of Internet Technology*, vol. 17, no. 1, pp. 81–89, 2016.

[7] M. Mandal, G. Sharma, and A. K. Verma, "A computational review of identity-based signcryption schemes," *International Journal of Network Security*, vol. 18, no. 5, pp. 969–977, 2016.

[8] Y. Ming and Y. M. Wang, "Cryptanalysis of an identity based signcryption scheme in the standard model," *International Journal of Network Security*, vol. 18, no. 1, pp. 165–171, 2016.

[9] M. Ramakrishnan and R. Sujatha, "Cf-huffman code based hybrid signcryption technique for secure data transmission in medical sensor network," *International Journal of Applied Engineering Research*, vol. 10, no. 4, pp. 11455–11474, 2015.

[10] S. S. D. Selvi, S. S. Vivek, and C. Pandu Rangan, "Breaking and re-building a certificateless hybrid signcryption scheme," *ePint IACR org/2009/62*, 2009.

[11] R. Sujatha, M. Ramakrishnan, N. Duraipandian, and B. Ramakrishnan, "Optimal adaptive genetic algorithm based hybrid signcryption algorithm for information security," *CMES: Computer Modeling in Engineering & Sciences*, vol. 105, no. 1, pp. 47–68, 2015.

[12] A. H. Yin and H. C. Liang, "On security of a certificateless hybrid signcryption scheme," *Wireless Personal Communications*, vol. 85, no. 4, pp. 1727–1739, 2015.

[13] Y. W. Zhou, B. Yang, and W. Z. Zhang, "Provably secure and efficient leakage-resilient certificateless signcryption scheme without bilinear pairing," *Discrete Applied Mathematics*, vol. 204, pp. 185–202, 2016.

Biography

Ming Luo received the B.E. and Ph.D degree from Northeastern University, Shenyang, China in 2004 and 2010, respectively. Now he is an associate professor at the School of Software, Nanchang University, Nanchang, China. He has won lots of scholarships in China and was supported by the National High-Tech Research and Development Plan of China, the National Natural Science Foundation of China and the Science and Technology Program of Jiangxi Province. His research interests are networks security, information security and cryptography.

Yuwei Wan is a M.S. candidate at the school of software, Nanchang University. Her current research interests include information security and cryptography.

Donghua Huang is a M.S. candidate at the school of software, Nanchang University. His current research interests include information security and cryptography.

Revocable ABE with Bounded Ciphertext in Cloud Computing

Mohamed Ali Hamza[1,2], Jianfei Sun[1], Xuyun Nie[1], Zhiquan Qin[1], and Hu Xiong[1,3]
(Corresponding author: Mohamed Ali Hamza)

School of Software Engineering, University of Electronic Science and Technology of China[1]
4 Jianshe North Rd 2nd Section, Chenghua Qu, Chengdu Shi, Sichuan Sheng 610051, China
(Email: mody231279@yahoo.com)
Department of Electronic Engineering, Karary University, Omdurman, Sudan[2]
State Key Laboratory of Cryptology, P. O. Box 5159, Beijing, 100878, China[3]
(Received June 24, 2016; revised and accepted Sept. 3 & Oct. 25, 2016)

Abstract

Revocable Attribute-Based Encryption (R-ABE) has received much concern recently due to its characteristic of capability on encrypting the Data, according to some attributes, whereas users can decrypt the ciphertexts if they own the credential of those attributes with ability to revoke the expired users. We propose a new practical Revocable Attribute Based Encryption which has a short ciphertext O(1) and private keys O(1) with efficient running time. In this scheme the users can effectively be revoked and added with backward and forward secrecy in the indirect mode, which can controlled by Key Authority Party without reseting the system parameter's or updating and redistributing the attributes private keys which has expense. Assuming the cloud provider is semi-honest and has been delegated by KA in order to apply dynamic processing on the data and controlling users. This scheme is secured against Chosen Plaintext Adversary (CPA), assuming the (Decision) Bilinear Diffie-Hellman Exponent assumption (n-BDHE) is being held.

Keywords: Access Control; Attributes Based Encryption; Bounded CipherText; Key Policy; Revocation; Revocable Storage Attribute-based Encryption

1 Introduction

Outsourcing is a movement has been influencing the global revolution of the information technology which gives effective solutions for data managing of the organizations, such as installations, data analysis, networks and data protection. It offers wonderful benefits such as better operating, reducing employment cost, delegating responsibilities to external agencies, as well as mitigating risk and resource's scalability.

Moreover, delegating responsibilities to other party as Cloud Service Party (CSP), there are data owner who still worrying about privacy preserving of their data and how they controlling the accessibility, in order to guarantee secure offshoring.

ABE is one of popular accessing control techniques and has been appeared firstly with Sahai and Waters [2] where they aim to encrypt Ciphertexts one-to-many. However, the users can decrypt if they have certain requirements, although ABE algorithms suffer from two significant drawbacks. For instance, growing of the Ciphertext impractically, and the revoking mechanism of expired or dishonest user.

In fact there are two types of revocations, direct and indirect models [13]. the first scenario revocation is enforced directly by the sender who determines the revoked list during encryption stage, whilst indirect revocation are controlled by the key authority KA which issues an updated key, such that only non-revoked users can update their keys.

We present a novel way of an Indirect R-ABE technique with bounded Ciphertext that overcomes the revocation challenges, such as revoking users without resetting credentials of others users and preventing revoked users from accessing the data or collude with dishonest users.

The challenging areas which have been handled in this work are dynamic controlling of the users and shortening ciphertext, the scheme relays on broadcast encryption technique that proposed in [5] which has collusion resistant and short ciphertext features.

1.1 Related Works

Many Revocable-ABE [1, 4, 9, 17, 18] were introduced recently. However, most of them suffered from the growth of ciphertext's size proportionally, with number of users and attributes. Updating periodically the attributes private keys which is unaccepted for practicable applications, particularly when users have limited resources.

Revocable Storage is a challenging task where the third

party can modify the existing ciphertext to block the repealed users from accessing the stored data in outsourcing storage without intermediating of the data owner while the other user can keep accessing it. Sahai and Waters innovated et al. [3] the first revocable storage when they used only publicly available information with periodically updating the ciphertexts and private keys. However, the size of ciphertext increased linearly with number of associated attributes, also needs to re-distribute periodically the private keys for all non-revoked users.

Nuttapong and Hideki proposed a Conjunctive Broadcast and Attributed Based Encryption, where the private key Conjuncted with a user index and the ciphertext associated also with a user index set S, the decryption can achieve if the condition on attributes of the ABE hold and, in addition, $ID \in S$ and KeyGen used ID with Linear Secret Sharing Schemes, but the size of CT and private keys were large [14].

Junbeom and Dong proposed et al. [11] designed revocable CP-ABE Schemes with periodic or timed revocation with the help of the semi-trusted proxy deployed in the Cloud Services Provider (CSP). The main drawback of these schemes is relayed on other part for re-encryption.

David and Thomas presented et al. [8] a broadcast encryption scheme, with attribute-based mechanisms that lets the Data Owner to add/revoke groups of users were defined by their attributes, also the size of private keys is grown with the number of attributes that are related to the user and size of Ciphertext is also increased linearly with the number of attributes used in the access policy whereas the public key is somewhat large.

1.2 Our Result

This section gives a comparison between state-of-art schemes and our novel approach which realizes shrinking of the private keys's and ciphertext size without influences with number of users or associated attributes, also the performance is enhanced by applying precomputed algorithms and cached the computation in secure memory. Table 1 shows the comparison.

We denote for the parameters of table1 as follows: **U** is universe attributes or all possible attributes in the scheme, **S** is set of attributes that have assigned to the user, **Y** is set of possible attributes have associated to CT, **r** is number of revoked user, **Nmax** the number of leaf nodes in **I** where The total number of all nodes in the circuit is $2Nmax?1$, **SD** is Subset Difference Method, (**RSABE**) Revocable Storage Attribute-based Encryption, **DO**, **nx** is the number of rows that selected by map function p(x) for all x in Y, **aMSE-DDH** augmented multi-sequence of exponents decisional Diffie Hellman problem, **IBBE** identity-based broadcast encryption, M number of total users in the schemes, **COBG** Composite-Order Bilinear Groups.

1.3 Contributions

We proposed a concrete R-ABE with following achievements (I) Short ciphertext and independent from number of users or attributes. (II) The key authority KA has ability to revoke or add users efficiently (III) The revocation processes did not need reseting user credentials or redistributing of the private keys or the public key (only the updated key) (IV) There a proxy re-encryption to prevent the existing data, however, each ciphertext will re-encrypt once before storing it in semi-trust third party TTP (V) The scheme prevents a repealed user from accessing the old ciphertext by modify tiny part of ciphertext (about **25%** of the original ciphertext).

1.4 Organization

Section 2 will present preliminaries and definitions of some security notions, Section 3 describes the scheme's constructions and correctness of the scheme, Section 4 introduces the security game and proving of the system's security, Section 5 presents some enhancement technique and applying precomputed algorithm to improve the cost of system's computations, Section 6 presents implementation and result, and in Section 7 concludes and shows the open problems and future work.

2 Preliminaries and Definitions

This section shows the preliminaries and definitions of some security tools which will use to construct the scheme.

2.1 Bilinear Mapping

We review some facts associated to bilinear map cycle groups which are efficient and computable, introduced by Boneh and Frankin [7], both groups have the same prime order group p, the map function must satisfy the following properties:

Computability: There exist polynomial time algorithm when given $g_1, g_2 \in G$ that can compute $e(g_1, g_2) \in G_T$.

Bilinear: For any $a, b \in Z_p$ the bilinear function is such that $e(g_1^a, g_2^b) = e(g_1, g_2)^{\text{a.b}} \in G_T$ are Non-Degenerate where g is generator of G and $e(g, g)$ generator of G_T where $e(g, g) \neq 1$.

Access structure: Suppose $\{P_1, P_2, \ldots, P_n\}$ is a set of attributes, we say a selection of attribution $S \in 2^{\{P_1, P_2, \ldots, P_n\}}$ is monotone if $\forall B, C$: $B \in A$ and $B \subseteq C$ then $C \in S$, a monotone access structure is a group collection of non-empty subsets $S \in 2^{\{P_1, P_2, \ldots, P_n\}} \backslash \{\emptyset\}$, the authorized sets is in S or qualified set, and the sets are not in S called the unauthorized sets. We emphasize on restriction that using monotone access structures in our system.

Table 1: Comparison between other ABE schemes

Scheme	**[3]**	**[20]**	[16]	**[10]**	[15]	Our
PK	$O(\|U\|)$	6	112	$O(2\|U\|)$	$O(2\|U\|)$	$O(2\|M\|)$
Pr	$O(2\|S\|)$	$2\|S\|+2$	$5+16L+16[log2Nmax]$ $+logNmax]$	$O(\|S\|)$	$O(2\|U\|)$ $+O(\|S\|)$	$O(\|S\|)$
CT	$O(\|Y\|)$	$2\|Y\|+2$	$O(\|RL\|)$	$O(3.nx*L)$	$O(1)$	$O(1)$
Updated	$(Pr+CT)$	$(PK+Pr$ $+MSK)$	$(St+RL)$	Pr	Pr	CT
Security Assumption	COBG	(DBDH)	(DLIN)	COBG	aMSE-DDH	Decisional qBDHE
Security Game	Oracles CPA	Selective CPA	Full CPA	Selective Oracle-CPA	Selective-CPA Non-interactive	Selective CPA
Access Structure	LSSS $L\times n$	LSSS $L\times n$	(SD) LSSS	LSSS $L\times n$	LSSS $L\times n+IBBE$	Fine Grained
Policy	KP-ABE	CP-ABE	KP-ABE	CP-ABE	CP-ABE KP-ABE	KP-ABE
Revocation Delegation	KA to TTP	KA to TTP	KA to DO	KA to AAs	KA only	KA to TTP
Revocation Methodology	Periodically	On Demands	On Demands	Periodically	On Demands	On Demands
Supporting RSABE	Yes	NO	NO	NO	NO	Yes

Access Circuit: Let C be a circuit represents accessing control of attributes holders, which contains mainly from (AND-Gate, OR-Gate) nodes, we denote to $\{att_i\}_{i\in k}$ as the set of attributes which are given to the user k, $Nmax$ is total number of attributes which input to the circuit (leaf nodes), d is a depth of the circuit and equals generally to the number of circuit's layers, $node_x$ is an indexed node which starts from initial node (the root) $node_1$ or the output of circuit down to the last node $node_l$ notice that $l \leq 2Nmax-1$ is total number of nodes in the circuit, a non-leaf nodes are the attribute nodes nod_x where $(l-Nmax) \leq x \geq l$, an input to the $node_x$ are $input(node_x)=(A,B)$ where A and B are the direct inputs to the node, an output of $node_x$ is denoted by $output(node_x)$, namely if $\{att_i\}_{i\in k}$ is set of attributes which assigned to user k so we say $C(\{att_i\}_{i\in k})$ =true obviously if $\{att_i\}_{i\in k}$ satisfied the access circuit C, also any $node_x$ is satisfied if its output is true $C(\{att_i\}_{i\in k})=true|\forall input(node_x)\subseteq \{att_i\}_{i\in k}$.

3 Revocable ABE with Bounded Ciphertext Scheme

The proposed system is contained of six probabilistic algorithms which are setup, keygen, encrypt, re-encrypt, decrypt and updatekey as described in next paragraph assuming that there exist semi Trusted Third Party TTP who is responsible of re-encrypts the data and proceed the revocability tasks which are delegated from Key Authority KA who is accountable for keys managing, figure 1 shows the interactions between parties.

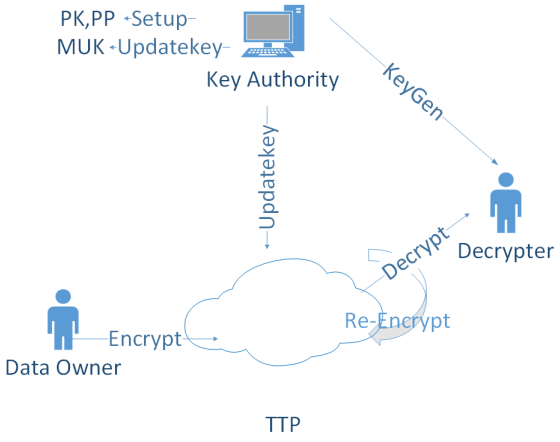


Figure 1: System model

3.1 Scheme Definition

Setup(n,λ)**:** This algorithm runs by KA after inputs the number of total users n with the security parameter λ and publishes out the public key PK, public parameters PP and keeps master secret key MSK secret.

KeyGen $(k,\{Attr\}_{\forall i\in S_k},MSK)$**:** KA takes the user in-

dex $k \in [n]$, set of user's attributes $\{Attr_i\}_{\forall i \in S_k}$ and the master secret key MSK and outputs the private key for each attribute $\{pr_{k,i}\}$.

UpdateKey$(k, \{Attr_i\}_{\forall i \in S_k}, MSK)$**:** The KA uses k the current user index, $\{Attr_i\}_{i \in S_k}$ user's attributes that were assigned to user k and MSK master secret key, this algorithm will output refreshed master updated key MUK and submits the delegation keyDK that can depute TTP to handle either adding or revoking users, an UpDateKey algorithm might run if one of the four actions happened:

1) Adds new user with new attributes and submits the new attributes private key$\{pr_{k',i}\}$ and outs a new index k'.
2) Adds new attributes for existing user k.
3) Revokes existing user k permanently from the scheme.
4) Revokes some Attributes?$\{pr_{k,i}\}$ from user k.

Encrypt(M, PP, PK, MUK)**:** This a probabilistic algorithm works in very straightforward ways by taken the message M, public parameter PP, public key PK and master updated key MUK, the algorithm outputs succinct ciphertext CT_0, we emphasis on the size of ciphertext is not impacted neither number of legitimate user nor valid attributes, size of CT_0 precisely $O(CT_0) = 1$ which offered efficient property.

Re-encrypt(CT_0, PP, PK, MUK)**:** The TTP is allowed to modify the existing ciphertext either for preventing expired users from accessing it or allowing new user to permit accessing these encrypted data, this algorithm runs after TTP received the CT_0 directly or on demand of KA after receiving delegation keys, and outs the CT.

Decrypt$(CT, k, \{pr_{k,i}\}, PP, MUK)$**:** The decrypter uses this algorithm and inputs ciphertext CT, user index k, attributes private keys $\{pr_{k,i}\}$, public parameter PP and master updated key MUK, then it decrypts out the message M.

Correctness: Required that the system to be correct, specifically as follows:

$$\begin{aligned}\Pr[Dec(CT, S, k, \{pr_{k,i}\}, PP, MUK) = M | \forall S, k,\\ (PP, PK, MSK) \leftarrow Setup(n, \lambda),\\ \{pr_{k,i}\} \leftarrow KeyGen(k, \{Attr_i\}, MSK)\\ MUK \leftarrow UpDatekey(k, \{Attr_i\}, MSK)\\ CT_0 \leftarrow Encrypt\,(M, PP, PK, MUK)\\ CT \leftarrow ReEncrypt\,(CT_0, PP, PK, MUK)\,,\\ \forall i \in S_k] = 1.\end{aligned} \tag{1}$$

3.2 Security Game

Revocable ABE with Bounded Ciphertext is secure against selective chosen plaintext adversary CPA, where the security game is made up between an adversary A and a challenger B as follows.

Setup: B selects at the beginning the authorized set S_0, also selects the revoked set S_r such that $S_r \subset S$, then runs setup and UpDatekey algorithms and submits public key PK, public parameter PP and master updated key MUK to A whereas keeps master secret key MSK hidden from A, the adversary selects set of users and submits them to B as challenged set S.

Phase 1: Adversary A is asking adaptively the challenger B queries about attribute private keys for number of users $S_q = \{q_1, q_2, ..., q_r\}$ with one of the two restrictions:

1) **Case 1:** In this case the adversary A chose an user $k \in S_q$, that must not belong to authorized sets $k \notin S_0$ and each attributes $\{att_i\}$ of user k satisfied the access circuit C commonly $C(\{att_i\}_{i \in k}) = true$.
2) **Case 2:** A asks for the user $k \in S_q$ belonged to authorized sets $k \in S_0$ and he/she has been revoked $k \in S_r$. Also user k satisfied the access circuit $C(\{att_i\}_{i \in k}) = true$.

Then challenger obtains attributes private keys by running KeyGen algorithm and responds to adversary A with attributes private keys.

Challenge: After adversary A satisfied from asking queries then will pick up two random messages m_0, m_1 where $|m_0| = |m_1|$ and submits two message to challenger who will toe coin $b \in \{0, 1\}$ and applies encryption algorithm on $CT_b = Encrypt(m_b, PK, PP, MUK)$ and sends CT_b to adversary as challenge.

Phase 2: The adversary A is continuing adaptively queries the challenger β in similar way of Phase1 by sending request for other attributes private keys $S_2 = \{q_{r+1}, q_{r+2}, \ldots, q_m\}$ and we recall same phase1's restrictions.

Guess: Eventually adversary A outs the guessing of b' and wins iff $b = b'$.

3.3 Security Assumption

Our system's security is based on the complexity of (Decisional) Bilinear Diffie-Hellman Exponent Assumption (n–$BDHE$) [6, 19] relays on choosing a symmetric pairing $e\colon G \times G \Rightarrow G_T$ where G is a bilinear multiplicative group of prime order P, G_T is target group of prime order P. The (decisional) $n - BDHE$ problem described when given to an algorithm B this tuples

$(h, g, g_1, g_2, \ldots, g_n, g_{(n+2)}, \ldots, g_{2n}) \in G^{(2n+1)}$ then the algorithm B can output $b \in \{0,1\}$ with advantage ζ, in breaking decisional n–$BDHE$ in G_T if

$$|\Pr[B(h, g, g_1, \ldots, g_n, g_{(n+2)}, \ldots, g_{2n}, e(g_{n+1}, h)) = 0] - Pr[B(h, g, g_1, \cdots, g_n, g_{(n+2)}, \ldots, g_{2n}, T) = 0]| \geq \zeta.$$

With probability over the random choice of generator $g, h \in G$, $\alpha \in Z_p$, $T \in G_T$ and the random bits used by B, the left part of above equation is valid distribution and is denoted V–$BDHE$ and the right part invalid random distribution and denoted R–$BDHE$.

4 Construction

We describe in this section the constructing of revocable ABE, as far we assume there exist Key Authority (**KA**) that in charges for creating users attributes private keys and revokes or adds users, Semi Trusted Third Party (**TTP**) that will re-encrypt the ciphertext and applies revocation or addition of users, Data Owner (DO) and decrypter, all of the above parties are participating as follows.

Setup(n, λ)**:** Setup algorithm is running by KA to generate the public parameters PP, public key PK and master secret key MSK, n is the input for this algorithm which is a number of expected users and λ is security parameter, the algorithm chooses $g \Leftarrow G$ uniformly as generator of source group G and $\alpha, \gamma \Leftarrow Z_p$, we denote $g_k = g^{(\alpha^k)}$ the public parameter is PP to compute public PK, the KA picks random $\beta \Leftarrow Z_p$ and computes the following tuple:

$$\begin{aligned} PP &= (g, g_1, g_2, \ldots, g_n, g_{(n+2)}, \ldots, g_{2n}, v = g^{\gamma}) \in G^{(2n+1)} \\ PK &= \left(g' = g^{\beta}, w = e(g_n, g_1)^{\beta}, S_0\right), \\ MSK &= (\alpha, \gamma, \beta) \end{aligned} \quad (2)$$

where S_0, S are initial authorized and current authorized set respectively, then KA publishes PP and PK while Master Secret key MSK are kept secret.

KeyGen $(k, \{att_i\}_{(i \in S_k)}, MSK)$**:** For each user k the TTP computes $d_k = g^{(\alpha^{k^{\gamma}})} = v^{(\alpha^k)}$ and sets $Y = d_k$ as final output of the circuit and assumes $Y = d_k = g^{(\alpha^{k^{\gamma}})} = g^{(\alpha^y)}$ where $y = k^{\gamma}$ is the final output of the root gate. Now to compute the attributes keys of user k the KeyGen algorithm is inspired from fine-grained structure so if the next gate is OR-Gate it just pass same value to the two next fans and if the next gate is AND-Gate it chooses random $r_{(l,A)} \in Z_p$ uniformly for A's input and sets $r_{(l,B)} = y - r_{(l,A)} \in Z_p$ where l is root gate index, for B's input again it works same as above if the gate is OR-Gate $r_{(l-1,B)} = r_{(l-1,A)} = r_l$ and if the gate is AND-Gate then chooses random $r_{(l-1,A)} \in Z_p$ and sets $r_{(l-1,B)} = r_l - r_{(l-1,A)} \in Z_p$ and it continues in the same way until reaches the inputs of the circuit (leaves) which are the attributes of this circuit $\{r_{(i,j)}\}_{i \in [m], j \in \{A,B\}}$. KeyGen algorithm computes private key of each attribute as $\{g^{(\alpha^{r_{(i,j)}})}\}_{i \in [m], j \in \{A,B\}}$ and sends attributes private keys $S_k \subseteq S$ of user k and sends to k via secured channel the values $\{g^{(\alpha^{r_{(i,j)}})}\}_{i \in ([m] \cap S_k), j \in \{A,B\}}$. We denote the attribute private of user k as $pr_{(k,i)} = g^{(\alpha^{r_{(i,j)}})}$.

UpDatekey $(k, \{att_i\}_{(i \in S_k)}, PK, MSK)$**:** This algorithm runs by Key Authority when decides to revoke certain user u or adding new user k' or after setup algorithm the output of this is master updated key (MUK) and delegated key DK which will be submit to TTP as follows.

$$MUK = \left(S, v' = (v. \prod_{(\forall j \in S)} g_{(n+1-j)})^{\beta}\right) \quad (3)$$

$$DK = (\{g^{-\beta}_{n+1-u_i}\}_{i \in S_r}, \{g^{\beta}_{n+1-k'_i}\}_{i \in S_a}) \quad (4)$$

where S_r, S_a are set of revoked and added users list respectively, then DK will send to TTP as delegated key to run Re-encrypt algorithm and updates the existing ciphertext for modification and publishes MUK. Our scheme is flexible for efficient key management processing in the following way:

1) Removing all user's attributes (revoke an user u), in this case the key authority refreshes MUK and updates v' in particularly

$$v' \leftarrow \left(\frac{v'}{g_{n+1-u}}\right)^{\beta} \quad (5)$$

simultaneously in other side the TTP will update (small part only C_1) the existing ciphertext as follows:

$$\begin{aligned} C_1 \Leftarrow \left(C_1 \cdot (DK_u)^{t'}\right) &= \left(C_1 \cdot (g^{-\beta}_{n+1-u})^{t'}\right) \\ &= \left(\frac{C_1}{\left(g^{\beta}_{n+1-u}\right)^{t'}}\right) \end{aligned}$$

Here $DK_u \in Dk$ is assigned to user u. TTP will update some part of the ciphertext for existing data, however for future encryption the TTP is preventing from adding revoked user with two safe guards MUK and DK, also re-encryption algorithm guaranties forward and backward secrecy.

2) Removing some part of user's attributes in this case the key authority first applies the above step (remove the user u) then run KeyGen algorithm again to create keys as a new user with new index k' and this index must be unique

$k' \notin S$ and for allowing the new user k' to decrypt previous ciphertext update some part of ciphertext as follows:

$$\begin{aligned} C_1 \Leftarrow \left(C_1 \cdot (DK)^{t'}\right) &= C_1 \cdot \left(g^{\beta}_{n+1-k'} \cdot g^{-\beta}_{n+1-u}\right)^{t'} \\ &= C_1 \cdot \left(\frac{g^{\beta}_{n+1-k'}}{g^{\beta}_{n+1-u}}\right)^{t'} \end{aligned}$$

3) For adding a new user the key authority runs KeyGen algorithm for obtaining attributes private keys and submits the keys through secure channel.

Encrypt(M, PK, MUK)**:** Data owner intends to encrypt the data before outsourced in the TTP environment, starts by chooses random $t \in Z_p$ uniformly, and inputs the plaintext $M \in G_T$ then computes

$$CT = (C_0, C_1, C_2) = \left(g'^t, (v')^t, M.(w)^t\right) \quad (6)$$

Re-encrypt(CT, PK, PP, MUK)**:** This algorithm runs by TTP after receiving the ciphertext this step can assist to revoke the expired user u or adds new user k' to the scheme where u, k' are the index of revoke and added user respectively. Then the TTP will select randomly $t' \in Z_p$ and recomputes the ciphertext:

$$\begin{aligned} C_1 &\leftarrow \left(C_1.v'^{t'}\right) = v'^t.v'^{t'} = v'^{(t+t')} \quad (7) \\ C_2 &\leftarrow \left(C_2.w^{t'}\right) \end{aligned}$$

Note that there are difference between the authorized current set S which were chosen to compute v' and the authorized initial set S_0, so S is used to re-encryption whereas S_0 is used for encryption, is obviously to notice that the exponent of ciphertext t is shifted to $t + t'$ with this algorithm as follows.

$$C_0 = g'^t, C_3 = g'^{t'}. \quad (8)$$

From above we realize no changing in C_0 and TTP added new part the ciphertext C_3

$$\begin{aligned} C_1 &= \left(v. \prod_{(\forall j \in S_0)} g_{(n+1-j)}\right)^{\beta.t} \cdot \left(v. \prod_{(\forall j \in S)} g_{(n+1-j)}\right)^{\beta.t'} \\ C_2 &= \left(C_2.w^{t'}\right) = \left(M.w^t.w^{t'}\right) = \left(M.w^{t+t'}\right). \end{aligned}$$

We emphasize this algorithm is run once so will not effect the performance of scheme.

Decrypt$(CT, k, \{pr_{k,i}\}, PP, MUK)$**:** If the decrypter k has enough attribute's private keys that can fulfill the circuit's requirement, then user k is capable to compose d_k and decrypts the ciphertext $CT = (C_0, C_1, C_2)$ according to the ABE circuit, the decrypter starts in reverse way beginning from the circuit's input (leaves) until final gate(root), at first k inputs the attributes private keys $pr_{(k,i)} = g^{(\alpha^{r_{(i,j)}})}$ then if the gate is OR-Gate then chooses any one of A or B as input of the gates (leaves)

$$\{y_{i+1} = g^{(\alpha^{r_{(i,A)}})} or = g^{(\alpha^{r_{(i,B)}})}\}_{i \in ([m] \cap S_k)}$$

where S_k is set of attributes belong to user k. Or the gate might be AND-Gate, then multiply the two inputs as

$$\begin{aligned} \{y_{i+1} &= g^{(\alpha^{r_{(i,A)}})} \times g^{(\alpha^{r_{(i,B)}})} \\ &= g^{(\alpha^{r_{(i,A)}}) + (\alpha^{r_{(i,B)}})}\}_{i \in ([m] \cap S_k)} \end{aligned}$$

for remaining gates acts in same way namely for OR-Gate:

$$\begin{aligned} \{y_{i+1} &= y_{(i,A)} = g^{(\alpha^{r_{(i,A)}})} \text{ or} \\ &= y_{(i,B)} = g^{(\alpha^{r_{(i,B)}})}\}_{\forall i \in \{(m+1),\ldots,(l-1)\}} \end{aligned}$$

and for AND-Gate the decrypter follows the circuit rules and computes:

$$\begin{aligned} \{y_{i+1} &= y_{(i,A)} \times y_{(i,B)} \\ &= g^{(\alpha^{r_{(i,A)}})} \times g^{(\alpha^{r_{(i,B)}})} \\ &= g^{(\alpha^{r_{(i,A)} + r_{(i,B)}})}\}_{\forall i \in \{(m+1),\ldots,(l-1)\}} \end{aligned}$$

until reaches final gate $Y = g^{\alpha^y} = g^{(\alpha^{(k^\gamma)})} = d_k$. Hence decrypter gets d_k.

This above steps run once and not in each decryption processing and decrypter will store d_k in secure place, then to decrypt CT which has been re-encrypted with TTP:

$$\begin{aligned} T &= \frac{e(g_k, C_1)}{e(d_k \cdot \prod_{j \in S_0, j \neq k} g_{n+1-j+k}, C_0)} \\ &\quad \times \frac{1}{e(d_k \cdot \prod_{j \in S, j \neq k} g_{n+1-j+k}, C_3)} \end{aligned}$$

Recall the ciphertext CT is composed from:

$$\begin{aligned} C_0 &= g^{\beta(t)}, \\ C_1 &= \left(v. \prod_{(\forall j \in S_0)} g_{(n+1-j)}\right)^{\beta.t} \cdot \left(v. \prod_{(\forall j \in S)} g_{(n+1-j)}\right)^{\beta.t'} \\ C_2 &= M \cdot e\left(g_{n+1}, g\right)^{\beta \cdot t} \cdot e\left(g_{n+1}, g\right)^{\beta \cdot t'}, C_3 = g^{t'} \end{aligned}$$

First we reduce the numerator of decryption equation:

$$\begin{aligned} T &= \frac{e\left(g_k, \left(v. \prod_{(\forall j \in S_0)} g_{(n+1-j)}\right)^{\beta.t}\right)}{e\left(d_k \cdot \prod_{j \in S_0, j \neq k} g_{n+1-j+k}, C_0\right)} \\ &\quad \times \frac{e\left(g_k, \left(v. \prod_{(\forall j \in S)} g_{(n+1-j)}\right)^{\beta.t'}\right)}{e\left(d_k \cdot \prod_{j \in S, j \neq k} g_{n+1-j+k}, C_3\right)} \end{aligned}$$

$$= \frac{e\left(g_k, g_{n+1-k}^{\beta\cdot t}\right)\cdot e\left(g_k, \left(v.\prod_{(\forall j\in S_0, j\neq k)} g_{(n+1-j)}\right)^{\beta.t}\right)}{e\left(d_k\cdot\prod_{j\in S_0, j\neq k} g_{n+1-j+k}, C_0\right)} \times \frac{e\left(g_k, \left(v.\prod_{(\forall j\in S)} g_{(n+1-j)}\right)^{\beta.t'}\right)}{e\left(d_k\cdot\prod_{j\in S, j\neq k} g_{n+1-j+k}, C_3\right)}$$

$$= \frac{e\left(g, g_{n+1}\right)^{\beta\cdot t}\cdot e\left(g_k, \left(v.\prod_{(\forall j\in S_0, j\neq k)} g_{(n+1-j)}\right)^{\beta.t}\right)}{e\left(d_k\cdot\prod_{j\in S_0, j\neq k} g_{n+1-j+k}, g^{\beta\cdot t}\right)} \times \frac{e\left(g_k, \left(v.\prod_{(\forall j\in S)} g_{(n+1-j)}\right)^{\beta.t'}\right)}{e\left(d_k\cdot\prod_{j\in S, j\neq k} g_{n+1-j+k}, g^{\beta\cdot t'}\right)}$$

$$= \frac{e\left(g, g_{n+1}\right)^{\beta\cdot t}\cdot e\left(g_k, \left(v.\prod_{(\forall j\in S_0, j\neq k)} g_{(n+1-j)}\right)\right)^{\beta.t}}{e\left(g^{\gamma}\cdot\prod_{j\in S_0, j\neq k} g_{n+1-j}, g_k\right)^{\beta.t}} \times \frac{e\left(g_k, \left(v.\prod_{(\forall j\in S)} g_{(n+1-j)}\right)\right)^{\beta.t'}}{e\left(g^{\gamma}\cdot\prod_{j\in S, j\neq k} g_{n+1-j}, g_k\right)^{\beta.t'}}$$

$$= \frac{e\left(g, g_{n+1}\right)^{\beta\cdot t}\cdot e\left(g_k, \left(v.\prod_{(\forall j\in S)} g_{(n+1-j)}\right)\right)^{\beta.t'}}{e\left(v\cdot\prod_{j\in S, j\neq k} g_{n+1-j}, g_k\right)^{\beta\cdot t'}}$$

The right part of equation will be reduce as same manner:

$$T = \left(g, g_{n+1}\right)^{\beta\cdot t}\cdot\left(g, g_{n+1}\right)^{\beta\cdot t'}$$

then the decrypter can obtain the plaintext by computes $M = \frac{C_2}{T}$ in blow equation.

$$M = \frac{C_2}{T} = \frac{M.e\left(g_{n+1}, g\right)^{\beta.t}.e\left(g_{n+1}, g\right)^{\beta.t'}}{e\left(g_{n+1}, g\right)^{\beta.t}.e\left(g_{n+1}, g\right)^{\beta.t'}}$$

5 Security

In the following theorem we prove semantic security of the Bounded R-ABE scheme assuming the hardness of the (Decisional) n-BDHE assumption holds, which are:

Theorem 1. *let G be bilinear group of order p where p is prime and $n > 1$ our proposed Bounded R-ABE scheme is (n) semantically secure if the decision n-BDHE assumption holds in G_T.*

Proof. Assuming there exists PPT adversary algorithm A that can breakdown our scheme with advantage $AdvBRABE_(A, n) > \zeta$ in time t, also there exist algorithm B has advantage ζ to break n-BDHE problem in G_T, B calls algorithm A which selects the set S of users that A wishes to be challenged on.

Setup: B selects the initial authorized set S_0, current set S_c, chooses randomly $g, \alpha, \gamma, \beta \in Z_p$, then computes PP and publishes the public keys PK as $PP = (g, g_1, g_2, \ldots, g_n, g_{(n+2)}, \ldots, g_{2n}, v = g\cdot\left(\prod_{j\in S_c} g_{n+1-j}\right)^{-1}$, where $PK = (g' = g^{\beta}, w = e(g_n, g_1)^{\beta}, S_0))$ hence g, α, β, γ were chosen randomly then PK and PP have uniform distribution same as original scheme. $MUK = (S_c, v')$ where $v' = \left(v\cdot\prod_{j\in S} g_{n+1-j}\right)^{\beta}$ as in Equation (3).

Phase 1: The adversary A asks the algorithm B in this phase for attribute's private keys of users $S_q = \{q_1, q_2, \ldots, q_r\}$ recall that there are two possible scenarios:

1) Recall case1 when the user's index is not in authorized set S_0 such that $\forall i\ q_i \in S_q$ and $q_i \notin S_0$, and each attributes $\{att_i\}$ of user k satisfied the access circuit $C(\{att_i\}_{i\in k}) = true$. In this case the algorithm B computes:

$$\begin{aligned} d_k &= g_k\cdot\left(\prod_{j\in S_c} g_{n+1-j+k}\right)^{-1} \\ &= \left(g\cdot\left(\prod_{j\in S_c} g_{n+1-j+k}\right)^{-1}\right)^{\alpha^k} \\ &= v^{\alpha^k} \end{aligned}$$

Challenger B sets $y = k^{\gamma}$ then follows fine-grained tree to compute $\left\{r_{(i,j)}\right\}_{i\in Att_k, k\notin S, j\in\{A,B\}}$ similar to original scheme and then responds to A With attributes private keys $\left\{pr_{(k,i)} = g^{(\alpha^{r_{(i,j)}})}\right\}_{i\in Att_k, j\in\{A,B\}}$ such that:

$$C\left(\left\{pr_{(k,i)} = g^{(\alpha^{r_{(i,j)}})}\right\}_{i\in Att_k, j\in\{A,B\}}\right) = true$$

Note that the output for root node is $output(node_1) = d_k$.

2) In other hand for Case 2 when A is asking for the user $k \in S_q$ belonged to authorized sets $k \in S_0$ and he/she has been revoked $k \in S_r$. Also user k satisfied the access circuit $C(\{att_i\}_{i\in k}) = true$. Then B will update MUK.

Then B continues in computing the attribute private keys as in case1.

Challenge: After adversary A finished from the query phase then will submit to B two equal random messages m_0, m_1 where $|m_0| = |m_1|$, so B choses $\beta \in Z_p$ and toes fair coin $b \in \{0, 1\}$ to select one message, then B will simulate the running of encrypt and re-encrypt algorithms sequentially on the message $CT = Re - Encrypt(encrypt(m_b))$, while picks random value for $CT_{1-b} \in G_T$ and computes $C_{2,b} =$

$m_b \cdot e\left(g_{n+1}, h\right)$ WOLOG:

$$\begin{cases} C_{2,1-b} \in G_T \\ \quad \text{if ciphertext is chosen randomly(invalid))} \\ C_{2,b} = m_b \cdot e\left(g_{n+1}, h\right) \text{ is valid n-BDHE.} \end{cases}$$

where $h = g^{\beta(t+t')}$ for the remaining ciphertext, B computes C_1 similar to the real scheme as in Equation (7) recall $C_1 = v'^{(t'+t)}$ and from the simulated value v' as in Equation (7) then C_1 is computed as follows:

$$\begin{aligned} C_1 &= (v')^{(t+t')} \\ &= \left(v \cdot \prod_{j \in S_c} g_{n+1-j}\right)^{\beta(t+t')} \\ &= \left(g \cdot \left(\prod_{j \in S_c} g_{n+1-j}\right)^{-1} \cdot \prod_{j \in S_c} g_{n+1-j}\right)^{\beta(t+t')} \\ &= g^{\beta(t+t')} \\ &= h. \end{aligned}$$

For other part of the ciphertext $C_0 = g'^t = g^{\beta \cdot t}, C_3 = g^{\beta \cdot t'}$ from Equation (8), then the algorithm B will submit the challenging ciphertext CT to A where $CT = (C_0, C_1, C_{2,0}, C_{2,1}, C_3)$.

Phase 2: The game between A and B will play identically as in phase1 with same restrictions.

Guess: Eventually the Algorithm A submits out b' guessing of the challenging ciphertext if $b' = b$ then B outs 0 showing that $T = C_{2,b} \div m_b = m_b \cdot e(g_{n+1}, h) \div (m_b) = e\left(g_{n+1}, h\right)$, else B outs 1 and that refers T is chosen randomly in G_T, Note that $|\Pr[B(h, g, g_1, \ldots, g_n, g_{(n+2)}, \ldots, g_{2n}, e(g_{(n+1)}), h)) = 0] - \Pr[B(h, g, g_1, ., g_n, g_{(n+2)}, \ldots, g_{2n}, T) = 0]| \geq \zeta$, is same to (Decisional)n-BDHE assumption and that is the proof of Theorem 1.

□

6 Implementation and Result

This section examines the system performance and tests the computations complexities, also we resolve the problem of complexity increasing significantly with number of users by applying powerful tool of pre-computation method.

6.1 Enhancing Performance

The proposed scheme has a dramatic increase of computation when KA is updating the keys or during user is decrypting the ciphertext which they need to multiply about $|S_0| + |S|$ times for every process and that consumes the resources especially on the users's side which are limited resources, so to overcome this problem we implement cached algorithm in both side without impacts the security, following equations shows the caching steps as:

1) When MUK is updated KA runs UpDateKey algorithm obtaining $MUK = \left(S, v' = (v. \prod_{(\forall j \in S)} g_{(n+1-j)})^{\beta}\right)$ to reduce the overhead multiplication KA can pre-compute $v' = \left(v. \prod_{(\forall j \in S)} g_{(n+1-j)}\right)^{\beta}$ and stores v' in cache memory so the MUK will reconstruct from v' and if the kA intends to add new user k' in this case will update only $v' = v'.g^{\beta}_{(n+1-k')}$, and in case of revoking existing user u then $v' = v'/g^{\beta}_{(n+1-k)}$, this will lead efficient calculation for MUK and reduces the computation cost from $O(|S|)$ to $O(1))$ for each time we run UpDateKey.

2) When decrypter user aims to decrypt some ciphertext according to decryption algorithm.

$$\begin{aligned} T &= \frac{e\left(g_k, C_1\right)}{e\left(d_k \cdot \prod_{j \in S_0, j \neq k} g_{n+1-j+k}, C_0\right)} \\ &\quad \times \frac{1}{e\left(d_k \cdot \prod_{j \in S, j \neq k} g_{n+1-j+k}, C_3\right)} \end{aligned}$$

There exit overhead computation on the client side who also has limited resource so this calculation can consume huge part from decrypter resources, again to handle this problem we apply caching algorithm by the client k as pre-computing parameter.

$$\begin{aligned} z_1 &= \prod_{j \in S_0, j \neq k} g_{n+1-j+k}, \\ z_2 &= \prod_{j \in S, j \neq k} g_{n+1-j+k} \end{aligned}$$

and stores z in fast cache memory, then for each decryption process the client k computes:

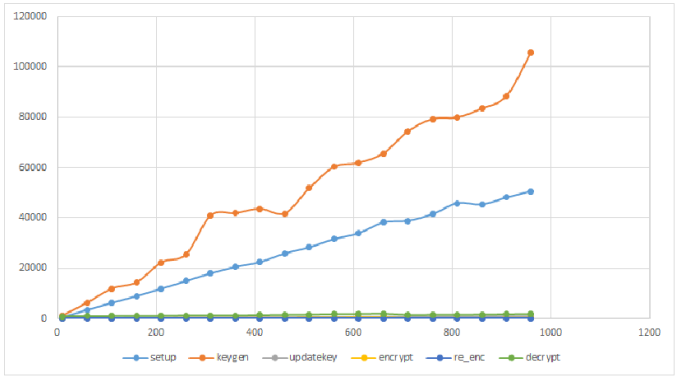
$$T = \frac{e\left(g_k, C_1\right)}{e\left(d_k \cdot z_1, C_0\right)} \times \frac{1}{\cdot e\left(d_k \cdot z_2, C_3\right)}$$

which again minimizes the overhead computing from $O(|S_0| + |S|)$ to $O(1))$.

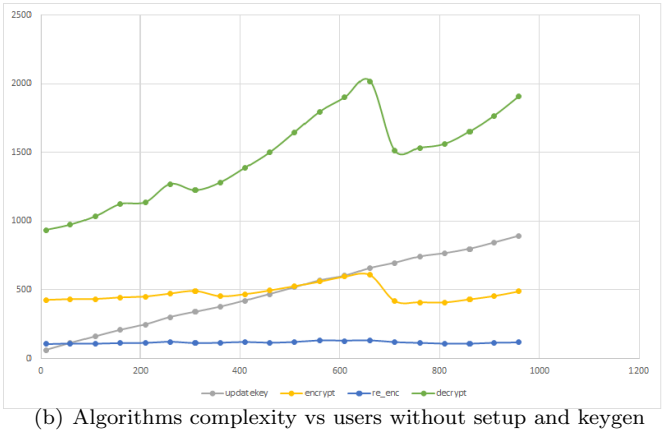
3) There also overhead computation and communication between the KA and TTP when is sending the delegated keys DK, suppose there are many users were wanted to revoke and adds so each time KA sends:

$$DK = (\{g^{-\beta}_{n+1-u_i}\}_{i \in S_r}, \{g^{\beta}_{n+1-k'_i}\}_{i \in S_a}),$$

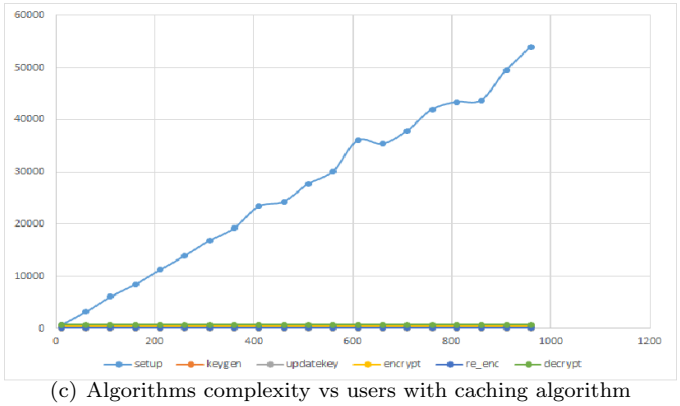
so we need about $O(|S_a| + S_r)$ iterations for computation and communication in both side (KA and



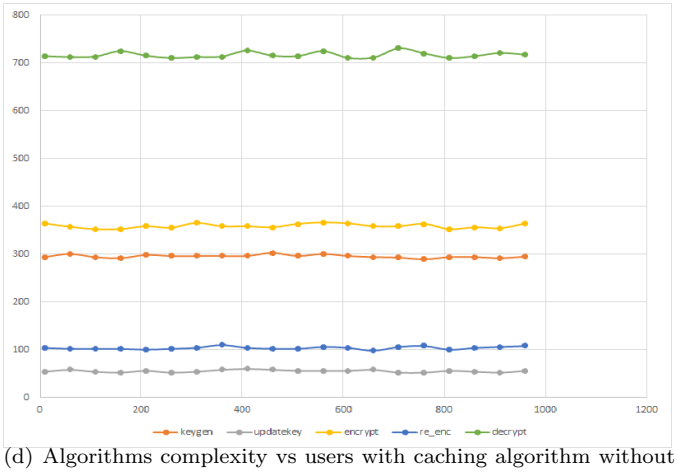
(a) Algorithms complexity vs users



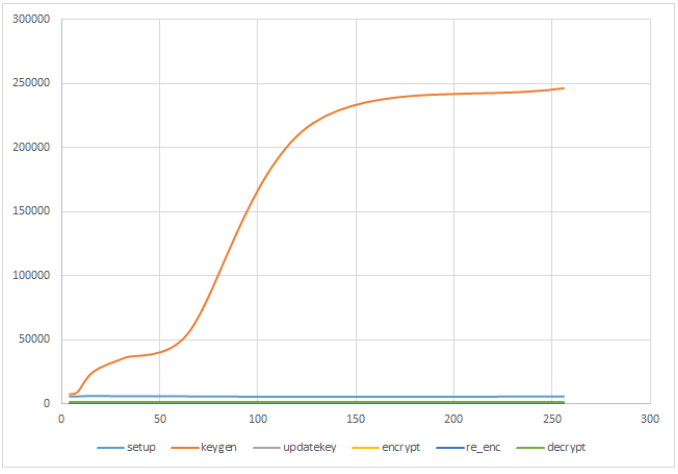
(b) Algorithms complexity vs users without setup and keygen



(c) Algorithms complexity vs users with caching algorithm



(d) Algorithms complexity vs users with caching algorithm without setup



(e) Algorithms complexity vs users with caching algorithm

Figure 2: Comparison of computational efficiency

TTP) and that can be reduced by mixing all in one:

$$DK = \left(\prod_{i \in S_r} g_{n+1-u_i} \cdot \prod_{i \in S_a} g_{n+1-k'_i} \right)^{\beta}$$

6.2 Implementation

We demonstrate the proposed scheme and analyze the performance, we uses the useful MIRACLE library and runs under visual stdio 2012 C++ platform [12].

Figure a shows the growing of users affections only with setup, keygen, updatekey and decrypt (small affections). In figure b setup and keygen were hidden to presents there small correlation in decryption process.

Figure c shows the enhancing of system and powerful reduction of algorithms's complexity when we apply caching algorithm and that leads most of algorithms are running in few computation cost except setup algorithm, the setup algorithm is committed in figure d and is clear that all algorithms are running independently from number of users with low cost.

Setup algorithm is effected only with increasing exponentially with number of attributes as in figure e, whereas the remained algorithms are not effected.

7 Conclusion

We could be concluded that R-ABE with bounded ciphertext has short ciphertext and private keys in addition low computations complexity in both sides client users (encrypter and decrypter) and could be operated in limited resources environment, also we overcome the updating private keys problem, and we avoid the obstacle of stateless problem. The open problem to reduce the large size of the public keys and in our future work also we will intend to design multi key authorities R-ABE.

Acknowledgments

This work was supported by National Natural Science Foundation of China under Grant No 61370026 and the National High Technology Research and Development Program of China (863) under Grant 2015AA016007, the Major International (Regional) Joint Research Project of China National Science Foundation under grant No.61520106007 and the science and technology foundation of Sichuan Province under Grant 2014GZ0109.

References

[1] B. Alexandra, G. Vipul, and K. Virendra, "Identity-based encryption with efficient revocation," in *Proceedings of the 15th ACM Conference on Computer and Communications Security*, pp. 417–426, 2008.

[2] S. Amit and W. Brent, "Fuzzy identity-based encryption," in *Advances in Cryptology (EUROCRYPT'05)*, pp. 457–473, 2005.

[3] S. Amit, S. Hakan, and W. Brent, "Dynamic credentials and ciphertext delegation for attribute-based encryption," in *Advances in Cryptology (CRYPTO'12)*, pp. 199–217, 2012.

[4] N. Dalit, N. Oni, and L. Jeff, "Revocation and tracing schemes for stateless receivers," in *Advances in Cryptology (CRYPTO'01)*, pp. 41–62. Springer, 2001.

[5] B. Dan, G. Craig, and W. Brent, "Collusion resistant broadcast encryption with short ciphertexts and private keys," in *Advances in Cryptology (CRYPTO'05)*, pp. 258–275, 2005.

[6] B. Dan and K. Jonathan, "Improved efficiency for cca-secure cryptosystems built using identity-based encryption," in *Topics in Cryptology (CT-RSA'05)*, pp. 87–103, 2005.

[7] B. Dan and F. Matt, "Identity-based encryption from the weil pairing," in *Advances in Cryptology (CRYPTO'01)*, pp. 213–229, 2001.

[8] L. David and S. Thomas, "Attribute-based broadcast encryption scheme made efficient," in *Progress in Cryptology (AFRICACRYPT'08)*, pp. 325–342, 2008.

[9] Z. Fengli, L. Qinyi, and X. Hu, "Efficient revocable key-policy attribute based encryption with full security," in *IEEE Eighth International Conference on Computational Intelligence and Security (CIS'12)*, pp. 477–481, 2012.

[10] C. Hui and D. Robert, "Revocable and decentralized attribute-based encryption," *The Computer Journal*, vol. 59, no. 8, pp. 1220–1235, 2016.

[11] H. Junbeom and N. Dong, "Attribute-based access control with efficient revocation in data outsourcing systems," *IEEE Transactions on Parallel and Distributed Systems*, vol. 22, no. 7, pp. 1214–1221, 2011.

[12] MIRACL, *Users Manual Shamus Software Ltd*, No. 4 Foster Place North, Aug. 2006. (`http://docs.miracl.com`)

[13] A. Nuttapong and I. Hideki, "Attribute-based encryption supporting direct/indirect revocation modes," in *Cryptography and Coding*, pp. 278–300, 2009.

[14] A. Nuttapong and I. Hideki, "Conjunctive broadcast and attribute-based encryption," in *Pairing-Based Cryptography (Pairing'09)*, pp. 248–265, 2009.

[15] A. Nuttapong, H. Javier, L. Abien, L. Benoît, D. Panafieu, and R. Carla, "Attribute-based encryption schemes with constant-size ciphertexts," *Theoretical Computer Science*, vol. 422, pp. 15–38, 2012.

[16] D. Pratish, D. Ratna, and M. Sourav, "Adaptively secure unrestricted attribute-based encryption with subset difference revocation in bilinear groups of prime order," in *International Conference on Cryptology in Africa*, pp. 325–345, 2016.

[17] M. Silvio, "Efficient certificate revocation," US Patent 6,487,658, Nov. 26 2002.

[18] A. William, L. Sachin, and O. Rafail, "Fast digital identity revocation," in *Advances in Cryptology (CRYPTO'98)*, pp. 137–152, 1998.

[19] D. Yevgeniy and Y. Aleksandr, "A verifiable random function with short proofs and keys," in *Public Key Cryptography (PKC'05)*, pp. 416–431, 2005.

[20] X. Zhiqian and M. Keith, "Dynamic user revocation and key refreshing for attribute-based encryption in cloud storage," in *IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom'12)*, pp. 844–849, 2012.

Biography

Mohamed Ali Hamza received the master degree in the School of Computer Science and Engineering from the University of Electronic Science and Technology of China (UESTC) in Dec 2013, Now he is Ph.D candidate, His research areas in cryptography and information security.

Jianfei Sun is pursuing his Master degree from the Department of Information and Software Engineering, University of Electronic Science and Technology of China (UESTC). His current research interests include cryptographic protocols 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.

Hu Xiong received his PhD degree in the School of Computer Science and Engineering from the University of Electronic Science and Technology of China (UESTC) in Dec 2009. He is currently State Key Laboratory of Cryptology, P. O. Box 5159, Beijing, 100878, China and an associate professor in the School of Information and Software Engineering and School of Computer Science and Engineering, UESTC. His research interests include cryptographic protocols and network security.

Protecting Large Size Medical Images with Logistic Map Using Dynamic Parameters and Key Image

M. Y. Mohamed Parvees[1], J. Abdul Samath[2], and B. Parameswaran Bose[3]

(Corresponding author: M.Y. Mohamed Parvees)

Research and Development Centre, Bharathiar University[1]
Maruthamalai Rd., Coimbatore-641046, India
(Email: yparvees@gmail.com)
Department of Computer Science, Government Arts College, Elayamuthur Rd., Udumalpet-642126, India[2]
No. 35, 1st Main, 3rd Cross, Indiragandhi St., Udaynagar, Bangalore-560016, Karnataka, India[3]

(Received July 22, 2016; revised and accepted Nov. 15 & Dec. 25, 2016)

Abstract

This paper presents a faster and efficient algorithm to encrypt large size 24-bit color medical images. The algorithm generates permutation sequences which shuffles the image pixels and color bytes. The generated masking sequences alters the value of the pixels and color bytes. The dynamic chaotic parameters initialisation is one of the key factors in the proposed cryptosystem. Further, the key image enhances efficiency of the encryption algorithm and the whole cryptosystem becomes complex by executing XOR operation with the shuffled pixels and bytes. The proposed algorithm yields final cipher images of the test medical images that have good confusion - diffusion properties and validated through series of tests to ensure the high security level.

Keywords: Chaotic Map; Immunohistochemical Images; Logistic Map; Medical Image Encryption

1 Introduction

Telemedicine and E-health are becoming more popular at times as there is a necessity for quick and secure diagnosis of diseases by transmitting and storing information among the experts. Since the internet is open to everyone, data protection is most important thing to avoid unauthorized access. The rapid growth of internet and E-techniques allows transmitting large files such as image, video and audio files. The medical experts, educationists and particularly insurance companies need methods to protect the patient's medical information [9, 10, 20]. The patient's information should be accessed only by the authorized personnel during the transmission and storage.

The microscopic colour images such as immunohistochemical (IHC) images have been transmitted among the experts for confirming the status of patient's disease. Further, the experts or doctors always need a second opinion to confirm the absolute status of the disease. Similarly, the health insurance companies wanted to know the real information of the patients by transmitting the images to the experts and getting opinion from them [2, 24, 32, 33]. The main idea is to protect the large medical images (colour) during the transmission and storage using chaotic maps. Since the telepathology and high-content whole slide imaging (WSI) techniques (virtual slides) are growing rapidly [5], the large size IHC images are considered for the analysis of proposed algorithm.

Chaos theory has elucidated considerable interest in computer science for several years because of its applications in the field of cryptography where the traditional algorithms DES, 3DES, RSA are lacking to encrypt bulky data efficiently [7, 18]. Many cryptographic protocols are designed in the literature and chaos theory was proposed in 1989 [17]. Chaos has lot of important qualities *i.e.* high reactive to initial conditions, aperiodicity and topological transitivity [30]. Further, the chaos theory can be use in symmetric key encryption due to their better computational performance over public key encryption [14]. Chaos theory involves in encryption based on the two processes, one is 'confusion' and another one is 'diffusion'. The chaotic maps and matrix manipulations are useful in generating confusion and diffusion processes [11]. In the chaotic image encryption, some of the algorithms with one round of shuffling and diffusion are also reported [6]. But for a cipher image with good confusion and diffusion property, the processing round should be more than three rounds [1]. So far, several researchers have been proposed different chaotic models to encrypt images using single choatic map[22], combinations of one or two maps[21], higher order maps [25] and even the chaotic

maps are combined with other techniques [19] to provide higher level of security. In this study, the chaotic image protection scheme is intended to encrypt immensely colossal size medical images using a Logistic map with dynamic parameters. The large size immunohistochemical images have been analyzed towards the proposed encryption algorithm.

2 The Mathematical Background of Logistic Map

The Logistic map is a simple one dimensional map and efficient in studying nonlinear dynamic systems which can able to reveal chaos behaviour. It is defined as follows:

$$x_{n+1} = r \times x_n \times (1 - x_n), x_n \in (0, 1), r \in (0, 4) \qquad (1)$$

where, x_n is an independent variable; r is the control parameter of logistic map; $n = 1, 2, 3 \ldots$. When the value of the control parameter lies between 3.5699456 and 4 that is, $(3.5699456 < r \leq 4)$ and $x_n \in (0, 1)$, this logistic map is chaotic which can able to produce n length sequences.

A bifurcation diagram represents the possible period orbits of a chaotic map based on the bifurcation parameter values. For a diminutive change in the bifurcation parameter values the possible periodic orbits of the logistic map represented by Equation (1) will be more complicated. From the bifurcation diagram the X-axis represents the bifurcation parameter and Y-axis represents the possible population values x of the logistic map function. As the control parameter value of the Logistic map increases, the bifurcation occurs. From the Figure 1 it is possible to say that many bifurcations occurred for the value of r lies between 3.6 to 4.0. The positive Lyapunov number of the Logistic map influences the non periodic orbits significantly which is shown in Figure 2.

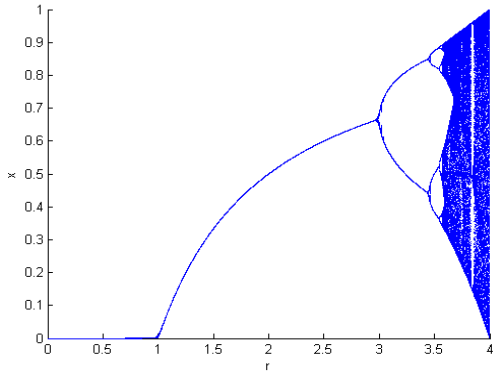


Figure 1: Bifurcation diagram of logistic map

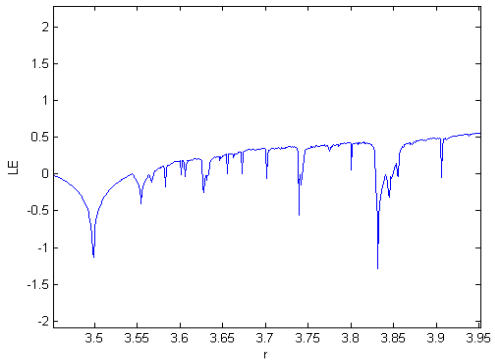


Figure 2: Lyapunov exponent diagram of logistic map

3 The Proposed Cryptosystem

3.1 Dynamic Control and Initial Parameters

In the proposed image encryption algorithm, the eight different dynamic control and initial parameters helps to generate chaotic sequences, thereby scrambles the image pixels efficiently. The eight dynamic control and initial parameters are generated from the two master dynamic control and initial parameters.

The master parameter $r_1 = 3.612345678901234$ and $x_1 = 0.112345678901234$ are used to generate a chaotic sequence $O = \{o_1, o_2, o_3, .o_{10000}\}$ to get dynamic control parameters by using Equation (1). Similarly, the master parameter $r_2 = 3.712345678901234$ and $x_2 = 0.212345678901234$ are used to generate a chaotic sequence $I = \{i_1, i_2, i_3, \ldots, i_{10000}\}$ to get dynamic initial parameters by using Equation (1).

The sequence element (o_{n*s}) is found and checked for sequence element $(o_{n*s}) \geq 0.6$, if true, $dr_n = 3.0 + o_{n*s}$ else moved to o_{n*s+1} until the sequence element value is ≥ 0.6 and added with 3.0 to get dr_n. The sequence element i_{n*s} is found and assigned as $dx_n = i_{n*s}$. Since the number of parameters are eight, the technique produces 8 different control $(dr_1, dr_2, dr_3, dr_4, dr_5, dr_6, dr_7, dr_8)$ and initial $(dx_1, dx_2, dx_3, dx_4, dx_5, dx_6, dx_7, dx_8)$ parameters. The 8 dynamic chaotic parameters are used to produce the permutation and masking sequences which are mainly used for image encryption. By varying the values of $S(diff), r_1, x_1, r_2, x_2$, it is possible to generate entirely different control and initial parameters set. The pseudocode to generate the dynamic control and initial parameters are shown in Algorithm 1.

3.2 Chaotic Sequence

A chaotic map produces non-converging and non-periodic sequences while assigning values for the initial and control parameters. These sequences are termed as chaotic sequences. By using the real valued chaotic sequences, it is able to produce an integer valued sequence. This se-

Algorithm 1 Generation of dynamic chaotic parameters

BEGIN
Intitialise the chaos parameters
SET $no\_params \leftarrow 8$
SET $diff \leftarrow 888$
SET $seq\_no \leftarrow 100000$
SET $c\_miu \leftarrow 3.612345678901234$
SET $c\_pre \leftarrow 0.112345678901234$
SET $i\_miu \leftarrow 3.712345678901234$
SET $i\_pre \leftarrow 0.212345678901234$
Generate control sequences using Equation (1)
Method: genSeqArray(seq_no, c_pre, c_miu)
INPUT: chaos parameters seq_no, c_pre, c_miu.
OUTPUT: Array of control sequences(c_seq).
$c\_seq[0] \leftarrow c\_pre$
for $i \leftarrow 1$ to seq_no **do**
  $c\_seq[i] = c\_miu \times c\_seq[i-1] \times (1 - c\_seq[i-1])$
  RETURN c_seq
end for
Generate initial sequences using Equation (1)
inisequences $\leftarrow$ genSeqArray(seq_no, i_pre, i_miu)
Method: generateDynamicParams()
for $i \leftarrow 1$ to no_params **do**
  $control \leftarrow c\_seq[(i+1) \times diff]$
  $c\_count \leftarrow 1$
  while $control < 0.6$ **do**
    $control \leftarrow c\_seq[((i+1) \times diff) + c\_count]$
    $c\_count + 1$
  end while
  $cparams.dynamiccontrol \leftarrow control + 3.0$
  $controlparameterlist \leftarrow dynamiccontrol$
  $dynamicinitial \leftarrow i\_seq[(i+1) \times diff]$
  $cparams.setpreviouselement \leftarrow dynamicinitial$
  $cparamslist \leftarrow cparams$
end for
RETURN cparamslist.
END

quence is called permutation sequence and it can be used for image encryption and decryption. The sample chaotic sequences and the sorted chaotic sequences for the given values are clearly illustrated in Figures 3(a) - (d).

3.3 Generation of Permutation Sequence by Linear Search

The chaotic sequences are real valued sequences, where as the permutation sequences are integer valued sequences. Generating permutation sequence is the vital part of the proposed algorithm. This integer valued permutation sequence is used to shuffle the pixels and color bytes. A chaotic sequence is generated using Equation (1). The first 1000 chaotic elements are discarded to avoid the transient effect. Then, the chaotic sequence $C = \{c_1, c_2, c_3, \ldots, c_n\}$ is chosen from the 1001st chaotic element and sorted in ascending order to get $S = \{s_1, s_2, s_3, \ldots s_n\}$.

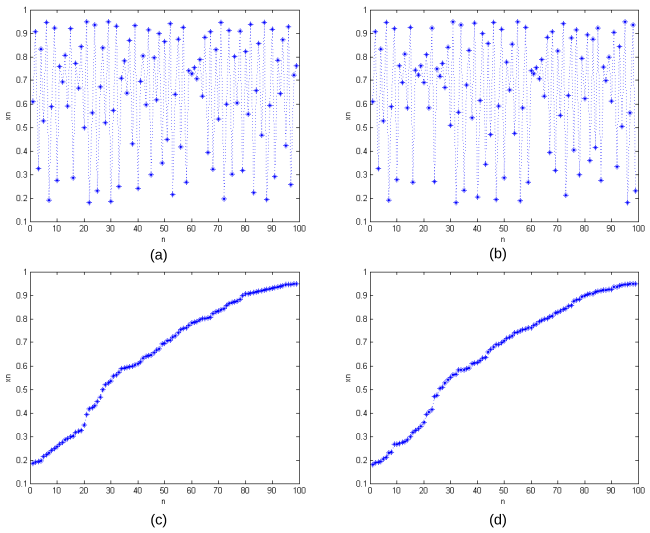


Figure 3: The sample chaotic sequences (a) x_n for $r = 3.80001$, $x_0 = 0.80001$, (b)x_n for $r = 3.80002, x_0 = 0.80002$ and the sorted chaotic sequences (c) x_n for $r = 3.80001, x_0 = 0.80001$, (d) x_n for $r = 3.80002, x_0 = 0.80002$

Further, iterated through S to find the index position of each element in C to get $P = p_1, p_2, p_3, \ldots, p_n$. The indexing an element in an array can be done through linear search. But, the linear search is time consuming. The pseudocode to create permutation sequence using linear search is shown in Algorithm 2 and the sample permutation sequences for the given values are shown in Figures 4(a)-(b).

Algorithm 2 Permutation sequence generation by linear search

BEGIN
$C_{\{c_1,c_2,\ldots,c_n\}} \leftarrow genSeqArray(seq\_no, pre, miu)$
$S_{\{s_1,s_2,\ldots,s_n\}} \leftarrow sortSeqArray(C_{\{c_1,c_2,\ldots,c_n\}})$
$P_{\{p_1,p_2,\ldots,p_n\}} \leftarrow linearSearch(S_{\{s_1,s_2,\ldots,s_n\}}, c(i))$
END

3.4 Generation of Permutation Sequence by Binary Search

The indexing of an array element is done using the binary search which is very faster while comparing to linear search. The S has been iterated to find the index position of each element in C to get $P = p_1, p_2, p_3, \ldots, p_n$ using binary search. The pseudocode for generating permutation sequence using binary search is given in Algorithm 3. The sample permutation sequences for the given values are shown in Figures 4(c)-(d).

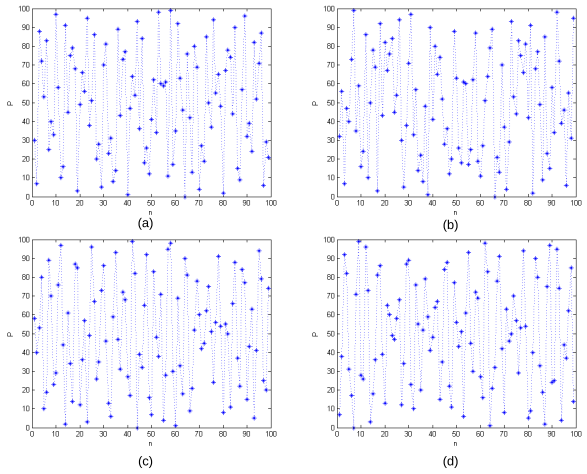


Figure 4: The permutation sequences for the given values (a) P for $r = 3.80001, x_0 = 0.80001$, (b) P for $r = 3.80002, x_0 = 0.80002$ and the permutation sequence using binary search for the given values (c) P for $r = 3.80001, x_0 = 0.80001$, (d) P for $r = 3.80002, x_0 = 0.80002$

Algorithm 3 Permutation sequence generation by binary search

BEGIN
$C_{\{c_1,c_2,...,c_n\}} \leftarrow genSeqArray(seq\_no, pre, miu)$
$S_{\{s_1,s_2,...,s_n\}} \leftarrow sortSeqArray(C_{\{c_1,c_2,...,c_n\}})$
$A_{\{a_1,...,a_n\}} \leftarrow C_{\{c_1,c_n,c_2,c_{n-1},...\}} \leftarrow cSeAr(C_{\{c_1,...,c_n\}})$
$P_{\{p_1,p_2,...,p_n\}} \leftarrow binarySearch(S_{\{s_1,s_2,...,s_n\}}, a(i))$
END

Algorithm 4 Masking sequence generation for pixel shuffling

BEGIN
$C_{\{c_1,c_2,...,c_n\}} \leftarrow genSeqArray(seq\_no, pre, miu)$
$S_{\{s_1,s_2,...,s_n\}} \leftarrow sortSeqArray(C_{\{c_1,c_2,...,c_n\}})$
$n_{max} \leftarrow S_{\{s_1,s_2,...,s_n\}}[S_{\{s_1,s_2,...,s_n\}}.length - 1]$
for $i \leftarrow 0$ to $C_{\{c_1,c_2,...,c_n\}}.length$ **do**
  $curr\_value \leftarrow Math.abs(C[i]/n_{max}) \times (2^{24} - 1)$
  $M[i] \leftarrow curr\_value$
end for
RETURN $M_{\{m_1,m_2,m_3...,m_n\}}$
END

3.5 Comparison of Time Complexity for Permutation Sequence Generation

In this section, the time complexity to generate the permutation sequence by linear search as well as by binary search method with various pixel lengths is investigated. The time taken to generate the permutation sequence are tabulated in the Table 1. The investigation clearly shows that the time taken to generate permutation sequence by binary search method is very little when compared to linear search method.

Algorithm 5 Masking sequence generation for colour byte shuffling

BEGIN
$C_{\{c_1,c_2,...,c_n\}} \leftarrow genSeqArray(seq\_no, pre, miu)$
$S_{\{s_1,s_2,...,s_n\}} \leftarrow sortSeqArray(C_{\{c_1,c_2,...,c_n\}})$
$n_{max} \leftarrow S_{\{s_1,s_2,...,s_n\}}[S_{\{s_1,s_2,...,s_n\}}.length - 1]$
for $i \leftarrow 0$ to $C_{\{c_1,c_2,...,c_n\}}.length$ **do**
  $curr\_value \leftarrow Math.abs(C[i]/n_{max}) \times 255$
  $M[i] \leftarrow curr\_value$
end for
RETURN $M_{\{m_1,m_2,m_3...,m_n\}}$
END

3.6 Proposed Algorithm for Image Encryption

The whole cryptosystem comprises of dynamic key generation, generation of permutation and diffusion sequences, confusion-diffusion operations and XOR operation of source image pixels with key image pixels. All these operation together makes the proposed system more secure. The dynamic control and initial parameters are generated which make the algorithm more secured. Thereby, the key become more secure and difficult to guess. The permutation sequences are generated using dynamically controlled logistic map. During the permutation sequence generation, the sequence element is indexed using binary search which makes the algorithm to behave faster. The colour byte scrambling and pixel scrambling are done iteratively with respect to different permutation sequences. Similarly, the masking sequences are generated to alters the values of pixels as well as colour bytes. The confusion and diffusion operations are done iteratively to makes the algorithm more complex.

The test image pixels are read and divided into small chunks of length n and the chunks are stored into a container map CM with index. The total number of pixels of test medial image are 9000000. But, the permutation and diffusion sequences are generated for 100000. It is not necessary to generate 9000000 sequences which takes more time. Further, the chunk size is 100000. So, the time taken for sequence generation is very less. The permutation sequences P_1, P_3 generated where length $length = n$ and P_2 & P_4 where $length = n \times 3$ to scramble image pixels' and colour bytes' positions respectively. Similarly the masking sequences M_1 & M_3 are generated to alter the 24-bit pixel values and colour bytes where $m_i = Int\left[(|(m_i/max\,(m_i))|) \times (2^{24} - 1)\right]$ and M_2 & M_4 where $m_i = Int\left[(|(m_i/max\,(m_i))|) \times 255\right]$. The pseudocode for generating masking sequence is given in Algorithms 4 and 5.

Then, the one dimensional array of pixels P is retrieved from container map. The pixels P is shuffled using permutation sequence P_1 and a bitwise XOR operation is done between P, M_1 and K $(P \oplus M_1 \oplus K)$. The K is the pixels of key image. The size of the key image 1000×1000 which consists random pixel and byte values. Then, the

Table 1: Time taken to generate the permutation sequence for various pixel lengths

$Width \times Height$	No. of pixels	Linear Search (sec)	Binary Search (sec)
256×256	65536	7.4	0.132
512×512	262144	117.29	0.596
1024×1024	1048576	1894.009	3.246
1920×1080	2073600	1921.561	5.815
3000×3000	9000000	38542.067	38.205

color bytes C, Y is separated from pixels P, K respectively. Now C is shuffled using permutation sequence P_2.

Then, the bitwise XOR operation is done between C, M_2 and Y. Similarly, the sequences P_3, P_4, M_3 and M_4 are employed for pixel confusion, colour byte confusion, pixel diffusion and colour byte diffusion respectively. The bitwise XOR operation between shuffled image with key image enhance the randomness of the cipher image. The alogorithm becomes efficient due to the number of rounds of scrambling done. The decryption is the reverse process of encryption. The pseudocode for the proposed algorithm is given in Algorithm 6.

The block diagram of the proposed scheme is given in Figure 5.

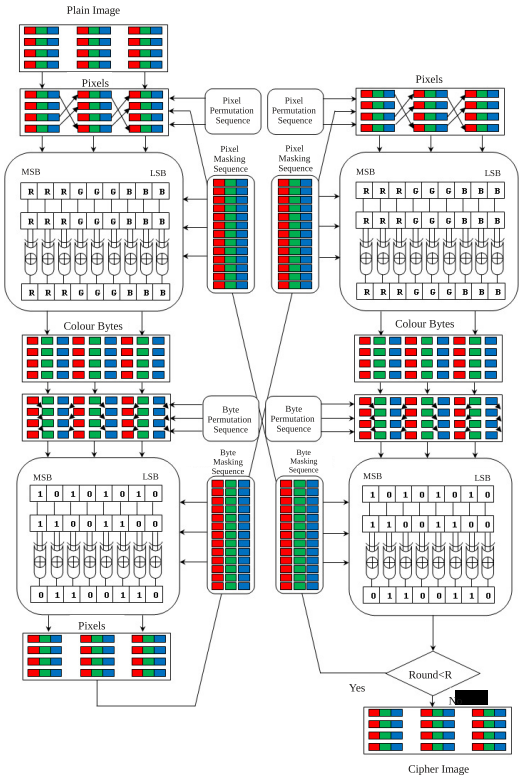


Figure 5: The block diagram of proposed algorithm

4 Result and Discussion

The proposed algorithm is experimented using a system with the Intel Core i5-3230M processor with a speed of 2.60 GHz, and the RAM capacity is 4GB. The algorithm has been coded using Java JDK 1.7.0 (64 bit). To analyze the proposed algorithm, seven large size and 24-bit colour images with size 3000 3000 are considered and they are downloaded from The Human Protein Atlas [23, 27].

The experimental results are given to illustrate the efficiency of the proposed cryptosystem with high resolution images. The selected initial and control parameters are $r_1 = 3.612345678901234, x_1 = 0.112345678901234, r_2 = 3.712345678901234, x_2 = 0.212345678901234$. According to the proposed algorithm, the values of $l = 3000 \times 3000$, $n = 100000$, and $R = 4$. The experimental plain, cipher and key images are given in Figure 6.

4.1 Security Analysis

The systematic security analyses have been done on proposed encryption scheme to verify that the system is much protective against the most frequent attacks. The key space and sensitivity analysis, statistical analysis, histogram analysis, mean variance color byte, information entropy analysis have been done to prove that the system withstands on different attacks.

4.1.1 Key Space Analysis

The sensitivity of the proposed cryptosystem is around 10^{-16} based on the control parameter and the sensitivity is around 10^{-15} based on the initial condition. The cryptosystem is designed in such a way that the encryption of colour components rely on each other. The keys of colour component are independent from each other. The key space is calculated by the control and the initial parameters, $K = (10^{16} \times 10^{16} \times 10^{16} \times 10^{16} \times 10^{16} \times 10^{16} \times 10^{16} \times 10^{16}) \times (10^{15} \times 10^{15} \times 10^{15} \times 10^{15} \times 10^{15} \times 10^{15} \times 10^{15} \times 10^{15}) = 10^{248}$. The key space is much higher while comparing with other literatures [3, 35, 26] as given in Table 2. This proves that the system defy against different brute-force attacks.

4.1.2 Differential Attack Analysis

The original image can be modified slightly (e.g., change a single pixel value) and encrypted to study the sensitiv-

Algorithm 6 Proposed algorithm for image encryption

BEGIN
READ image
SET $W \leftarrow$ image width
SET $H \leftarrow$ image height
CALCULATE $l \leftarrow W \times H$
SET I as one dimensional array of image pixels
READ key image
SET K as one dimensional array of key image pixels
GENERATE indexed Container Map(CM)contains small chunks of length n from pixels I of length l
GENERATE Permutation sequences P_1 and P_3 where $length = n$ and P_2 and P_4 where $length = n \times 3$
GENERATE Masking sequences $M_1\&M_3$ where masking sequence array $(m_i) \leftarrow [max(m_i) \times (2^{24} - 1)]$ and $M_2\&M_4$ where masking sequence array $m_i = [max(m_i) \times 255]\{m_i$-sequence Array[i], $max(m_i)$−maximum Sorted Sequence Array Value$\}$
READ one dimensional array P from Container map
for $i \leftarrow 0$ to 4 **do**
  for $i \leftarrow 0$ to length of the permutation sequence array index P_1 **do**
    Encrypted pixels[i] ← Pixel array[sequence Array Index [i]]
  end for
  COMPUTE Encrypted pixels P $\leftarrow P \oplus M_1 \oplus K$
  GENERATE colour bytes $C\&Y$ from pixels $P\&K$
  for $i \leftarrow 0$ to length of the permutation sequence array index P_2 **do**
    Encrypted bytes[i] $(C) \leftarrow$ byte array[sequence Array Index [i]]
  end for
  COMPUTE Encrypted Colour bytes $\leftarrow C \oplus M_2 \oplus Y$
  GENERATE pixels $P\&K$ from the color bytes $C\&Y$
  for $i \leftarrow 0$ to length of the permutation sequence array index P_3 **do**
    Encrypted pixels[i] ← Pixel array[sequence Array Index [i]]
  end for
  COMPUTE Encrypted pixels$\leftarrow P \oplus M_3 \oplus K$
  GENERATE colour bytes $C\&Y$ from pixels $P\&K$
  for $i \leftarrow 0$ to length of the permutation sequence array index P_4 **do**
    Encrypted bytes[i] $(C) \leftarrow$ byte array[sequence Array Index [i]]
  end for
  COMPUTE Encrypted Colour bytes $(C) \leftarrow C \oplus M_4 \oplus Y$
  GENERATE pixels $(P) \leftarrow$ colour bytes (C)
end for
COMPUTE resultant pixels I (Cipher image) ← chunks of pixels assembled in container map (CM) of length l.
END

Table 2: The key space value comparisons

Encryption Algorithm	Key Space
Proposed cryptosystem	$2^{720} \approx 10^{248}$
Ref. [26]	$2^{298} \approx 10^{90}$
Ref. [3]	2^{292}
Ref. [35]	2^{400}

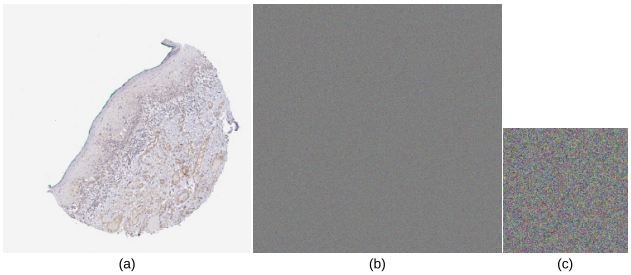


Figure 6: (a) The original plain image (3000 × 3000), (b) cipher image (3000×3000) and (c) key image (1000×1000)

ity of the encryption algorithm. If a slight change in the original image makes a significant change in the cipher image, then the algorithm is efficient and can withstands on any differential attacks. The differential analysis has been carried out by calculating the NPCR (Net Pixel Change Rate) and UACI (Unified Average Changing Intensity) using Equations (2) and (3).

$$NPCR = \frac{1}{W \times H} \sum_{p=1}^{W} \sum_{q=1}^{H} D(p,q) \times 100\% \quad (2)$$

$$UACI = \frac{1}{W \times H} \sum_{p=1}^{W} \sum_{q=1}^{H} \frac{|E_1(p,q) - E_2(p,q)|}{255} \times 100\% \quad (3)$$

where $D(p,\ q)$ represent the difference between $E_1(p,q)$ and $E_2(p,q)$. If $E_1(p,q) = E_2(p,q)$ then $D(p,\ q) = 0$, else $D(p,\ q) = 1$.

For an 8-bit grey image, the expected approximate are $NPCR_E = 99.6094\%$ and $UACI_E = 33.4635\%$. The calculated values of proposed algorithm are given in Table 3. The calculated NPCR and UACI values of immunohistochemical images are higher than the given in [4, 19]. Table 3 evidences that the proposed cryptosystem withstands strongly against differential attacks.

4.1.3 Correlation Coefficient Analysis

The correlation in between the two neighbouring pixels in the original and cipher image is tested by considering the whole pairs of neighbouring pixels in different directions (in horizontal, vertical and diagonal) from the original and cipher images and the correlation coefficients are es-

Table 3: The NPCR and UACI values of the experimental images

File name	NPCR%			UACI%		
	Red	Green	Blue	Red	Green	Blue
109225_A_8_1	99.6022	99.6119	99.6095	41.8200	41.3991	41.8393
109225_A_9_1	99.6066	99.6119	99.6077	42.0469	41.9835	43.2646
7436_A_7_1	99.6090	99.6117	99.6115	36.2793	36.3419	38.2842
7436_A_8_1	99.6070	99.6130	99.6159	39.6193	39.6151	40.9691
7436_A_9_1	99.6050	99.6085	99.6120	36.7789	36.7797	38.5883
54482_A_8_1	99.6096	99.6014	99.6094	41.2274	41.2965	42.0912
54482_A_9_1	99.6098	99.6035	99.6125	41.8828	41.9248	42.5739
Lena	99.6219	99.6044	99.6067	33.0119	30.6249	27.63317
Ref. [12] Image 1	99.6253	99.6332	99.6259	33.4312	33.3385	33.3522
Ref. [19] Lena	98.8373	99.6277	99.6445	33.0753	30.7349	28.0029
Ref. [4] Lena	99.6013	99.6131	99.6226	33.4210	33.4485	33.4815

timated using the equations given below.

$$
\begin{aligned}
E(x) &= \frac{1}{W \times H} \sum_{p=1}^{W} \sum_{q=1}^{H} x(p,q) \\
D(x) &= \frac{1}{W \times H} \sum_{p=1}^{W} \sum_{q=1}^{H} [x(p,q) - E(x)]^2 , \\
Conv(x,y) &= \frac{1}{W \times H} \sum_{p=1}^{W} \sum_{q=1}^{H} [x\,(p,q) - E(x)] \\
&\qquad [y\,(p,q) - E(y)] \\
\gamma_{xy} &= \frac{Conv(x,y)}{\sqrt{D(x)}\sqrt{D(y)}} \qquad (4)
\end{aligned}
$$

where x and y are the Red, Green, Blue values of two-adjacent pixels in the image and γ_{xy} is the correlation coefficient of two adjacent pixels. The correlation coefficients in horizontal (HC), vertical (VC) and diagonal (DC) directions of cipher images are shown in Table 4. From Table 4, the cipher images' correlation analysis values are nearing '0' and they are farther apart from the correlation analysis values of plain images. The correlation coefficient value of plain images are nearing 1. For the image 109225_A_8_1, the horizontal correlation coefficient values of plain and cipher images are 0.9 and -0.00034 respectively which proves the superiority of confusion and diffusion characteristics. Further, the values are identical with other literatures [12, 19, 34].

The values of the two neighbouring pixels for different directions are plotted as shown in the following Figures 7(a)-(f).

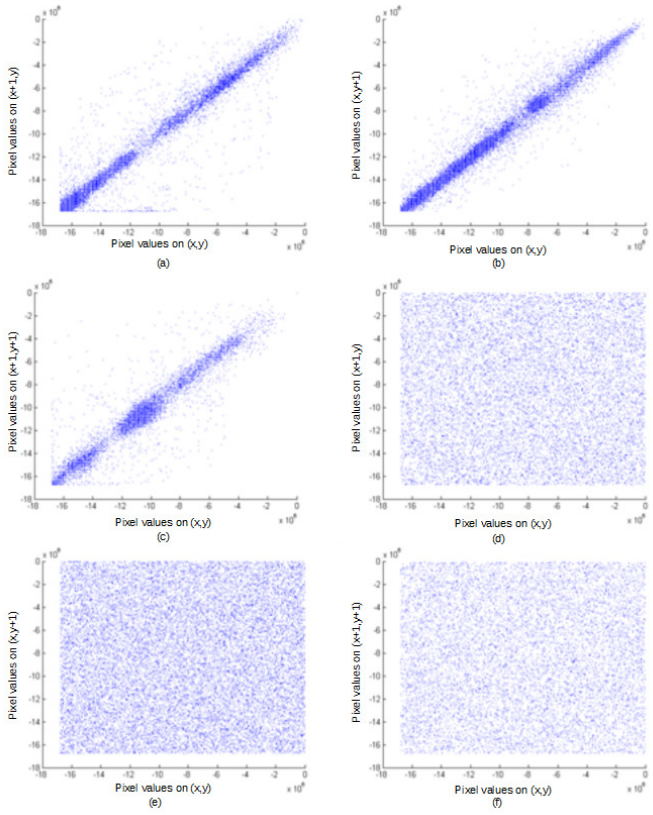


Figure 7: (a), (b), and (c) represents the horizontal, vertical, and diagonal correlation coefficients of original plain image; and (d), (e), and (f) represents horizontal, vertical, and diagonal correlation coefficients of cipher image.

4.1.4 Key Sensitivity Analysis

The control and the initial parameters used for medical image encryption in the experimental part are recalled and attempted to decrypt the ciphered images with different key. The different key is prearranged with the value of $r_1 = 3.612345678901235$ which have the tiny difference from the encryption key and then the consequent decrypted images are shown in Figures 8(a)-(b). It is evident that the use of slightly different key for decryption results in the entirely different decrypted image from the original which shows that the proposed system is highly key sensitive.

4.1.5 Histogram Analysis

A histogram is a graph which is used to show the pixel value distribution of an image. An opponent can guess

Table 4: The correlation coefficients (CC) in horizontal (HC), vertical (VC) and diagonal (DC) directions of cipher images

Component	CC	109225_A_8_1	109225_A_9_1	7436_A_7_1	7436_A_8_1	7436_A_9_1	54482_A_8_1	54482_A_9_1	Lena	Ref.[19]	Ref.[34]
Red	HC	-0.00034	-0.00027	0.00125	0.00238	0.00143	0.00280	0.00296	0.00298	-0.00110	-0.00201
	VC	-0.00043	-0.00028	-0.00270	-0.00157	-0.00290	-0.00160	-0.00108	0.00242	-0.01110	0.00293
	DC	0.00052	0.00015	-0.00012	-0.00080	0.00001	-0.00109	-0.00074	0.00312	0.00120	-0.00164
Green	HC	-0.00005	-0.00002	-0.00213	-0.00153	-0.00202	-0.00114	-0.00102	0.00105	-0.00008	-0.00149
	VC	-0.00095	-0.00088	-0.00447	-0.00491	-0.00507	-0.00451	-0.00462	0.00377	0.00130	-0.00221
	DC	0.00003	0.00003	0.00215	0.00192	0.00242	0.00141	0.00140	0.00148	0.00370	0.00349
Blue	HC	0.00012	0.00011	-0.00061	-0.00165	-0.00071	-0.00200	-0.00213	0.00247	0.00120	0.00055
	VC	-0.00147	-0.00151	0.00248	0.00235	0.00271	0.00200	0.00189	0.00103	0.00540	0.00037
	DC	0.00007	-0.00015	0.00332	0.00512	0.00294	0.00565	0.00602	0.00019	0.00110	-0.00092

Figure 8: (a)Decrypted cipher 109225_A_8_1 with wrong key (b) Decrypted cipher 109225_A_9_1 with wrong key

certain amount of data if the histogram is not flat enough. This is known as cipher only attack by analyzing the statistical property of the ciphered image. So the flat distribution is obtained during the medical image encryption. The distribution diagram of the color bytes Red, Green and Blue of the plain image and the cipher images are shown in the Figures 9(a)-(f). From the figures, it is proved that the cipher images are capable to resist the cipher only attack.

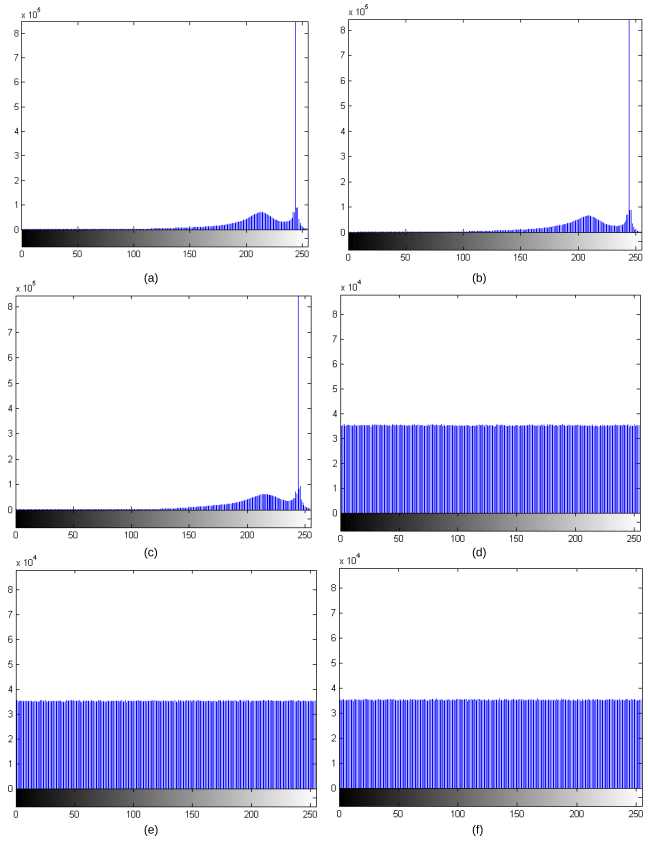


Figure 9: (a), (b), and (c) represents the R, G and B distribution of original plain image; and (d), (e), and (f) represents the R, G and B distribution of cipher image.

4.1.6 Mean-variance Color Byte Analysis

The definition of mean-variance color byte value is,

$$C = \frac{1}{W \times H} \sum_{p=1}^{W} \sum_{q=1}^{H} \left| B\left(p,q\right) - \overline{B} \right| \tag{5}$$

where $\bar{B}$ denotes the average of all the color values of image pixels, and $W \times H$ is the size of the plain image. The mean-variance colour byte value of cipher image is higher than the plain image and it proves that the scramble performance of the proposed algorithm is better. The red component's mean-variance color byte values of the plain and encrypted image are 20.16568 and 64.03818 respectively as shown in Table 5.

4.1.7 Information Entropy Analysis

To measure the randomness of encryption, information entropy is the metric which is given by Shannon is used. It is a mathematical theory which is used for data communication and storage. The entropy value of the source one is smaller than the ideal one. Let m be the information source, and the formula for calculating information

Table 5: Mean-variance color byte value of images

File name	Size	Plain Image			Cipher Image		
		Red	Green	Blue	Red	Green	Blue
109225_A_8_1	3000 × 3000	20.16568	22.89676	19.77083	64.03818	64.00136	63.98380
109225_A_9_1	3000 × 3000	18.75448	19.47938	12.89327	64.02837	64.00637	63.98659
7436_A_7_1	3000 × 3000	38.53359	37.90420	26.13839	63.97741	64.03066	63.93428
7436_A_8_1	3000 × 3000	26.88287	26.93553	18.96092	63.99242	64.06941	63.89905
7436_A_9_1	3000 × 3000	34.76804	34.54240	24.29289	63.97604	64.02882	63.94772
54482_A_8_1	3000 × 3000	19.59453	19.12174	15.27504	64.01059	64.07322	63.88637
54482_A_9_1	3000 × 3000	18.52384	18.19927	15.11005	64.00498	64.06246	63.87084
Lena	512 × 512	41.44229	43.78703	27.63417	64.00606	63.97906	64.15340

entropy is,

$$H(m) = \sum_{i=1}^{M} p(m_i) log \frac{1}{p(m_i)} \quad (6)$$

where M is the total number of symbols $m_i \in m; p(m_i)$ represents the probability of occurrence of symbol m_i and log denotes the base 2 logarithm so that entropy is expressed. For a random source emitting 256 symbols, image's entropy is $H(m) = 8$ bits. The information entropy values of the plain images and the cipher images are shown in Table 6.

4.2 Encryption Speed Analysis

The encryption speed is an important factor for describing the performance of the algorithm. Though the colour image size is 3000×3000, the algorithm executes the encryption efficiently by means of splitting the image into small chunks. Only one lakh permutation and diffusion sequences are generated to confuse and diffuse the chunks of image iteratively. Hence, the time taken for sequence generation is very less while comparing to generate sequences for whole image. Thereby, the algorithm encrypts faster. Further, the binary search algorithm speeds up the system and it is employed in indexing the sequences during permutation sequence generation. The encryption time for test medical images are calculated and compared with existing literatures (Table 7). The encryption speed is 12.68 Mbit/s which is much better than the existing methods, thereby it proves that the cryptosystem is very fast and efficient.

Table 7: The encryption speed comparisons

Encryption Algorithm	Encrption speed (Mbit/s)
Proposed algorithm	12.68
Ref. [13]	9.89
Ref. [22]	9.39
Ref. [25]	9.12

5 Discussion

The chaotic image encryption literature provides many encryption techniques proposed by targeting grey scale and colour images with smaller sizes [8, 16, 28, 29, 31]. Since the grey scale images are compiled with single channel pixels of 8-bit depth, the image encryption schemes are dealt with pixel shuffling and masking only. In case of 24-bit color images, the encryption is dealt with pixels and color bytes. So the proposed scheme is employing confusion and diffusion operations on both pixels and color bytes of an image. Also by using the key image, it is possible to add more randomness by executing XOR operation with the scrambled pixels and bytes. The resultant cipher image is having a good quality of confusion and diffusion properties. Further, the encryption standard is not only based on the logistic map, it also rigorously depends on the algorithm designed in the manner that involved in step by step operations on pixels and number of iterations in shuffling the pixels and colour bytes. The key space of the proposed system is sufficient and better than the other approaches [3, 26, 35]. The correlation coefficient, NPCR and UACI values of plain and cipher images are optimal and identical with existing approaches [4, 12, 19, 34]. The randomness of the cipher image is proved by better entropy values than the existing methods [12, 15, 25, 35]. The encryption speed is faster than the reported approach [13]. From the numerical results, it is evident that the proposed scheme possesses the highest security against various forms of threats. Hence, it is faster and efficient when compared to other existing methods.

6 Conclusion

This paper presents a faster and efficient symmetric cryptosystem using logistic map to encrypt large size 24-bit colour medical images. The proposed cryptosystem will be able to process different kinds of medical images as well as images of any size. Security analyses and experimental results demonstrated the effectiveness of the given scheme. The large key space supports the system to resist against different brute-force attacks. Statistical analysis reveals that the scheme protects the image from

Table 6: Information entropy values of plain and cipher images

File name	Plain Image			Cipher Image		
	Red	Green	Blue	Red	Green	Blue
109225_A_8_1	3.66064	3.71358	3.66988	7.99993	7.99994	7.99994
109225_A_9_1	4.30507	4.32594	4.15458	7.99995	7.99996	7.99996
7436_A_7_1	5.94937	5.90641	5.69366	7.99989	7.99989	7.99991
7436_A_8_1	4.98705	4.96496	4.79376	7.99978	7.99979	7.99983
7436_A_9_1	5.93084	5.89377	5.63771	7.99988	7.99992	7.99991
54482_A_8_1	3.97933	3.95653	3.86004	7.99948	7.99947	7.99955
54482_A_9_1	3.88024	3.86895	3.79533	7.99944	7.99945	7.99953
Lena	7.25310	7.59403	6.96842	7.99936	7.99926	7.99932
Ref. [12] Image 1	7.4567	7.2863	6.9628	7.9893	7.9879	7.9897
Ref. [35] Lena	-	-	-	7.99936	7.99935	7.99936
Ref. [15] Lena	-	-	-	7.98970	7.98770	7.98960
Ref. [25] Lena	-	-	-	7.99927	7.99924	7.99911

any form of statistical attack. The scheme has high sensitivity to plain image and key, so it can withstand on differential attack. Based on the performance and results of security analsyses, one can conclude that the algorithm is much faster and efficient. On the whole, the proposed encryption scheme has high-level of security and it can be utilized in secure medical image storage and communications.

Acknowledgements

The authors would like to thank the authorities of Bharathiar University, Coimbatore, India for providing the necessary laboratory facilities to carry out this study.

References

[1] G. Alvarez and S. Li, "Some basic cryptographic requirements for chaos-based cryptosystems," *International Journal of Bifurcation and Chaos*, vol. 16, no. 8, pp. 2129–2151, 2006.

[2] G. Bellandi, M. Giannini, and C. Grande, "Mobile ehealth technology and healthcare quality impacts in italy," *International Journal of Healthcare Management*, vol. 6, no. 3, pp. 192–200, 2013.

[3] A. Diaconu and K. Loukhaoukha, "An improved secure image encryption algorithm based on rubiks cube principle and digital chaotic cipher," *Mathematical Problems in Engineering*, pp. 1–10, 2013.

[4] C. Dong, "Color image encryption using one-time keys and coupled chaotic systems," *Signal Processing Image Communication*, vol. 29, no. 5, pp. 628–640, 2014.

[5] J. Galvez, "Whole slide imaging and telepathology," *Breast Cancer Research*, vol. 5, pp. 1–2, 2003.

[6] T. G. Gao and Z. Q. Chen, "Image encryption based on a new total shuffling algorithm," *Chaos Solitons Fractals*, vol. 38, no. 1, pp. 213–220, 2008.

[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] S. K. A. Hafiz, A. G. Radwan, S. H. Abdel Haleem, and M. L. Barakat, "A fractal-based image encryption system," *IET Image Processing*, vol. 8, no. 12, pp. 742–752, 2014.

[9] L. C. Huang, L. Y. Tseng, and M. S. Hwang, "The study on data hiding in medical images," *International Journal of Network Security*, vol. 14, no. 6, pp. 301–309, 2012.

[10] L. C. Huang, L. Y. Tseng, and M. S. Hwang, "A reversible data hiding method by histogram shifting in high quality medical images," *Journal of Systems and Software*, vol. 86, no. 3, pp. 716–727, 2013.

[11] I. Hussain and T. Shah, "Application of s-box and chaotic map for image encryption," *Mathematical and Computer Modelling*, vol. 57, no. 9-10, pp. 2576–2579, 2013.

[12] G. Liu, J. Li, and H. Liu, "Chaos-based color pathological image encryption scheme using one-time keys," *Computers in Biology and Medicine*, vol. 45, pp. 111–117, 2014.

[13] H. Liu and C. Jin, "A color image encryption scheme based on arnold scrambling and quantum chaotic," *International Journal of Network Security*, vol. 19, no. 3, pp. 347–357, 2017.

[14] L. Liu and Z. Cao, "Analysis of two condentiality-preserving image search schemes based on additive homomorphic encryption," *International Journal of Electronics and Information Engineering*, vol. 5, no. 1, pp. 1–5, 2016.

[15] L. Liu, Q. Zhang, and X. Wei, "A rgb image encryption algorithm based on dna encoding and chaos map," *Computers and Electrical Engineering*, vol. 38, pp. 1240–1248, 2012.

[16] A. Masmoudi and W. Puech, "Lossless chaos-based crypto-compression scheme for image protection," *IET Image Processing*, vol. 8, no. 12, pp. 671–686, 2014.

[17] R. Matthes, "On the derivation of a chaotic encryption algorithm," *Cryptologia*, vol. 13, no. 1, pp. 29–42, 1989.

[18] A. Mersaid, T. Gulom, "The encryption algorithm AES-RFWKIDEA32-1 based on network RFWKIDEA32-1," *International Journal of Electronics and Information Engineering*, vol. 4, no. 1, pp. 1–11, 2016.

[19] P. Padmapriya, R. Amirtharajan, K. Thenmozhi, and J. B. B. Rayappan, "Triple chaotic image scrambling on rgb a random image encryption approach," *Security and Communication Networks*, vol. 8, no. 18, pp. 3335–3345, 2015.

[20] M. Y. M. Parvees, J. A. Samath, and B. P. Bose, "Secured medical images - a chaotic pixel scrambling approach," *Journal of Medical Systems*, vol. 40, no. 11, pp. 232:1–232:11, 2016.

[21] M. Y. M. Parvees, J. A. Samath, I. K. Raj, and B. P. Bose, "A colour byte scrambling technique for efficient image encryption based on combined chaotic map," in *Proceedings of the International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT'16)*, pp. 1067–1072, Mar. 2016.

[22] V. Patidar, N. K. Pareek, and K. K. Sud, "A new substitution-diffusion based image cipher using chaotic standard and logistic maps," *Communications in Nonlinear Science and Numerical Simulation*, vol. 14, no. 7, pp. 3056–3075, 2009.

[23] Protein Atlas, *The Human Protein Atlas Project*, Apr. 15, 2017. (`http://www.proteinatlas.org`)

[24] E. Ryhan, "Telemedicine: Current and future perspectives," *International Journal of Computer Science Issues*, vol. 10, no. 6, pp. 242–249, 2013.

[25] S. M. Seyedzadeh and S. Mirzakuchaki, "A fast color image encryption algorithm based on coupled two-dimensional piecewise chaotic map," *Signal Processing*, vol. 92, pp. 1202–1215, 2012.

[26] B. Stoyanov and K. Kordov, "Image encryption using chebyshev map and rotation equation," *Entropy*, vol. 17, pp. 2117–2139, 2015.

[27] M. Uhlen, P. Oksvold, L. Fagerberg, E. Lundberg, K. Jonasson, M. Forsberg, M. Zwahlen, C. Kampf, K. Wester, S. Hober, H. Wernerus, L. Bjorling, and F. Ponten, "Towards a knowledge-based human protein atlas," *Nature Biotechnology*, vol. 28, no. 12, pp. 1248–1250, 2010.

[28] X. Wang and K. Guo, "A new image alternate encryption algorithm based on chaotic map," *Nonlinear Dynamics*, vol. 76, pp. 1943–1950, 2014.

[29] X. Y. Wang, "New chaotic encryption algorithm based on chaotic sequence and plain text," *IET Information Security*, vol. 8, no. 3, pp. 213–216, 2014.

[30] X. Y. Wang and J. F. Zhao, "A new image encryption algorithm based on chaos," *Optics Communications*, vol. 285, no. 5, pp. 562–566, 2012.

[31] Y. Wang, K. W. Wong, X. F. Liao, and G. R. Chen, "A new chaos-based fast image encryption algorithm," *Applied Soft Computing*, vol. 11, no. 1, pp. 514–522, 2011.

[32] J. D. Webster and R. W. Dunstan, "Whole-slide imaging and automated image analysis: considerations and opportunities in the practice of pathology," *Veterinary Pathology*, vol. 51, no. 1, pp. 211–223, 2014.

[33] M. Yang, M. Trifas, and L. Chen, "Secure patient information and privacy in medical imaging, systemics," *Cybernetics and Informatics*, vol. 8, no. 3, pp. 63–66, 2010.

[34] W. Zhang, K. W. Wong, H. Yu, and Z. L. Zhu, "A symmetric color image encryption algorithm using the intrinsic features of bit distributions," *Communications in Nonlinear Science and Numerical Simulation*, vol. 18, no. 3, pp. 584–600, 2013.

[35] Y. Q. Zhang and X. Y. Wang, "A symmetric image encryption algorithm based on mixed linear-nonlinear coupled map lattice," *Information Sciences*, vol. 273, pp. 329–351, 2014.

Biography

M. Y. Mohamed Parvees received his M.Sc. (Information Technology) in 2002 from Gandhigram Rural Institute - Deemed University and completed M.Phil. (Computer Science) in 2004 from Annamalai University, India. Presently, he is a faculty in Department of Computer and Information Science, Annamalai University. He pursues his Ph.D. degree in Bharathiar University. He has few international and national publications. His research interests include cryptography, multimedia security and medical information systems.

Dr. J. Abdul Samath received his Ph.D. (Computer Science) from Gandhigram Rural Institute - Deemed University. Currently, he is working as an Assistant Professor in Government Arts College, Udumalpet. He has 10 years of teaching experience and he has published 12 research articles in international journals. His research interests include neural networks, image processing, control theory, cryptography and medical image analysis.

B. Parameswaran Bose received his M.Sc. (Information Technology) in 2002 from Gandhigram Rural Institute - Deemed University. He has 7 years of experience in software research and development mainly in the field of application programming, information security and web technologies with knowledge in analyzing, developing and deploying critical applications. Presently, he does research on cryptography and information security.

Confidentiality-Preserving Personal Health Records in Tele-Healthcare System Using Authenticated Certificateless Encryption

Rui Guo[1,2], Huixian Shi[3]

(Corresponding author: Rui Guo)

National Engineering Laboratory for Wireless Security, Xi'an University of Posts and Telecommunications[1]
Xi'an 710121, P.R. China
(Email: guorui@xupt.edu.cn)
State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications[2]
School of Mathematics and Information Science, Shaanxi Normal University[3]

(Received Aug. 3, 2016; revised and accepted Nov. 15 & Dec. 25, 2016)

Abstract

Wireless Medical Sensor Networks (WMSN) facilitate the traditional healthcare systems, however, due to the public transmission, the healthcare system in WMSN also faces some serious security and privacy challenges. These are major concerns in the Health Insurance Portability and Accountability Act. Especially, integrity and confidentiality of patient physiological data are two key issues in privacy protection, which must be considered and addressed firstly. Therefore, the security and privacy in such systems should be enforced via authentication as well as encryption. This paper presents an authenticated certificateless public key encryption scheme for protecting the integrity and confidentiality of the patient sensitive information in tele-healthcare system simultaneously. The security of this protocol is based on the hardness of the bilinear Diffie-Hellman problem, and we prove that it is secure in the random oracle model. Our analysis and comparisons with related protocols show that this scheme is a viable encryption for tele-healthcare system.

Keywords: Authentication; Bilinear Pairing; Certificateless Public Key Encryption; Privacy; Tele-healthcare System

1 Introduction

Wireless Medical Sensor Networks (WMSN) are the networks of medical sensors with small, limited memory and low battery power, for enabling to offer professional, individualized and real-time medical services [9]. In WMSN, the wearable sensor in the patient's body transmits his/her physiological signals (e.g., blood pressure, pulse oxi-meter and temperature, etc.) to the doctor via a wireless channel. In the transmission, lacking of necessary security protection may divulge the patient's privacy, and then cause that the adversary eavesdrops and distorts actual data to misadvise the patients with these false diagnoses and treatments [22].

The Health Insurance Portability and Accountability Act (HIPAA) [3], as a guideline for privacy and security regulations, was presented in 1996. This Act stated that the integrity and confidentiality of the personal health records (PHR) between patient and doctor should be ensured. Therefore, in order to protect the patient's privacy, authentication mechanisms and encryption protocols among patient, the medical server (MS) and doctor are essential in tele-healthcare systems (THS).

1.1 Related Works

Wu et al. [26] proposed an authentication protocol with a new phase named the pre-computing phase for the telecare medicine information system (TMIS). In that phase, the entity computes costly and time consuming exponential operations and stores them into a smart card. When these values are needed, the entity enables to ex-tract them from the device rapidly to raise performance. In 2012, He et al. [7] pointed out that Wu et al.'s scheme suffered from the impersonation attack to the insider attack. In order to overcome this weakness, they also proposed a more secure authentication scheme for TMIS. Following these two works, the different authentication protocols [14, 23, 28] were presented to ensure that the data cannot be distorted by an illegal entity.

With regard to the confidentiality of data, in public key cryptography, a public key infrastructure (PKI) is responsible for providing an assurance through the certificates issued by a certification authority (CA). However, this PKI must manage the certificate in revocation, storage, distribution and verification, which places a huge cost on

the entity [10]. To avoid these disadvantages, Shamir [19] put forward the notion of identity based public key cryptography (ID-PKC) by deriving the user's public key directly from its identity information, such as email address and IP address. Moreover, Boneh and Franklin [4] presented a practical identity based encryption (IBE) firstly. Nevertheless, the inherent key escrow problem in ID-PKC is a great drawback [15]. Al-Riyami and Paterson [1] introduced a new paradigm called certificateless public key cryptography (CL-PKC) to get rid of the above flaws. Then, they proved that their certificateless encryption (CLE) is secure in the random oracle model. Based on that scheme, Guo et al. [6] proposed a provably secure CLE scheme for TMIS, which protects the confidentiality of the PHR efficiently.

A common characteristic of above schemes is that each protocol satisfies only one requirement of HIPAA. In 2002, Lynn [12] proposed an authenticated IBE firstly, which integrated the authentication with encryption on the basis of Boneh-Franklin's IBE system [4] and ensured the integrity and confidentiality of the data simultaneously. Unfortunately, there is a security defect that the private key generator (PKG) has the ability of impersonating any user to recover confidential messages. After that, Cheng and Comley [5] constructed an authenticated CLE to prevent the malicious PKG from eavesdropping the privacy information. In addition, as a special authenticated encryption, the signcryption also achieves the same purpose [27]. Barbosa and Farshim [2] proposed the first certificateless signcryption scheme in 2008. However, their construction is vulnerable to the malicious-but-passive key generation center (KGC) at-tacks. In the same year, Wu and Chen [25] designed a more efficient certificateless signcryption scheme and introduced the public verifiability into it. Shamila et al. [17] claimed that the scheme in [25] could not provide the confidentiality of data. Liu et al. [11] introduced an efficient certificateless signcryption scheme with the security proof in the standard model. But Sharmila et al. and Weng et al. [18, 24] pointed out that their security proof is not sound and the scheme is in fact insecure. In 2015, Huang et al. [8] proposed a new efficient convertible multi-authenticated encryption scheme for mobile communication which the signature was cooperatively produced by a group of signers instead of a signal signer. Based on factoring and discrete logarithms,Tsai et al. [21] recently designed a publicly verifiable authenticated encryption scheme. They also claimed that even if either factoring or discrete logarithms is broken, their scheme still could keep the authentication, integration and confidentiality of the message.

1.2 Our Contributions

In this paper, we put forward an authenticated CLE (Auth-CLE) scheme in THS for protecting PHR. The authentication phase is added in decryption to protect the integrity and confidentiality of ciphertext at the same time. Furthermore, we prove that our scheme is secure in the random oracle model, provided that the bilinear Diffie-Hellman (BDH) problem is intractable. At last, we compare the cost of the computation and communication between our proposal and others by the evaluations and experiments and it concludes that our protocol offers better performances in efficiency.

The remainder of this paper is organized as follows. Section 2 addresses some preliminaries such as bilinear pairing, complexity assumption and the model of Auth-CLE. Section 3 proposes an Auth-CLE scheme and proves its security in the random oracle model. Section 4 compares the proposed scheme with some other related schemes from two points. Finally, we conclude the paper in Section 5.

2 Preliminaries

2.1 Bilinear Pairing

Let G_1 be a cyclic additive group generated by a point P, whose order is p, G_2 be a multiplicative group of the same order. Assuming that the bilinear pairing is a map $\hat{e} : G_1 \times G_1 \to G_2$ with the following properties:

Bilinearity: For any $X, Y \in G_1$ and $a, b \in Z_p$, we have $\hat{e}(aX, bY) = \hat{e}(X, Y)^{ab}$.

Non-degeneracy: For any $X, Y \in G_1$, $\hat{e}(X, Y) \neq 1_{G_2}$, where 1_{G_2} denotes the identity element of the group G_2.

Computability: There exists an efficient algorithm to compute $\hat{e}(X, Y)$ for any $X, Y \in G_1$.

2.2 Complexity Assumption

Considering the following computational hardness assumption in $< G_1, G_2, \hat{e} >$ as above, which is the basis of our scheme's security.

Definition 1. *Bilinear Diffie-Hellman (BDH) problem: Given* $< P, xP, yP, zP > \in G_1$ *with uniformly random choices of* $x, y, z \in Z_p^*$, *compute* $\hat{e}(P, P)^{xyz} \in G_2$.

The BDH assumption is that there is no polynomial time algorithm that can solve the BDH problem with non-negligible probability.

Let algorithm $\mathcal{A}$ be a BDH adversary who has an advantage ε in solving the BDH problem if $Pr[\mathcal{A}(< P, xP, yP, zP >) = \hat{e}(P, P)^{xyz}] = \varepsilon$. This probability is measured over random choices of $x, y, z \in Z_p^*$ and the point P. Adversary $\mathcal{A}$ solves the BDH problem with ε if and only if the advantage of $\mathcal{A}$ is greater than ε. The BDH problem is said to be ε-intractable if there is no algorithm that $\mathcal{A}$ solves this problem with ε.

2.3 Syntax

Different from the traditional CL-PKE scheme in [1], an Auth-CLE scheme consists of seven probabilistic, polyno-

mial time (PPT) algorithms: *Setup*, *Partial-Private-Key-Extract*, *Set-Secret-Value*, *Set-Private-Key*, *Set-Public-Key*, *Authenticated-Encrypt* and *Authenticated-Decrypt*. These algorithms are defined as follows:

Setup: On input a security parameter 1^k, this algorithm returns the system parameters *params*, master public key *mpk* and the master secret key *msk*. The system parameters *params* include the plaintext space $\mathcal{M}$ and the ciphertext space $\mathcal{C}$. After this algorithm is over, the KGC publishes *params* and *mpk*, then keeps the *msk* secretly.

Partial-Private-Key-Extract: On input *params*, *msk* and an identity ID for the entity, KGC executes this algorithm and returns the partial private key D_{ID} to entity via a confidential and authentic channel.

Set-Secret-Value: On input *params* and an identity ID, entity executes this algorithm and returns entity's secret value x_{ID}.

Set-Private-Key: On input *params*, entity's partial private key D_{ID} and secret value x_{ID}, this algorithm returns the entity's full private key SK_{ID}.

Set-Public-Key: On input *params*, *mpk* and entity's secret value x_{ID}, this algorithm re-turns the public key PK_{ID} to the entity.

Authenticated-Encrypt: Running by a sender. On input *params*, message $M \in \mathcal{M}$, the receiver's identity ID_R, the public keys of receiver PK_{ID_R} and the secret value of sender x_{ID_S}, this algorithm returns a ciphertext $C \in \mathcal{C}$.

Authenticated-Decrypt: Running this deterministic algorithm by a receiver. On input *params*, ciphertext $C \in \mathcal{C}$, the sender's public key PK_{ID_S} and a private key of receiver's SK_{ID_R}, this algorithm returns and verifies a message $M \in \mathcal{M}$, which is either a plaintext message or a "*Reject*" message.

2.4 Security model for Auth-CLE

In the Auth-CLE, there are two types of adversaries with different capabilities, Type I and Type II adversaries. A difference between these two attackers is that $\mathcal{A}_I$ does not have access to the master secret key of KGC while $\mathcal{A}_{II}$ does have. Specifically, the adversary $\mathcal{A}_I$ in Type I represents a normal third party attacker against the Auth-CLE scheme. That is, $\mathcal{A}_I$ is not allowed to access to the master secret key but it may request the public keys and replace them with the values of its choice. By contrast, adversary $\mathcal{A}_{II}$ in Type II represents a malicious KGC who can generate the partial private keys of users, and it is allowed to have access to the master secret key but not replace a public key.

Definition 2. *An Auth-CLE scheme is IND-CCA secure if neither polynomial bounded adversary $\mathcal{A}$ of Type I nor Type II has a non-negligible advantage against the challenger in the following game:*

Setup: The challenger $\mathcal{CH}$ takes a security parameter 1^k as inputs and runs the *Setup* algorithm, then it sends the resulting system parameters *params* and *mpk* to $\mathcal{A}$. If $\mathcal{A}$ is of Type I, $\mathcal{CH}$ keeps the master secret key *msk* to itself. Otherwise, returns *msk* to $\mathcal{A}$.

Phase 1: $\mathcal{A}$ is given access to the following oracles:

1) Partial-Key-Extract-Oracle: Upon receiving a partial key query for a user's identity ID, $\mathcal{CH}$ computes D_{ID} and returns it to $\mathcal{A}$. (Note that it is only useful to Type I adversary.)
2) Private-Key-Request-Oracle: Upon receiving a private key query for a user's identity ID, $\mathcal{CH}$ computes SK_{ID} and returns it to $\mathcal{A}$. It outputs $\perp$ (denotes failure) if the user's public key has been replaced (in the case of Type I adversary).
3) Public-Key-Request-Oracle: Upon receiving a public key query for a user's identity ID, $\mathcal{CH}$ computes PK_{ID} and returns it to $\mathcal{A}$.
4) Public-Key-Replace-Oracle: For identity ID and a valid public key, $\mathcal{A}$ replaces the associated user's public key with the new one of its choice (this is only for Type I adversary). The new value will be recorded and used by $\mathcal{CH}$ in the coming computations or responses to the adversary's queries.
5) Authenticated-Decryption-Oracle: On input a ciphertext and an identity, $\mathcal{CH}$ returns the correct decryption of ciphertext using the private key corresponding to the current value of the public key associated with the identity of the user, even if the corresponding public key for the user ID has been replaced.

Challenge Phase: Once $\mathcal{A}$ decides that *Phase 1* is over, it outputs and submits two messages (M_0, M_1), together with a challenge identity ID^* of uncorrupted secret key. Note that $\mathcal{A}$ is not allowed to know the private key of ID^* in anyway. The challenger $\mathcal{CH}$ picks a random bit $\beta \in 0, 1$ and computes C^*, which is the encryption of M_β under the current public key PK_{ID^*} for ID^*. If the output of the encryption is $\perp$, $\mathcal{A}$ immediately losses the game. Otherwise, C^* is delivered to $\mathcal{A}$.

Phase 2: $\mathcal{A}$ issues a second sequence of queries as in *Phase* 1. A decryption query on the challenge ciphertext for C^* the combination of ID^* and PK_{ID^*} is not allowed.

Guess: Finally, $\mathcal{A}$ outputs its guess β' for β. The adversary wins the game if $\beta' = \beta$ and the advantage of $\mathcal{A}$ in this game is defined to be $Adv(\mathcal{A}) = \mid Pr(\beta' = \beta) - \frac{1}{2} \mid$. The adversary $\mathcal{A}$ breaks an IND-CCA secure Auth-CLE scheme

with $(q_H, q_{par}, q_{pub}, q_{prv}, q_D, \varepsilon)$ if and only if the guessing advantage of $\mathcal{A}$ that makes q_H times to the random oracle $H(\cdot)$, q_{par} times *Partial-Key-Extract-Oracle*, q_{pub} times *Public-Key-Request-Oracle*, q_{prv} times *Private-Key-Request-Oracle* and q_D times *Authenticated-Decryption-Oracle* queries is greater than ε. The scheme is said to be $(q_H, q_{par}, q_{pub}, q_{prv}, q_D, \varepsilon)$-IND-CCA secure if there is no attacker $\mathcal{A}$ that breaks IND-CCA secure scheme with $(q_H, q_{par}, q_{pub}, q_{prv}, q_D, \varepsilon)$.

3 Our Protocol

In this section, we propose an Auth-CLE scheme to protect the integrity and confiden-tiality of data between the patient and the doctor.

3.1 Construction

The proposed Auth-CLE scheme consists of the following seven PPT algorithms.

Setup: Let G_1, G_2 be cyclic groups of prime order p with an arbitrary generator $P \in G_1$, $\hat{e} : G_1 \times G_1 \to G_2$ be a bilinear pairing. The MS selects $s \in Z_p^*$ at random and computes $P_{pub} = sP$ as master public key. Then, it chooses three collision resistant hash functions $H_1 : \{0,1\}^l \to G_1^*$, $H_2 : G_2 \to \{0,1\}^m$ and $H_3 : G_1 \times \{0,1\}^m \to G_1^*$, where l, m denotes the bit-length of identity and plaintext respectively. The system parameters are $params = \{G_1, G_2, \hat{e}, P, P_{pub}, H_1, H_2, H_3\}$ and the master secret key is $msk = s$.

Partial-Private-Key-Extract: On input patient's identity $\mathrm{ID}_P \in \{0,1\}^l$, MS computes $Q_{ID_P} = H_1(\mathrm{ID}_p)$ and sends the partial private key $D_{ID_P} = s \cdot Q_{ID_P} \in G_1^*$ to patient via a secure channel.

Set-Secret-Value: On input *params*, doctor's identity ID_D and patient's identity ID_P, doctor picks a secret value $\omega \in Z_p^*$ and returns $x_{ID_D} = \omega$ as his/her secret value. Correspondingly, the patient chooses $x_{ID_P} = \upsilon \in Z_p^*$ as his/her secret value.

Set-Private-Key: On input *params*, D_{ID_D} and x_{ID_D}, D_{ID_P} and x_{ID_P}, the doctor obtains the private key SK_{ID_D} by computing $SK_{ID_D} = \omega \cdot D_{ID_D}$. The patient gets his/her private key $SK_{ID_P} = \upsilon \cdot D_{ID_P}$.

Set-Public-Key: On input *params*, *mpk*, x_{ID_D} and x_{ID_P}, this algorithm returns $PK_{ID_D} = \omega P_{pub} = \omega sP$, $PK_{ID_P} = \upsilon P_{pub} = \upsilon sP$ as the public keys of doctor and patient respectively.

Authenticated-Encrypt: To encrypt $M \in \{0,1\}^m$, the doctor selects a random value $r \in Z_p^*$ and computes

$$Q_{ID_P} = H_1(ID_P),$$

$$\begin{aligned} c_1 &= r \cdot P, \\ c_2 &= M \oplus H_2(\hat{e}(H_1(ID_P), PK_{ID_P})^r), \\ c_3 &= H_3(\omega \cdot PK_{ID_P}, M). \end{aligned}$$

Then, set the ciphertext to $C = (c_1, c_2, c_3)$ and transmit it to the patient via the WMSN.

Authenticated-Decrypt: To decrypt ciphertext $C = (c_1, c_2, c_3)$ for the patient with private key SK_{ID_P}, he/she computes

$$M' = c_2 \oplus H_2(\hat{e}(SK_{ID_P}, c_1)).$$

After that, check $c_3 = H_3(\upsilon \cdot PK_{ID_D}, M')$. If not, reject the ciphertext. Otherwise, output M' as plaintext. Consistency of the scheme is clear since

$$\begin{aligned} \hat{e}(H_1(ID_P), PK_{ID_P})^r &= \hat{e}(H_1(ID_P), \upsilon sP)^r \\ &= \hat{e}(H_1(ID_P), P)^{\upsilon sr} \\ &= \hat{e}(\upsilon s H_1(ID_P), rP) \\ &= \hat{e}(SK_{ID_P}, c_1) \end{aligned}$$

by bilinearity.

3.2 Confidentiality Analysis

Theorem 1. *Given H_1, H_2 and H_3 are three collision resistant hash functions. The Auth-CLE scheme is IND-CCA secure in the random oracle model assuming that the BDH problem is intractable.*

This theorem following from two lemmas will show that our Auth-CLE scheme is secure against the Type I and Type II attacker whose behaviors are as described in **Definition 2**.

Lemma 1. *The Auth-CLE scheme is $(q_{H_1}, q_{H_2}, q_{par}, q_{pub}, q_{prv}, q_D, \varepsilon_I)$-IND-CCA secure against the Type I attacker $\mathcal{A}$ in the random oracle assuming the BDH problem is ε_I'-intractable, where*

$$\varepsilon_I' > \frac{1}{q_{H_2}}\left(\frac{2\varepsilon_I}{e(q_{prv} + q_{par} + 1)} - \frac{q_D q_{H_1}}{2^l} - \frac{q_D}{p}\right).$$

Proof. In this lemma, a Type I $\mathcal{A}$ models an "outside" adversary $\mathcal{A}_I$, who replaces the public key of arbitrary identities but cannot corrupt the master secret key.

Let $\mathcal{A}_I$ be a Type I IND-CCA adversary against our scheme. Suppose $\mathcal{A}_I$ has the advantage ε_I', makes q_{H_i} queries to random oracle $H_i (i = 1, 2)$ and q_D decryption queries. We show how to construct a PPT algorithm $\mathcal{B}$ to solve the BDH problem with instance of (P, aP, bP, cP) by interacting with $\mathcal{A}_I$.

At the beginning, $\mathcal{B}$ simulates the algorithm *Setup* for $\mathcal{A}_I$ by supplying $\mathcal{A}_I$ with $params = \{G_1, G_2, \hat{e}, P, P_{pub}, H_1, H_2, H_3\}$, where H_1, H_2 and H_3 are random oracles that will be controlled by $\mathcal{B}$. $\mathcal{B}$ chooses an index I uniformly at random with $1 \le I \le q_{H_1}$.

The $\mathcal{A}_I$ adversary may make queries of the random oracles $H_i(i = 1, 2)$ at any time during its attack. $\mathcal{B}$ responds as follows:

H_1 **queries:** $\mathcal{B}$ maintains a list of tuples $< \mathrm{ID}_i, Q_i, t_i >$ in H_1-**List** L_1. On receiving a query ID_i to H_1, $\mathcal{B}$ responds as follows:

1) If ID_i already appears on the list L_1 in a tuple $< \mathrm{ID}_i, Q_i, t_i >$, $\mathcal{B}$ responds Q_i as an answer.
2) Otherwise, if $i \neq I$, choose $t_i \in Z_p^*$ at random and compute $Q_i = t_i P$, add $< \mathrm{ID}_i, Q_i, t_i >$ to L_1, then return Q_i as an answer.
3) If $i = I$, add $< \mathrm{ID}_i, Q_i = aP, * >$ to L_1 and return $Q_i = aP$ as an answer (where "$*$" denotes the arbitrary value).

H_2 **queries:** $\mathcal{B}$ maintains a list of tuples $< \mathrm{ID}_i, e_i, R_i >$ in H_2-List L_2. On receiving a query $< \mathrm{ID}_i, e_i >$ to H_2, $\mathcal{B}$ responds as follows:

1) If ID_i already appears on the list L_2 in a tuple $< \mathrm{ID}_i, e_i, R_i >$, $\mathcal{B}$ responds R_i as an answer.
2) Otherwise, pick $R_i \in \{0,1\}^m$at random, add $< \mathrm{ID}_i, e_i, R_i >$ to L_2 and return R_i as an answer.

Phase 1: $\mathcal{A}_I$ issues a sequence of polynomially bounded number of the following oracle queries.

Partial-Key-Extract-Oracle: $\mathcal{B}$ maintains a **PartialKeyList** of tuples $< \mathrm{ID}_i, D_i >$. On receiving a query ID_i, $\mathcal{B}$ responds as follows:

1) If $< \mathrm{ID}_i, D_i >$ exist in **PartialKeyList**, return D_i as an answer.
2) Otherwise, pick i at random, so that $\Pr[i \neq I]=\delta$ (δ will be determined later.). If $i \neq I$, search L_1 for a tuple $< \mathrm{ID}_i, Q_i, t_i >$, compute $D_i = t_i P_{pub}$, add $< \mathrm{ID}_i, D_i >$ to the **PartialKeyList** and return D_i as an answer.
3) If $i = I$, return "Abort" and terminate.

Private-Key-Request-Oracle: $\mathcal{B}$ maintains a **PrivateKeyList** of tuples $< \mathrm{ID}_i, x_i, D_i >$. On receiving a query ID_i, $\mathcal{B}$ responds as follows:

1) If $< \mathrm{ID}_i, x_i, D_i >$ exist in **PrivateKeyList**, return $< x_i, D_i >$ as an answer.
2) Otherwise, if $i \neq I$, run the simulation algorithm *Public-Key-Request-Oracle* to get a tuple $< \mathrm{ID}_i, x_i, PK_i >$ and *Partial-Key-Extract-Oracle* to get a tuple $< \mathrm{ID}_i, D_i >$, add $< \mathrm{ID}_i, x_i, D_i >$ to the **PrivateKeyList** and return $< x_i, D_i >$ as an answer. (Note that if the corresponding public key has been replaced, such a private key query is not allowed.)
3) If $i = I$, return "Abort" and terminate.

Public-Key-Request-Oracle: $\mathcal{B}$ maintains a **PublicKeyList** of tuples $< \mathrm{ID}_i, x_i, PK_i >$. On receiving a query ID_i, $\mathcal{B}$ responds as follows:

1) If $< \mathrm{ID}_i, x_i, PK_i >$ exist in **PublicKeyList**, return PK_i as an answer.
2) Otherwise, if $i \neq I$ choose $x_i \in Z_p^*$ and compute $PK_i = x_i P_{pub} = bP$, add $< \mathrm{ID}_i, x_i, PK_i >$ to the **PublicKeyList** and return PK_i as an answer.
3) If $i = I$, add $< \mathrm{ID}_i, *, PK_i = bQ_i >$ to **PublicKeyList** and return PK_i as an answer.

Public-Key-Replace-Oracle: $\mathcal{A}_I$ may replace any public key with a new value of its choice and $\mathcal{B}$ records all the changes.

Auth-Decryption-Oracle: On receiving a query $< \mathrm{ID}_i, PK_i, C >$, where $C = (c_1, c_2, c_3)$. $\mathcal{B}$ responds as follows:

1) If $i \neq I$ and PK_i is the correct public key (not a replaced one), $\mathcal{B}$ decrypts C by using the corresponding private key.
2) Otherwise, search L_2 for a tuple $< \mathrm{ID}_i, e_i, R_i >$. If such a tuple exists, $\mathcal{B}$ retrieves the related R_i to compute $M = c_2 \oplus R_i$ and returns M as an answer.
3) Otherwise, $\mathcal{B}$ picks $R_i \in \{0,1\}^m$ at random, computes $M = c_2 \oplus R_i$ and returns M as an answer. Add $< \mathrm{ID}_i, e_i, R_i >$ to L_2.

Challenge Phase: $\mathcal{A}_I$ then outputs two messages (M_0, M_1) and a challenge identity ID^*. On receiving a challenge query $< \mathrm{ID}^*, (M_0, M_1) >$:

1) If $\mathrm{ID}^* \neq \mathrm{ID}_i$, $\mathcal{B}$ aborts the game.
2) Otherwise, $\mathcal{B}$ sets $c_1^* = cP$ and defines $c_2^* = H_2(\hat{e}(SK_{ID^*}, c_1^*)) \oplus M_\beta$, $c_3^* = H_3(\omega^* \cdot PK_{ID'}, M_\beta)$ (note that $\mathcal{B}$ does not know c and ω^*, $PK_{ID'}$ is the sender's public key), returns $C^* = (c_1^*, c_2^*, c_3^*)$ as a target ciphertext.

Phase 2: $\mathcal{A}_I$ requests in the same way as in *Phase* 1. Moreover, no *Private-Key-Request-Oracle* on ID^* is allowed and no *Auth-Decryption-Oracle* can be made on the ciphertext C^* for the combination of identity ID^* and public key PK_{ID^*} that encrypted plaintext M_β.

Guess: $\mathcal{A}_I$ should make a guess β' for β. The adversary wins the game if $\beta' = \beta$. Then, $\mathcal{B}$ will be able to solve the BDH problem by computing

$$\hat{e}(PK_i, c_1^*) = \hat{e}(bQ_i, cP) = \hat{e}(baP, cP) = \hat{e}(P, P)^{abc}.$$

Analysis: By $\mathbf{Ask}H_2^*$, we denote the event that (ID_i^*, e_i^*) has been queried to H_2. Also, by $\mathbf{Ask}H_1^*$, we denote the event that ID_i^* has been queried to H_1. If $\mathbf{Ask}H_2^*$ happens, $\mathcal{B}$ will be able to solve the BDH problem by

choosing a tuple $< \mathrm{ID}_i, e_i, R_i >$ from L_2 and computing $H_2(e_i)$ with the probability at least $\frac{1}{q_{H_2}}$. Hence, we have $\varepsilon'_I \geq \frac{1}{q_{H_2}} \Pr[\mathbf{Ask}H_2^*]$.

It is easy to notice that if $\mathcal{B}$ does not abort these oracles, the simulations of *Partial-Key-Extract-Oracle*, *Private-Key-Request-Oracle*, *Public-Key-Request-Oracle* and the simulated target ciphertext is identically distributed as the real one from the construction.

Then, we evaluate the simulation of *Auth-Decryption-Oracle*. If a public key PK_i has not been replaced nor PK_i has not been produced by reselecting $x_i \in Z_p^*$, the simulation is perfect as $\mathcal{B}$ knows the private key SK_i corresponding to PK_i. Otherwise, a simulation error may occur while $\mathcal{B}$ running the decryption oracle simulation specified above. Let **DecErr** be this event. Suppose $< \mathrm{ID}_i, PK_i, C >$, where $C = (c_1, c_2, c_3)$ and $PK_i = x_i P_{pub}$, has been issued as a valid decryption query. Even if C is valid, there is a possibility that C can be produced without querying (ID_i, e_i) to H_2.

Let **Valid** be an event that C is valid, $\mathbf{Ask}H_2$ and $\mathbf{Ask}H_1$ respectively be events that (ID_i, e_i) has been queried to H_2 and ID_i has been queried to H_1. Since **DecErr** is an event that $\mathbf{Valid}|\neg\mathbf{Ask}H_2$ happens during the entire simulation and q_D *Auth-Decryption-Oracle* queries are made, we have $\Pr[\mathbf{DecErr}]=q_D\Pr[\mathbf{Valid}|\neg\mathbf{Ask}H_2]$. However,

$$\begin{aligned}\Pr[\mathbf{Valid}|\neg\mathbf{Ask}H_2] &\leq \Pr[\mathbf{Valid} \wedge \mathbf{Ask}H_1|\neg\mathbf{Ask}H_2] \\ &+ \Pr[\mathbf{Valid} \wedge \neg\mathbf{Ask}H_1|\neg\mathbf{Ask}H_2] \\ &\leq \Pr[\mathbf{Ask}H_1|\neg\mathbf{Ask}H_2] \\ &+ \Pr[\mathbf{Valid}|\neg\mathbf{Ask}H_1 \wedge \neg\mathbf{Ask}H_2] \\ &\leq \frac{q_{H_1}}{2^l} + \frac{1}{p}\end{aligned}$$

Let the event $(\mathbf{Ask}H_2^* \vee \mathbf{DecErr})|\neg\mathbf{Abort}$ be denoted by $\mathbf{E}$, where **Abort** denotes an event that $\mathcal{B}$ aborts during the simulation. The probability $\neg\mathbf{Abort}$ that happens is given by $\delta^{q_{prv}+q_{par}}(1-\delta)$ which is maximized at $\delta = 1 - \frac{1}{q_{prv}+q_{par}+1}$. Hence, we have $\Pr[\neg\mathbf{Abort}] \leq \frac{1}{e(q_{prv}+q_{par}+1)}$, where e denotes the base of the natural logarithm.

If $\mathbf{E}$ does not happen, it is clear that $\mathcal{A}_I$ does not gain any advantage greater than $\frac{1}{2}$ to guess β due to the randomness of the output of the random oracle H_2. Namely, we have $\Pr[\beta' = \beta|\neg\mathbf{E}] \leq \frac{1}{2}$.

By Definition 2, we have

$$\begin{aligned}\varepsilon_I &< |\Pr[\beta' = \beta] - \frac{1}{2}| \\ &= |\Pr[\beta' = \beta|\neg\mathbf{E}]\Pr[\neg\mathbf{E}] + \Pr[\beta' = \beta|\mathbf{E}]\Pr[\mathbf{E}] - \frac{1}{2}| \\ &\leq |\frac{1}{2}\Pr[\neg\mathbf{E}] + \Pr[\mathbf{E}] - \frac{1}{2}| \\ &= |\frac{1}{2}(1 - \Pr[\mathbf{E}]) + \Pr[\mathbf{E}] - \frac{1}{2}| \\ &= \frac{1}{2}\Pr[\mathbf{E}] \\ &\leq \frac{\Pr[\mathbf{Ask}H_2^*] + \Pr[\mathbf{Ask}H_1^*|\neg\mathbf{Ask}H_2^*] + \Pr[\mathbf{DecErr}]}{2\Pr[\neg\mathbf{Abort}]} \\ &\leq \frac{e(q_{prv}+q_{par}+1)}{2}(q_{H_2}\varepsilon'_I + \frac{q_D q_{H_1}}{2^l} + \frac{q_D}{p})\end{aligned}$$

Consequently, we obtain

$$\varepsilon'_I > \frac{1}{q_{H_2}}(\frac{2\varepsilon_I}{e(q_{prv}+q_{par}+1)} - \frac{q_D q_{H_1}}{2^l} - \frac{q_D}{p}).$$

□

Lemma 2. *The Auth-CLE scheme is* $(q_{H_1}, q_{H_2}, q_{pub}, q_{prv}, q_D, \varepsilon_{II})$*-IND-CCA secure against the Type II attacker* $\mathcal{A}$ *in the random oracle assuming the BDH problem is* ε'_{II}*-intractable, where*

$$\varepsilon'_{II} > \frac{1}{q_{H_2}}(\frac{2\varepsilon_{II}}{e(q_{prv}+1)} - \frac{q_D q_{H_1}}{2^l} - \frac{q_D}{p}).$$

Proof. In this lemma, a Type II $\mathcal{A}$ models an "inside" adversary $\mathcal{A}_{II}$, who has access to *msk* but cannot replace public key of entity.

Let $\mathcal{A}_{II}$ be a Type II IND-CCA adversary against our scheme. Suppose $\mathcal{A}_{II}$ has the advantage ε'_{II}, makes q_{H_i} queries to random oracle $H_i (i = 1, 2)$ and q_D decryption queries. We show how to construct a PPT algorithm $\mathcal{B}$ to solve the BDH problem with instance of (P, aP, bP, cP) by interacting with $\mathcal{A}_{II}$.

At the beginning, $\mathcal{B}$ simulates the algorithm *Setup* for $\mathcal{A}_{II}$ by supplying $\mathcal{A}_{II}$ with $params = \{G_1, G_2, \hat{e}, P, P_{pub}, H_1, H_2, H_3\}$, where H_1, H_2 and H_3 are random oracles that will be controlled by $\mathcal{B}$. $\mathcal{B}$ chooses an index I uniformly at random with $1 \leq I \leq q_{H_1}$.

The adversary $\mathcal{A}_{II}$ may make queries of the random oracles $H_i (i = 1, 2)$ at any time during its attack. $\mathcal{B}$ responds as follows:

H_1 **queries:** $\mathcal{B}$ maintains a list of tuples $< \mathrm{ID}_i, Q_i >$ in H_1-**List** L_1. On receiving a query ID_i to H_1, $\mathcal{B}$ responds as follows:

1) If ID_i already appears on the list L_1 in a tuple $< \mathrm{ID}_i, Q_i >$, $\mathcal{B}$ responds Q_i as an answer.
2) Otherwise, if $i \neq I$, choose $Q_i \in G_1^*$ at random and add $< \mathrm{ID}_i, Q_i >$ to L_1, return Q_i as an answer.

3) If $i = I$, add $< \mathrm{ID}_i, Q_i = aP, * >$ to L_1 and return $Q_i = aP$ as an answer (where "$*$" denotes the arbitrary value).

H_2 queries: $\mathcal{B}$ maintains a list of tuples $< \mathrm{ID}_i, e_i, R_i >$ in H_2-List L_2. On receiving a query $< \mathrm{ID}_i, e_i >$ to H_2, $\mathcal{B}$ responds as follows:

1) If ID_i already appears on the list L_2 in a tuple $< \mathrm{ID}_i, e_i, R_i >$, $\mathcal{B}$ responds R_i as an answer.
2) Otherwise, pick $R_i \in \{0,1\}^m$ at random, add $< \mathrm{ID}_i, e_i, R_i >$ to L_2 and return R_i as an answer.

Phase 1: $\mathcal{A}_{II}$ issues a sequence of polynomially bounded number of the following oracle queries.

Private-Key-Request-Oracle: $\mathcal{B}$ maintains a **PrivateKeyList** of tuples $< \mathrm{ID}_i, x_i, D_i >$. On receiving a query ID_i, $\mathcal{B}$ responds as follows:

1) If $< \mathrm{ID}_i, x_i, D_i >$ exist in **PrivateKeyList**, return $< x_i, D_i >$ as an answer.
2) Otherwise, pick i at random, so that $\Pr[i \neq I] = \delta$ (δ is the same as it in the proof of **Lemma 1**). If $i \neq I$, run the simulation algorithm *Public-Key-Request-Oracle* to get a tuple $< \mathrm{ID}_i, x_i, PK_i >$, pick $s \in Z_p^*$ and compute $D_i = sH_1(\mathrm{ID}_i)$, add $< \mathrm{ID}_i, x_i, D_i >$ to the **PrivateKeyList** and return $< x_i, D_i >$ as an answer.
3) If $i = I$, return "Abort" and terminate.

Public-Key-Request-Oracle: $\mathcal{B}$ maintains a **PublicKeyList** of tuples $< \mathrm{ID}_i, x_i, PK_i >$. On receiving a query ID_i, $\mathcal{B}$ responds as follows:

1) If $< \mathrm{ID}_i, x_i, PK_i >$ exist in **PublicKeyList**, return PK_i as an answer.
2) Otherwise, if $i \neq I$ choose $x_i \in Z_p^*$, compute $PK_i = x_i P$, add $< \mathrm{ID}_i, x_i, PK_i >$ to the **PublicKeyList**, return PK_i as an answer.
3) If $i = I$, set $PK_i = bP_{pub} = sbP$, add $< \mathrm{ID}_i, *, PK_i >$ to **PublicKeyList** and return PK_i as an answer.

Auth-Decryption-Oracle: On receiving a query $< \mathrm{ID}_i, PK_i, C >$, where $C = (c_1, c_2, c_3)$. $\mathcal{B}$ responds as follows:

1) If $i \neq I$, $\mathcal{B}$ decrypts C by using the private key $< x_i, D_i >$.
2) Otherwise, search L_2 for a tuple $< \mathrm{ID}_i, e_i, R_i >$. If such a tuple exists, $\mathcal{B}$ retrieves the related R_i to compute $M = c_2 \oplus R_i$ and returns M as an answer.
3) Otherwise, $\mathcal{B}$ picks $R_i \in \{0,1\}^m$ at random, computes $M = c_2 \oplus R_i$ and returns M as an answer. Add $< \mathrm{ID}_i, e_i, R_i >$ to L_2.

Challenge Phase: $\mathcal{A}_{II}$ then outputs two messages (M_0, M_1) and a challenge identity ID^*. On receiving a challenge query $< \mathrm{ID}^*, (M_0, M_1) >$:

1) If $\mathrm{ID}^* \neq \mathrm{ID}_i$, $\mathcal{B}$ aborts the game.
2) Otherwise, $\mathcal{B}$ sets $c_1^* = s^{-1}cP$ and defines $c_2^* = H_2(\hat{e}(SK_{ID^*}, c_1^*)) \oplus M_\beta$, $c_3^* = H_3(\omega^* \cdot PK_{ID'}, M_\beta)$ (note that $\mathcal{B}$ does not know c and ω^*, $PK_{ID'}$ is the sender's public key), returns $C^* = (c_1^*, c_2^*, c_3^*)$ as a target ciphertext.

Phase 2: $\mathcal{A}_{II}$ requests the same methods that it used in *Phase* 1. Moreover, no *Private-Key-Request-Oracle* on ID^* is allowed and no *Auth-Decryption-Oracle* can be made on the ciphertext C^* for the combination of identity ID^* and public key PK_{ID^*} that encrypted plaintext M_β.

Guess: $\mathcal{A}_{II}$ should make a guess β' for β. The adversary wins the game if $\beta' = \beta$. Then, $\mathcal{B}$ will be able to solve the BDH problem by computing

$\hat{e}(aPK_{ID^*}, c_1^*) = \hat{e}(absP, s^{-1}cP) = \hat{e}(P,P)^{abss^{-1}c} = \hat{e}(P,P)^{abc}$.

Analysis: Similar to *Analysis* in the proof of **Lemma 1**.

□

Consequently, we obtain

$$\varepsilon'_{II} > \frac{1}{q_{H_2}}\left(\frac{2\varepsilon_{II}}{e(q_{prv}+1)} - \frac{q_D q_{H_1}}{2^l} - \frac{q_D}{p}\right).$$

These two lemmas complete the proof of **Theorem 1**.

3.3 Unforgeability Analysis

Theorem 2. *Suppose H_1, H_2 and H_3 are three collision resistant hash functions, and $\mathcal{A}$ is an adversary that can forge a ciphertext with advantage ε by making q_{H_3} queries to random oracle H_3 and q_D queries to* Auth-Decryption-Oracle. *Then, there exists a PPT algorithm $\mathcal{B}$ that can solve the BDH problem with advantage at least*

$$Adv(\mathcal{B}) = \frac{\varepsilon}{(\frac{2}{q_{H_3}(q_{H_3}-1)})^2 q_D}$$

Proof. We show how to construct a PPT algorithm $\mathcal{B}$ to solve the BDH problem with instance of (P, aP, bP, cP) by interacting with $\mathcal{A}$.

At the beginning, $\mathcal{B}$ simulates the algorithm *Setup* for $\mathcal{A}$ by supplying $\mathcal{A}$ with $params = \{G_1, G_2, \hat{e}, P, P_{pub}, H_1, H_2, H_3\}$, where H_1, H_2 and H_3 are random oracles that will be controlled by $\mathcal{B}$. There are two lists L_i that store the answers on H_i queries ($i = 2, 3$) and a list of possible bilinear pairing answers L_e.

H_2 queries: $\mathcal{B}$ maintains a list of tuples $< \mathrm{ID}, e, R >$ in H_2-**List** L_2. On receiving a query $< \mathrm{ID}, e >$ to H_2, $\mathcal{B}$ responds as follows:

1) If ID already appears on the list L_2 in a tuple $< \mathrm{ID}, e, R >$, $\mathcal{B}$ responds R as an answer.
2) Otherwise, pick $R \in \{0,1\}^m$ randomly, add $< \mathrm{ID}, e, R >$ to L_2 and return R as an answer.

H_3 queries: $\mathcal{B}$ supplies $\mathcal{A}$ with (P, aP) and sets $1 \leq i \neq j \leq q_{H_3}$. $\mathcal{B}$ responds as follows:

1) If it is the ith query, respond with bP and call ID a guessed identity.
2) If it is the jth query, respond with cP and call ID a guessed identity.
3) Otherwise, choose a random $W \in G_1^*$ and add $< T, M, W >$ to L_3, return $W = dP$ as an answer.

Private-Key-Request-Oracle: On input identity ID, $\mathcal{B}$ responds as follows:

1) If ID is a guessed identity, $\mathcal{B}$ fails.
2) Otherwise, the list L_3 must contain the tuple $< T, M, W >$ for some $d \in Z_p^*$ and $\mathcal{B}$ outputs adP as its private key.

Authenticated-Encryption-Oracle: Suppose $\mathcal{A}$ issues an encryption query for a plaintext M between doctor ID_D and patient ID_P.

1) If ID_D and patient ID_P are the guessed identity,$\mathcal{B}$ picks three random values $\{R, T\} \in G_1^*$ and $S \in \{0,1\}^m$, return $C = (R, S, H_3(T, M))$ as a ciphertext.
2) Otherwise, assume ID_P is not a guessed identity, the list L_3 must contain the tuple $< T, M, W >$ for some $d \in Z_p^*$. Then, the patient's private key is adP and the ciphertext is computed as described in the *Authenticated-Encrypt*. Return the ciphertext to $\mathcal{A}$.

Authenticated-Decryption-Oracle: Suppose $\mathcal{A}$ issues a decryption query for a ciphertext $C = (c_1, c_2, c_3)$ between doctor and patient.

1) If ID_P is the guessed identity, L_2 is examined for an entry of the form $< \mathrm{ID}_P, e, R >$ for some e. If such an entry exists, e is added to the list L_e. $\mathcal{A}$ is notified that C is invalid even if it is valid.
2) Otherwise, assume ID_P is not a guessed identity, the list L_3 must contain the tuple $< T, M, W >$ for some $d \in Z_p^*$ and adP is its private key. Then, the ciphertext is decrypted as described in the *Authenticated-Decryption* algorithm. If this ciphertext is valid, the correspondingly plaintext is given to $\mathcal{A}$ and it wins.

Eventually, $\mathcal{A}$ decides that the game is over. If the list L_e is empty, $\mathcal{B}$ fails. Otherwise, $\mathcal{B}$ outputs a random element of L_e.

Analysis: The probability that $\mathcal{A}$ has never issued *Private-Key-Request-Oracle* query on one of the guessed identity is at least $C_{q_{H_3}}^2$. If $\mathcal{A}$ has submitted a valid ciphertext, it forges a ciphertext successfully between the guessed identity with at least probability $C_{q_{H_3}}^2$ (but this ciphertext is actually invalid).

If $e = \hat{e}(P,P)^{abc}$ is not on the list L_e, $\mathcal{A}$ cannot generate a correct forgery for H_2 is a random oracle. Therefore, the probability that $\mathcal{A}$ queries $H_2(e)$ is at least ε. If this happens, $\mathcal{B}$ cannot fail and output the correct value with probability at $\frac{1}{q_D}$. Then, we have

$$Adv(\mathcal{B}) = \frac{\varepsilon}{(\frac{2}{q_{H_3}(q_{H_3}-1)})^2 q_D}.$$

□

4 Comparisons

4.1 Computation Costs

First, we evaluate the computational cost of our scheme and others [2, 11, 17, 25] through combined implementation and simulation. We test the cryptographic operations in bilinear pairing, exponentiation and scalar multiplication (without considering the addition of two points, the hash function and exclusive-OR operations), and detailed time results on a PC with the Intel Core i5-2400 at a frequency of 3.1 GHz with 3 GB memory and Windows XP operating system, using the MIRACL (Version 5.6.1, [16]). For bilinear pairing, in order to implement it in practice efficiently, we employ the Fast-Tate-Pairing in MIRACL, which is defined over the MNT curve E/F_q [13] with characteristic a 160-bit prime and embedding degree 4. For ECC-based protocols, we choose the recommended parameters "secp192k1" [20]. Furthermore, we denote the length of an element in a multiplicative group to be 1024-bit. Based on the above parameter settings, the average running time of each operation in 100 times is obtained and demonstrated in Table 1. Then, the total running time to finish one round of "*Authenticated-Encrypt and Decrypt*" is illustrated in Table 2. For example, in *Authenticated-Encrypt* and *Decrypt* of our scheme, there are two bilinear pairing operations, one exponentiation and three scalar multication in the additive cyclic group in all; thus the total operation time is $2\times2.65+1\times3.75+3\times0.78 = 11.39$ ms. These indicate our scheme is more scalable and efficient than existing works.

4.2 Communication Costs

Next, we analyze the communication cost in terms of the bandwidth of the transmitted ciphertext (or signcrypted text). Suppose that the output of one-way hash function is 160-bit. In our protocol, the ciphertext contains two hash values and one point, thus the bandwidth of it is $(160 \times 2 + 192)/8 = 64$ bytes. In Barbosa and Farshim's

Table 1: Cryptography operation time

Fast-Tate-Pairing	Exponentiation	Scalar Multiplication
2.65 ms	3.75 ms	0.78 ms

Table 2: Comparison of the related schemes

Scheme	Auth-Enc	Auth-Dec	Bandwidth	Total Time
[2]	1P+1E+4S	5P+1S	68 bytes	23.55 ms
[11]	4E	5P	532 bytes	28.25 ms
[17]	5E	7E	276 bytes	45.00 ms
[25]	1P+4E+3S	3P+4S	108 bytes	31.06 ms
Ours	1P+1E+2S	1P+1S	64 bytes	11.39 ms

scheme [2], the signcrypted text contains two points and one hash, the bandwidth of it is $(192 \times 2 + 160)/8 = 68$ bytes. In Liu et al.'s scheme [11], the signcrypted text contains four elements of multiplicative group and one bilinear pairing, the bandwidth of it is $(1024 \times 4 + 160)/8 = 532$ bytes. In the scheme of [17], the signcrypted text contains two elements of multiplicative group and one hash value, the bandwidth of it is $(1024 \times 2 + 160)/8 = 276$ bytes. At last, in Wu and Chen's scheme [25], the signcrypted text contains two points, two hash values and one element in additive group, and therefore the bandwidth of it is $(192 \times 2 + 160 \times 2 + 160)/8 = 108$ bytes. The detailed comparison results are also listed in Table 2, which shows that the bandwidth of our scheme is the smallest.

5 Conclusions

In this paper, we propose an authenticated certificateless encryption scheme to ensure the confidentiality and integrity of the transmitted information between patient and doctor in THS, which satisfies the privacy requirements of HIPAA. Moreover, it is proved that our protocol is IND-CCA secure and the information cannot be forged in the random oracle model, relative to the hardness of the BDH problem. By the evaluation and simulation, a comparison in Table 2 concludes that the proposed scheme is advantageous over the related schemes in computation and communication cost.

Acknowledgments

This study was supported by the National Nature Science Foundation of China under grant NSFC 11501343, and the Open Foundation of State key Laboratory of Networking and Switching Technology (Beijing University of Posts and Telecommunications) under grant SKLNST-2016-2-11. The authors gratefully acknowledge the anonymous reviewers for their valuable comments.

References

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

[2] M. Barbosa and P. Farshim, "Certificateless signcryption," in *Proceedings of the 2008 ACM symposium on Information, Computer and Communications Security*, pp. 369–372, 2008.

[3] L. J. Blumberg, L. M. Nichols, "The health insurance portability and accountability act of 1996: Summary of provisions and anticipated effects", *Journal of Medical Practical Management*, vol. 14, no. 1, pp. pp. 13-8, 1998.

[4] D. Boneh and M. Franklin, "Identity based encryption from the Weil pairing," in *Advances in Cryptology (CRYPTO'01)*, LNCS, vol. 2139, pp. 213–229, 2001.

[5] Z. H. Cheng and R. Comley, "Efficient certificateless public key encryption," *IACR Cryptology ePrint Archive*, 2005. (`http://eprint.iacr.org/2005/012.pdf`)

[6] R. Guo, Q. Y. Wen, H. X. Shi, Z. P. Jin and H. Zhang, "An efficient and provably-secure certificateless public key encryption scheme for telecare medicine in-formation systems," *Journal of Medical Systems*, vol. 37, pp. 9965, 2013.

[7] D. B. He, J. H. Chen and R. Zhang, "A more secure authentication scheme for telecare medicine information systems," *Journal of Medical Systems*, vol. 36, pp. 1989–1995, 2012.

[8] H. F. Huang, P. H. Lin, and M. H. Tsai, "Convertible Multi-authenticated Encryption Scheme for Data Communication," *International Journal of Network Security*, vol. 17, no. 1, pp. 40–48, 2015.

[9] R. S. Istepanian, E. Jovanov and Y. T. Zhang, "Guest editorial introduction to the special section on m-health: Beyond seamless mobility and global wireless health-care connectivity," *IEEE Transac-*

tions on Information Technology in Biomedicine, vol. 8, no. 4, pp. 405–414, 2004.

[10] A. V. N. Krishna, A. H. Narayana, K. M. Vani, "Window method based cubic spline curve public key cryptography," *International Journal of Electronics and Information Engineering*, vol. 4, no. 2, pp. 94–102, 2016.

[11] Z. H. Liu, Y. P. Hu, X. S. Zhang and H. Ma, "Certificateless signcryption scheme in the standard model," *Information Sciences*, vol. 180, no. 3, pp. 452–464, 2010.

[12] B. Lynn, "Authenticated identity-based encryption," *IACR Cryptology ePrint Archive*, 2002. (`http://eprint.iacr.org/2002/072.pdf`)

[13] A. Miyaji, M. Nakavayashi and S. Takano, "New explicit conditions of elliptic curve traces for FR-reduction," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E84-A, no. 5, pp. 1234–1243, 2001.

[14] J. Moon, Y. Choi, J. Kim and D. Won, "An improvement of robust and efficient biometrics based password authentication scheme for telecare medicine information systems using extended chaotic maps," *Journal of Medical Systems*, vol. 40, no. 3, 2016.

[15] J. H. Oh, K. K. Lee and S. J. Moon, "How to solve key escrow and identity revocation in identity based encryption schemes," in *The First International Conference of Information Systems Security*, vol. 3803, pp. 290–303, 2005.

[16] M. Scott, *Miracl Library*, Apr. 15, 2017. (`http://certivox.com/`)

[17] S. S. D. Selvi, S. S. Vivek and C. P. Rangan, "Cryptanalysis of certificateless signcryption schemes and an efficient construction without pairing," *IACR Cryptology ePrint Archive*, 2009. (`http://eprint.iacr.org/2009/298.pdf`)

[18] S. S. D. Selvi, S. S. Vivek and C. P. Rangan, "Security weaknesses in two certificateless signcryption schemes," *IACR Cryptology ePrint Archive*, 2010. (`http://eprint.iacr.org/2010/092.pdf`)

[19] A. Shamir, "Identity-based cryptosystems and signature schemes," in *Advances in Cryptology (CRYPTO'84)*, LNCS, vol. 196, pp. 47–53, 1985.

[20] The Certicom Research, *SEC2: Recommended Elliptic Curve Domain Parameters*, Version 1.5, 2005.

[21] 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, 2017.

[22] K. K. Venkatasubramanian, A. Banerjee and S. K. S. Gupta, "Plethysmogram-based secure inter-sensor communication in body area networks," in *IEEE Military Communications Conference*, pp. 1–7, 2008.

[23] J. H. Wei, X. X. Hu and W. F. Liu, "An improved authentication scheme for telecare medicine information systems," *Journal of Medical Systems*, vol. 36, pp. 3597–3604, 2012.

[24] J. Weng, G. X. Yao, R. H. Deng, M. R. Chen and X. X. Li, "Cryptanalysis of a certificateless signcryption scheme in the standard model," *Information Sciences*, vol. 181, no. 3, pp. 661–667, 2011.

[25] C. H. Wu and Z. X. Chen, "A new efficient certificateless signcryption scheme," in *International Symposium on Information Science and Engineering*, vol. 1, pp. 661–664, 2008.

[26] Z. Y. Wu, Y. C. Lee, F. Lai, H. C. Lee and Y. Chung, "A secure authentication scheme for telecare medicine information systems," *Journal of Medical Systems*, vol. 36, pp. 1529–1535, 2012.

[27] Y. Zheng, "Digital signcryption or how to achieve cost (signature & encryption) $\leq$ cost(signature) + cost (encryption)," in *Advances in Cryptology (CRYPTO'97)*, LNCS, vol. 1294, pp. 165–179, 1997.

[28] Z. A. Zhu, "An efficient authentication scheme for telecare medicine information systems," *Journal of Medical Systems*, vol. 36, no. 6, pp. 3833–3838, 2012.

Biography

Rui Guo received the Ph.D degrees in the Department of State Key Laboratory of Networking and Switch Technology, Beijing University of Posts and Telecommunications in 2014. Now, he is a lecturer in National Engineering Laboratory for Wireless Security, Xi'an University of Posts and Telecommunications. His present research interests include cryptography, information security and WSN.

Huixian Shi received the B.S. and Ph.D degrees in Department of Mathematics and Information Science from Shaanxi Normal University, Xi'an, China, in 2007 and 2013, respectively. Now she is a associate professor in Department of Department of Mathematics and Information Science in Shaanxi normal University. Her present research interests include model checking, fuzzy logic and uncertainty reasoning.

Double Verifiable Lossless Secret Sharing Based on Hyper-chaos Generated Random Grid

Hang Gao[1], Mengting Hu[1], Tiegang Gao[2], and Renhong Cheng[1]
(Corresponding author: Tiegang Gao)
College of Computer and Control Engineering, Nankai University[1]
Tianjin Haihe Education Park, No. 38, Tongyan Road, Tianjin, China
(Email: gaotiegang@nankai.edu.cn)
College of Software, Nankai University[2]
Tianjin Haihe Education Park, No. 38, Tongyan Road, Tianjin, China
(Received July 27, 2016; revised and accepted Nov. 15 & Dec. 25, 2016)

Abstract

A novel multi secret sharing scheme is proposed in this paper. In the scheme, n secret images are firstly encrypted by hyper-chaos whose initial values are hash of the image itself. Then the encrypted images are shuffled into n confused images; lastly, the confused images are used to generate n sharing images using the random grid method. The proposed scheme has the advantage such that it can lossless restore the original secret image, and have the double verification ability, that is to say, it can verify whether the anyone of the sharing is modified, and it can also verify the whether the anyone of the original secret image is completely reconstructed. The above advantages make it especially suitable for secret sharing of important images such as medical and military images. Experimental results and some comparison analysis are given to testify the effectiveness of the proposed scheme.

Keywords: Hash-256; Hyper-chaos; Lossless Restoration; Random Grid; Visual Secret Sharing

1 Introduction

In general, a secret is regarded as safer when it is given two or more participants than that it is kept by only one person. Based on this kind of idea, Blakley and Shamir proposed the concept of secret sharing, respectively, in 1979 [2, 24]. As a kind of secret sharing, Naor and Shamir (1995) proposed a new type of secret sharing called visual cryptography (VC) or visual secret sharing (VSS) for images [21]. VC can encrypt a secret image into numerous meaningless sharing images, and anyone of the shared images does not reveal any information about the secret, In a general threshold VC scheme, a secret image is encoded into n meaningless random shares. The n shares are distributed to n corresponding participants. In order to reconstruct the original secret image, at least k or more participants are needed to share their shares, but any k?1 or fewer shares have no means to give clue about the secret.

In the last two decades, VC algorithm has aroused many interests among the researchers. Various proposals of the VC schemes for different situation have been proposed [6, 7, 9, 13, 20, 26, 31]. For example, people have presented extended VC schemes to encode the secret image into natural-looking shares [26, 31]; the progressive VC schemes is that the restored secret image can be shown in a progressive perceptual quality [9, 13]; prevention of cheating in VC is the scheme that break the misleading secret by dishonest participants [6, 7].

Another interesting VSS scheme is random grid (RG) based algorithm which has the advantages such as no pixel expansion in sharing secret. RG is firstly proposed by Kafri and Keren in 1987 [15]. Since then, some new researches have been made to probe extensive application scene of the RG based VSS scheme. Among them, Shyu proposed a RG-based scheme for gray-level and color images, and Chen et al. proposed a Threshold RG-based scheme for color images [3, 25]. In the study of meaningful shares in RG-based VCC, Chen and Tsao proposed a novel RG-based VSS scheme by skillfully designing a procedure of distinguishing different light transmissions on shared images, and the visual quality between meaningful random-grids and superimposed results can be adjusted to be more friendly [4]. Abd El-Latif, et al. proposed a scheme combines the error diffusion technique, RG and chaotic encryption to encode a secret binary image into meaningful shadow images [1]. Yan et al. proposed a generalized RG-based VC with meaningful shares which can support (k,n) threshold and provide adaptive visual quality, at the cost of slightly decreasing visual quality of shared images [28], and Chen proposed a lossless RG-based VSS scheme [5].

In recent years, security of VSS in different scenarios has aroused people's attentions [12, 16]. As different ap-

plication situations have different requirements. In 2006, Horng et al. showed that cheating is possible in (k,n) VSS scheme, the cheating problem will happen when dishonest participants collude to cheat honest ones by enabling the latter to accept the wrong secret information generated by the former [12]. Lee demonstrated that the inside collusion attacks to the RG-based VSS schemes is possible, and gives some examples for verifying feasibility of cheating [16]. In order to cope with cheating in VSS, some people proposed schemes to verify the genuineness of shares through generating extra verification shares [8, 18, 19, 27], for example, Wu et al. proposed a cheating immune method to provided extra ability of cheat-preventing for RG-based VSS [27]. Lin et al. exploited the hybrid codebook to hide the additional verification images into the share images to prevent cheating [18]; in order not to resort to any additional dedicated verification share, Lin et al. proposed a VC-based scheme with the ability to prevent cheating, the scheme designed an authentication pattern stamping process to detect the faked shares provided by malicious participants [19].

In real application of image encryption, for medical, satellite and military image, it is demanded that the encryption algorithm is safe. Moreover, the decrypted image must be lossless restored and verifiable. Inspired by the above algorithm, we proposed a RG-based VSS scheme for grey image, the main advantages of scheme include mini pixel expansions in sharing secret, verification of sharing secret and verification of restored secret image. Besides, the proposed algorithm can lossless restore secret image, the above features of the proposed scheme make it especially suitable for application of important images such as medical, satellite and military.

The rest of the paper is organized as follows. Some preliminaries are introduced in Section 2. The proposed scheme is describe in Section 3 and experimental results and analysis are given in Section 4, in the last, some conclusions are described in Section 5.

2 Preliminaries

In this section, some relative technology such as Hash function, RG-based VSS and the hyper-chaotic system based encryption algorithm are firstly introduced.

2.1 Hash Function

In computer science, one-way functions is a function that is easy to compute on every input, but it is very difficult to compute their inverse functions. That is to say, for a given data x, it is easy to calculate one-way function of x, on the other hand, knowing the value of one-way of x, it is quite difficult to calculate the value of x.

Hash functions is one-way function that converts input messages of any length into output sequences of fixed length, the output sequence is often called a hash value. A hash function has the properties of sensitivity to initial conditions, diffusion and confusion, collision resistance. This means that it should be very difficult to find two different sequences that produce the same hash value. These characteristics make it can be used for verification of data integrity. It has played an important role in the field of information security [34].

At present, some typical hash function includes MD5, SHA-1 and SHA-2. Among them, and the most important hash functions is the SHA family, which shares the same functional structure with some variation in the internal operations, message size, message block size, word size, number of security bits and message hash size [22]. In this paper, SHA-1 with the 256 bits output is used for generation of verification information.

2.2 RG-based VSS

The RG-based VSS was firstly proposed by Kafri, in the scheme, for a given secret binary image, two random grids and were generated, and anyone of them will leak no information about the binary image individually, yet they reveal the secret binary image when the two grids were superimposed. One of the three algorithms is given in follows.

1) For a binary secret image B with the size of $N \times M$, generate a random grid R_1 which includes only 0 and 1, the size of R_1 is the same as that of secret image.

2) For every pixel value $B_{(i,j)}$, $i = 1, 2, \cdots, N$, $j = 1, 2, \cdots, M$, if $B_{(i,j)}$ is 0, then the value of $R_2(i, j)$ is equal to $R_1(i, j)$, else $R_2(i, j)$ is equal to the complement of $R_1(i, j)$.

3) R_1 and R_2 are the random grid.

The original secret binary image can be restored by superimposing R_1 and R_2 together. For example, in Figure 1, Figure 1(a) is the original secret image, Figures 1(b) and (c) are random grids, and the Figure 1(d) is the restored image by Figures 1(b) and (c).

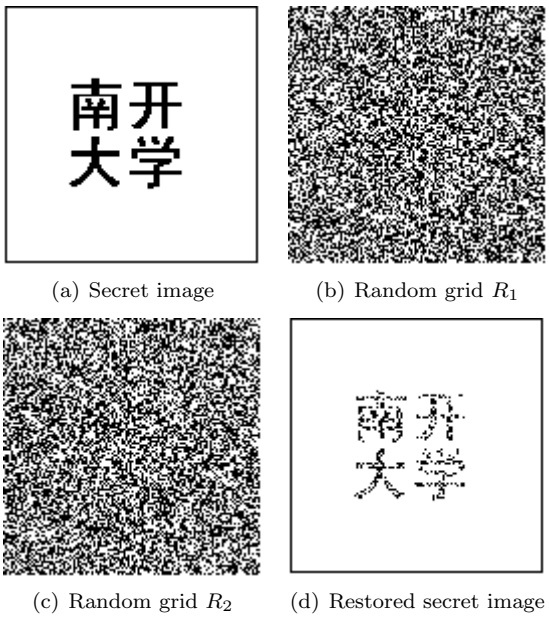
2.3 The Hyper-chaotic System

In the proposed scheme, a hyper-chaos system which is modeled by Equation (1) is used for generation of random grid.

$$\begin{cases} \dot{x}_1 = a(x_2 - x_1) \\ \dot{x}_2 = -x_1 x_3 + d x_1 + c x_2 - x_4 \\ \dot{x}_3 = -x_1 x_2 - b x_3 \\ \dot{x}_4 = x_1 + k \end{cases} \tag{1}$$

where a, b, c, d and k are parameters, when $a = 36, b = 3, c = 28, d = -16$ and $-0.7 \leq k \leq 0.7$, the system is hyper-chaotic, and its attractors are shown in Figure 2 with parameters $a = 36, b = 3, c = 28, d = -16$ and $k = 0.2$, its Lyapunov exponents are $\lambda_1 = 1.552, \lambda_2 = 0.023, \lambda_3 = 0, \lambda_4 = -12.573$.

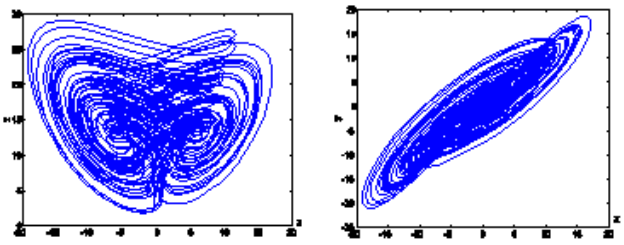
Because the hyper-chaos has two positive Lyapunov exponents, so the prediction time of a hyper-chaotic system



(a) Secret image (b) Random grid R_1

(c) Random grid R_2 (d) Restored secret image

Figure 1: Experimental results of random grid



(a) $x_1 - x_3$ plane (b) $x_1 - x_2$ plane

Figure 2: Hyper-chaos attractors of System (1) with $k = 0.2$

is shorter than that of a chaotic system [29], as a result, it is safer than chaos in security algorithm. For more detailed analysis of the complex dynamics of the system, please see relative reference [10].

As the hyper-chaos has the ergodicity, sensitive features on initial conditions and control parameters of chaotic maps and random-like behaviors, these features make it suitable for generating pseudo-random sequences and key sequences in cryptography [11, 23, 32, 33]. Among various image encryption algorithm based on hyper-chaos, it has been shown that some scheme can be effectively broken with known plaintext and chosen plaintext attacks, and people have given detailed both mathematical analysis and experimental results to testify the security weakness and potential risk of suffering statistical attacks [14, 17, 30].

In application of random grid based secret sharing, Ahmed A. proposed a novel secret image sharing scheme [1]. The scheme combines random grids (RG), error diffusion (ED) and chaotic permutation, it has the advantages of simple computation, alternative order of shadow images in recovery, avoids the design of complex codebook, and avoids the pixel expansion problem.

3 The Proposed Scheme

In this section, we give detailed description of the process of sharing secret generation and restoration of original secret image, and for simplicity,we only discuss the algorithm for grey image. The general flowchar of the scheme can be described in Figure 3.

3.1 The Generation of Sharing Secret

Step 1: For secret images $I_1, I_1, \cdots, I_m$, firstly, the message authentication code (MAC) with the size of 256 bit for every image is calculated, thus m message authentication codes (MAC) $H_1, H_2, \cdots, H_m$ are obtained.

Step 2: For every $H_i, i = 1, 2, \cdots, m$, assume its 256 bits are expressed by $h_1, h_2, \cdots, h_{256}$, then it is truncated into 64 bit by Equation (2).

$$\begin{cases} h_i^{'} = h_i \otimes h_{i+128}, i = 1, 2, \cdots, 128, \\ h_i^{''} = h_i^{'} \otimes h_{i+64}^{'}, i = 1, 2, \cdots, 64, \end{cases} \tag{2}$$

Step 3: For the produced 64 bit data, it is divided into 4 sections, every section includes 16 bits, apply Equation (3) to turn the 16 bits data into a integer which belongs to [0, 65535]. Thus, we can get 4 integer numbers.

$$\begin{cases} x_1 = Bin2dec(h_1^{''} h_2^{''}, \cdots, h_{16}^{''}) \\ x_2 = Bin2dec(h_{17}^{''} h_{18}^{''}, \cdots, h_{32}^{''}) \\ x_3 = Bin2dec(h_{33}^{''} h_{34}^{''}, \cdots, h_{48}^{''}) \\ x_4 = Bin2dec(h_{49}^{''} h_{50}^{''}, \cdots, h_{64}^{''}) \end{cases} \tag{3}$$

Step 4: Multiply the above generated 4 number by 10^{-5}, and give it to four initial values $x_1(0), x_2(0), x_3(0), x_4(0)$.

$$\begin{cases} x_1(0) = x_1 \times 10^{-5} \\ x_2(0) = x_2 \times 10^{-5} \\ x_3(0) = x_3 \times 10^{-5} \\ x_4(0) = x_4 \times 10^{-5} \end{cases} \tag{4}$$

Step 5: Implement encryption of secrets image $I_1, I_1, \cdots, I_m$ based on hyper-chaos using the method proposed by Gao [11], that is to say, for every image I_i, the following steps are done.

1) Iterate the hyper-chaotic system for N_0 times by Runge-Kutta algorithm to avoid the harmful effect of transient procedure.

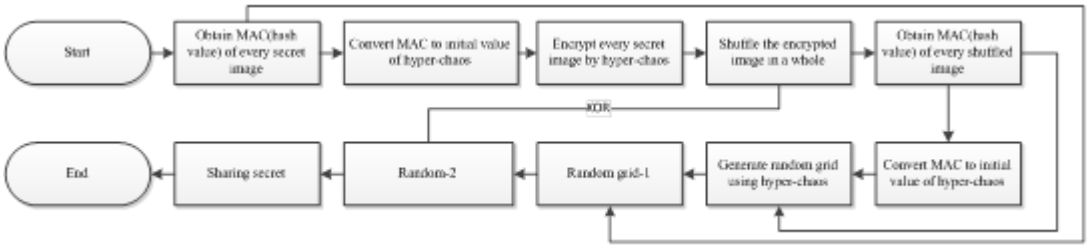


Figure 3: The general flowchart of the proposed scheme

2) The hyper-chaotic System (1) is iterated, and as a result, four decimal fractions x_1, x_2, x_3, x_4 will be generated. These decimal values are preprocessed firstly as follows

$$x_i = mod((Abs(x_i) - floor(Abs(x_i)) \times 10^{14}, 256) \tag{5}$$

where$Abs(x_i)$ returns the absolute value of x_i.$Floor(x)$ returns the value of x to the nearest integers less than or equal to x,$mod(x, y)$ returns the remainder after division.

3) Encrypt the image by formula according to Equation (6):

$$\begin{cases} C_{3\times(i-1)+1} = P_{3\times(i-1)+1} \otimes x_1 \\ C_{3\times(i-1)+2} = P_{3\times(i-1)+2} \otimes x_2 \\ C_{3\times(i-1)+3} = P_{3\times(i-1)+3} \otimes x_3 \\ C_{3\times(i-1)+4} = P_{3\times(i-1)+4} \otimes x_4 \end{cases} \tag{6}$$

where $i = 1, 2, \cdots$ represents the i^{th} iteration of the hyper-chaotic system. The symbol$\otimes$ represents the exclusive OR operation bit-by-bit.$P_i, i = 1, 2, \cdots, N \times M$ represents pixel values of the secret image, The process does not end until the set is all encrypted. Then the encrypted pixel set $C_i, i = 1, 2, \cdots, N \times M$ is written to the cipher-image.

When all the secret image are encrypted, m encrypted image $I_i^{'}, i = 1, 2, \cdots, m$ are gotten. Obviously, as the initial values of hyper-chaos are different, so it means that the encryption algorithm is One-Time Pad, which is information-theoretically secure in that the encrypted message provides no information about the original message to a cryptanalyst (except the length of the message).

Step 6: Put the images $I_i^{'}, i = 1, 2, \cdots, m$ in a row, so a matrix with the size M row and $N \times m$ column are presented. For convenience, it is assumed that every image is with the size of $M \times N$.

Step 7: Shuffle the encrypted image. Take one column from every image in turn until column are obtained, these columns are combined into a image $I_i^{''}$. Then in the same way, take the remainder columns in order from the m encrypted images, every column is combined into one image. Finally, m confused images $I_i^{''}, i = 1, 2, \cdots, m$ are generated.

Step 8: For every confused image of $I_i^{''}, i = 1, 2, \cdots, m$, repeat Step 1, then another m message authentication code (MAC) $H_i^{'}, i = 1, 2, \cdots, m$ which belongs to $I_i^{''}, i = 1, 2, \cdots, m$ are obtained.

Step 9: For every MAC $H_i^{'}, i = 1, 2, \cdots, m$, repeat Step 2 to Step 4, then four initial values of hyper-chaos are generated.

Step 10: For above four initial values $x_1(0)$, $x_2(0)$, $x_3(0)$, $x_4(0)$. Iterate the hyper-chaotic system by Runge-Kutta algorithm with initial parameters, the step of progression is 0.01, and the number of iterations is $M \times (N+1)$. Then we will get four decimal fractions in each iteration, and then preprocesses the four decimal fractions as follows

$$x_i^* = mod((Abs(x_i) - floor(Abs(x_i)) \times 10^{14}, 256) \tag{7}$$

where $Abs(x)$ returns the absolute value of x. $Floor(x)$ returns the value of x to the nearest integers less than or equal to x, $mod(x, y)$returns the remainder after division. As a result, x_1^*, x_2^*, x_3^* and x_4^*, which belong to will be used in generation of RG.

Step 11: Arrange the resulted x_1^* in M rows and $N+1$ columns. As a result, a random grid generated by hyper-chaos will be gotten.

Step 12: Repeat Steps 9 to 11 for every $I_i^{''}, i = 1, 2, \cdots, m$, m random grids will be obtained. They are called $R_i, i = 1, 2, \cdots, m$, with the size of $M \times (N+1)$.

Step 13: Generation of sharing secret. For confused image $I_1^{''}$. Firstly, convert the MAC H_1into 32 decimal values labeled by $a_i^1, i = 1, 2, \cdots, 32$. In the same way, $H_i^{'}, i = 1, 2, \cdots, m$ are also converted into

$b_i^1, i = 1, 2, \cdots, 32.$

$$\begin{cases} a_1^1 = Bin2dec(h_1^1 h_2^1, \cdots, h_8^1) \\ a_2^1 = Bin2dec(h_9^1 h_{10}^1, \cdots, h_{16}^1) \\ \cdots\cdots \\ a_{32}^1 = Bin2dec(h_{249}^1 h_{250}^1, \cdots, h_{256}^1) \\ b_1^1 = Bin2dec(h_1^{I'} h_2^{I'}, \cdots, h_8^{I'}) \\ b_2^1 = Bin2dec(h_9^{I'} h_{10}^{I'}, \cdots, h_{16}^{I'}) \\ \cdots\cdots \\ b_{32}^1 = Bin2dec(h_{249}^{I'} h_{250}^{I'}, \cdots, h_{56}^{I'}) \end{cases} \quad (8)$$

where as, $h_i^1, i = 1, 2, \cdots, 256$ represents 256 bits of H_1, and $h_i^{I'}, i = 1, 2, \cdots, 256$ is the 256 bits of H_1'. Obviously, $a_i^1, b_i^1 i = 1, 2, \cdots, 32$ all are in the scope of $[0, 255]$. Then, put them into the last position of the last column of R_1. Lastly, carry out the exclusive or operation between the ahead N columns of R_1 and I_1'', thus get the shared images E_1.

In the same way, for other confused images $I_i'', i = 2, 3, \cdots, m$, the sharing secret $E_i'', i = 2, 3, \cdots, m$ can be derived. all are with the size of $M \times (N + 1)$.

The detail flowchart of the generation of sharing secret is shown in Figure 4.

3.2 Process of Secret Image Restoration

For every shared image $E_i, i = 1, 2, \cdots, m$, the two groups MAC can be collected from the last column, then use the same way as that in the generation stage of sharing image, we can obtain random grid $R_i', i = 1, 2, \cdots, m$, then, the following steps are executed in order to restore the secret images.

Step 1: Implement the exclusive or operation between the ahead N columns of generated R_i' and $E_i, i = 1, 2, \cdots, m$, then, image $U_i'', i = 1, 2, \cdots, m$ with the size of $M \times N$ are produced.

Step 2: Transform $U_i'', i = 1, 2, \cdots, m$ into the $U_i', i = 1, 2, \cdots, m$ in the reverse order with the Step 7 of generation stage of sharing.

Step 3: Use the MAC to execute Steps 2-4 in the generation stage of sharing secret to perform decryption for every $U_i', i = 1, 2, \cdots, m$, then the original secret images $I_i, i = 1, 2, \cdots, m$ are restored.

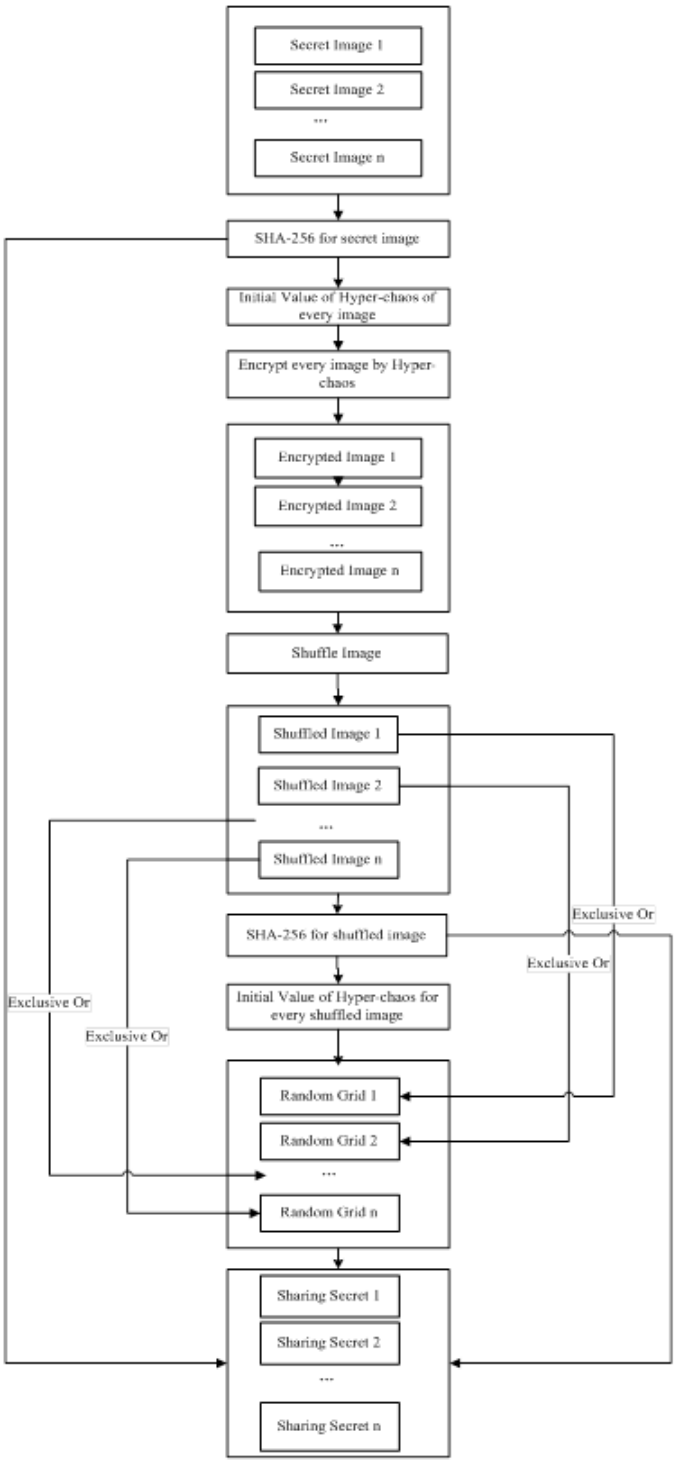


Figure 4: The flowchart of the generation of sharing secret

4 Experimental Results and Discussions

The experiment was done by Mathworks MATLAB version 12b. Some grey images such as "Lenna, Bird, Aerial and Camera" with the size of 256×256 are used for secret image; they are shown in Figure 5.

4.1 Experimental Results

Firstly, the secret images are calculated by MAC, and then MAC generated by the secret images is given in Table 1, respectively.

The initial values of hyper-chaos converted from the

Table 1: The MAC of original secret image

Secret Image	**MAC**
Lenna	CD77462213222E005C31767595E33417FB1C78DB8570837C769E8A7AB4E754F8
Bird	A0C7B83460A68AE1DB97FFD9F98E0CB993FFAAACAB595012DA784833365D65F7
Aerial	05B44DC1F9BEE57B656651A62E2A7389D8774AF93FF133A7E9C11D0AF8F67814
Camera	CB629BACC15E52D202567C47421D0368B92FE25B475FFABFFE9FBEB1C9415A17

above MAC are as follows:

$$\begin{cases} x_1(0) = 0.07364 \\ x_2(0) = 0.49910 \\ x_3(0) = 0.46934 \\ x_4(0) = 0.52627 \end{cases}$$

$$\begin{cases} x_1(0) = 0.13105 \\ x_2(0) = 0.42354 \\ x_3(0) = 0.01068 \\ x_4(0) = 0.46013 \end{cases}$$

$$\begin{cases} x_1(0) = 0.14164 \\ x_2(0) = 0.15432 \\ x_3(0) = 0.01257 \\ x_4(0) = 0.49200 \end{cases}$$

$$\begin{cases} x_1(0) = 0.36484 \\ x_2(0) = 0.47873 \\ x_3(0) = 0.03421 \\ x_4(0) = 0.61714 \end{cases} \quad (9)$$

So, the encrypted image and the final sharing secret image can be got and are shown in Figure 6. Here the size of encrypted image is with 256×256, and the sharing secret image is with the size of 256×257.

4.2 Security of the Proposed Scheme

In the proposed scheme, the security of algorithm lies in two aspects. One is that the hash values of original secret images are used to generate the initial values of the hyper-chaos system and initial values decides the security of encryption. It can be obviously seen from the generation processing of sharing secret, the initial values of hyper-chaos (secret keys) are strongly related to secret images. In the third stage of the scheme, the MAC of shuffled image is used to generate another group initial value of hyper-chaos, which affects the generation of final sharing secret. This is a combination of two one-time pads from the point of encryption, and the two MAC codes all have 512 bits, this makes the secret space reaches, which ensures the security of the scheme.

(a) Lenna (b) Bird

(c) Aerial (d) Camera

Figure 5: The image for test

Another aspects lies in the generation of sharing secret based on hyper-chaos diffusion. One has shown that image encryption algorithm based on hyper-chaotic system with only one round diffusion process has some kind of weak security [32]. The proposed scheme achieves the security through two rounds of diffusion process and one position confusion.

From the description of the proposed scheme, it can be seen that, the sharing secret can be used to restore the original secret image, but the scheme need all the sharing secret images to take part in. From the point of security, once anyone of the sharing secret images is modified, the original image will not be restored. The data position in the sharing secret image is shown in Figure 7.

Obviously, the some original secret images will not be completely restored if the "Sharing image data" is tampered with from the scheme. Even if only one bit are modified, some original secret images will not restored completely.

For example, the value of the 30^{th} position in Sharing image data in sharing secret 1 is 45. If it is modified to be 46, other data in all sharing secrets are all kept intact,

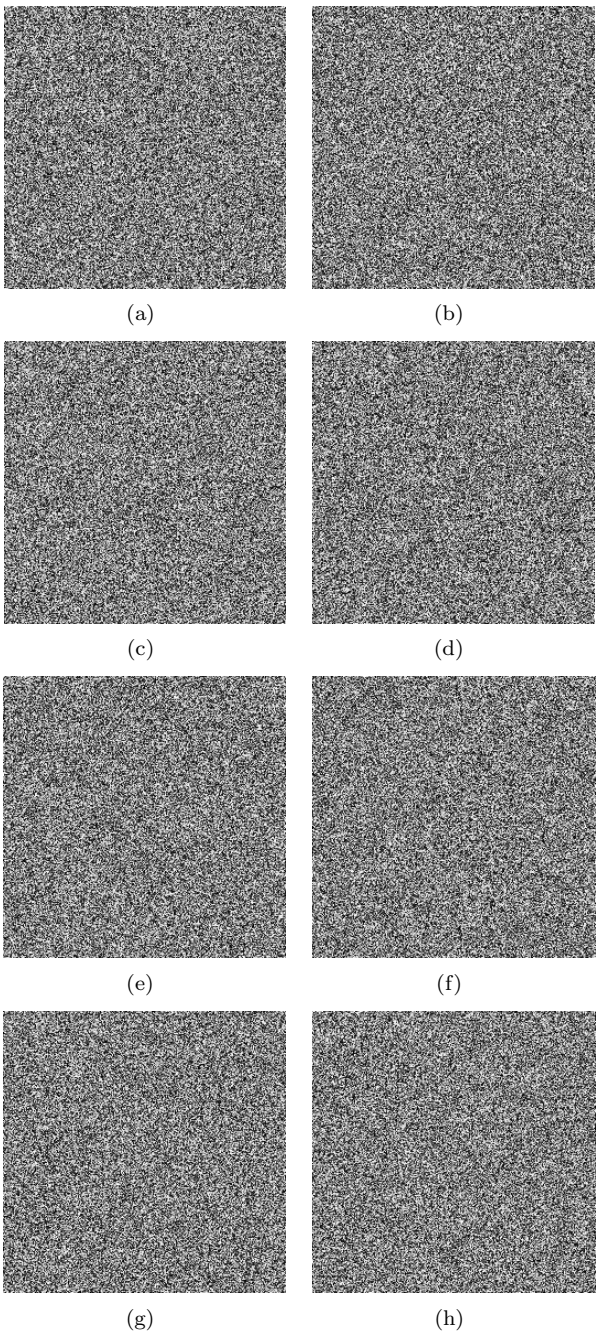


Figure 6: The encrypted and sharing secret images, (a)-(d) the encrypted image (e)-(h) the sharing secret

then it is found that the image "Aerial" is not lossless restored when we restore the original secret image, as the MAC of restored "Aerial" is

"1C3132BE45A60A4E544F39E29803B03A477C6
C4B2D9F0557AD217C6168BD3DA0"

It is totally different from that in table 1. Although the image has no distinct changes from the visual effect, it is indeed altered from the computation, so it is not lossless restore for the algorithm in real sense.

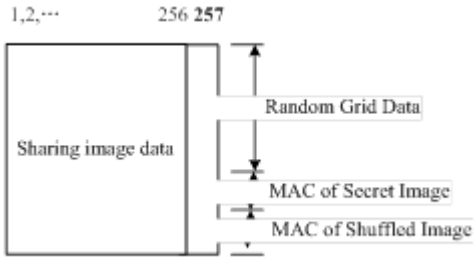


Figure 7: Data position information in sharing secret

"MAC of Shuffled Image" affects the generation of random grid, if it is falsified, the sharing secret will become useless, because if it is modified, the initial values of hyper-chaos will be altered, thus the random grid generated by hyper-chaos will be different from the original one, this will result in the mistakenly restoration of original secret image. In the same sense, "MAC of Secret Image" affects the decryption of original secret image, if it is falsified, the original secret image will not be correctly decrypted. "Random Grid Data" in the sharing secret is redundant; it can be replaced by some valuables, which will be studied in future works.

For example, for sharing secret 1, the data of "MAC of Secret Image" is "205, 119, 70, 34, 19, 34, 46, 0, 92, 49, 118, 117, 149, 227, 52, 23, 251, 28, 120, 219, 133, 112, 131, 124, 118, 158, 138, 122, 180, 231, 84, 248", if the first number "205" is modified to be "206", other data in sharing secret 1 and other sharing secret are all not modified. In this situation, the four shuffled images are all correctly restored, and when restore the original secret image, the initial values of hyper-chaos for the first group become:

$$\begin{cases} x_1(0) = 0.08132 \\ x_2(0) = 0.49910 \\ x_3(0) = 0.46934 \\ x_4(0) = 0.52627 \end{cases} \tag{10}$$

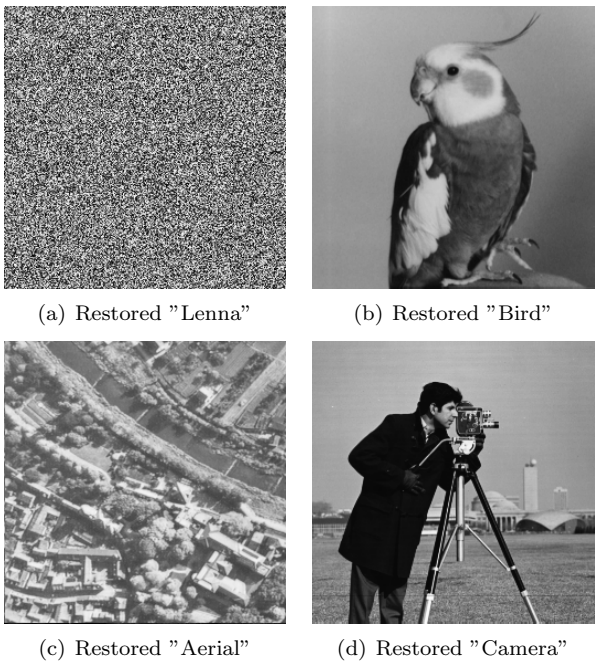
It is obviously different from the original initial value in Equation (7), so we get the restored secret image, which are shown in Figure 8.

4.3 Double Verification

It can be seen from the generation of sharing secret, the proposed scheme has verification of two stages. One is the integrity verification of sharing secret image; the other is the integrity verification of restored original secret image.

For anyone of sharing images, it is with the size of 256×257, in order to verify the integrity of sharing secret image; we can perform the following steps:

1) Convert the "MAC of Shuffled Image" into initial value of hyper-chaos, and generate the random grid R.



(a) Restored "Lenna" (b) Restored "Bird"

(c) Restored "Aerial" (d) Restored "Camera"

Figure 8: The recovered secret images when sharing secret is modified

2) Carry out the exclusive or operation between R and the sharing secret image, and the resulting image is represented by $I^{''}$.

3) Calculate the MAC of $I^{''}$, and compared it with "MAC of Shuffled Image", if they are identical, it proves the "MAC of Shuffled Image" is intact; else we can think the "MAC of Shuffled Image" is modified.

Of course, if the "Sharing image data" are modified, then the MAC of $I^{''}$ and that of "MAC of Shuffled Image" must be different from the properties of MAC. Similarly, we can also judge whether the Sharing secret is tampered with.

For the second verification stage, assume that the restored original image is I^s, then, calculate the MAC of the restored image, and compare it with that of the "MAC of Secret Image", thus we can also judge the integrity of the restored secret image.

The necessity of the second verification can be explained by the following example.

Assume the four sharing secret image are $I_i^r, i = 1, 2, 3, 4$, and the 20th column data in is tampered with, and other any data in are all not modified.

From the processing of sharing image generation and decryption, it can be concluded that

1) The first sharing secret image is modified through the verification of "MAC of Shuffled Image", and the other three sharing secret images are intact.

2) For the four shuffled images, only the first shuffled image is modified.

3) Convert to original secret image from the shuffled image, only the fourth image is altered, and the other three are lossless.

To clearly explain the verification process, two examples are given in the following. The example1 illustrates the verification of sharing secret, another one is for verification of restored secret image.

Example 1. *For sharing image 2, the 32 decimal values corresponding to "MAC of Shuffled Image" are "231, 30, 203, 231, 35, 111, 191, 127, 186, 197, 47, 240, 131, 114, 102, 156, 1, 136, 108, 248, 15, 215, 160, 183, 15, 170, 183, 44, 127, 237, 110, 231". If the last two numbers are modified into "111, 230" from "110 230", then, the original initial value of hyper-chaos and that of modified one will become the following, respectively:*

$$\begin{cases} x_1(0) = 0.21497 \\ x_2(0) = 0.16323 \\ x_3(0) = 0.53287 \\ x_4(0) = 0.06067 \end{cases}$$

$$\begin{cases} x_1(0) = 0.21497 \\ x_2(0) = 0.16323 \\ x_3(0) = 0.53287 \\ x_4(0) = 0.05810 \end{cases} \tag{11}$$

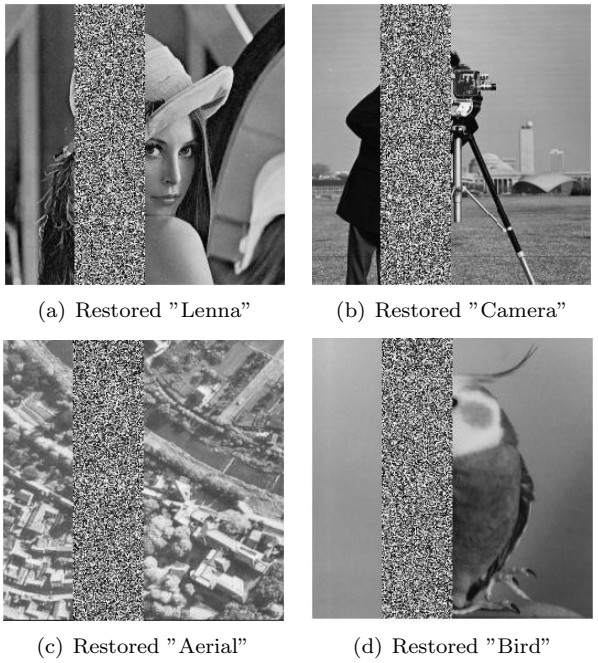
Thus, the random grid generated by hyper-chaos with the initial values of modified one is different from that generated by original initial values. In this case, the restored secret images are shown in Figure 9. it can also be verified that the restored images are damaging.

Example 2. *In order to testify the necessity of integrity verification of the restored secret image, it is assumed that the first sharing image is inserted by two white lines, such as shown in Figure 10. Then, the decryption process is used to retrieve the secret images; the restored secret images are shown in Figure 10. It can be concluded that from the verification process that, the image "Lenna" and "Camera" are lossless restored, because the MAC of the two restored image and that in the sharing images are the same, but the MAC of restored "Bird" and "Aerial " are:*

"$E3D53BB3C4324FA6E023F4CDEE4A7F4CD956$
$CBB36E9523403B83C436049552D4$"

"$F8F7944EFE837387F9C56C0CB8EEA19077337$
$09115D0FDAA23C6505B8D31ADC0$"

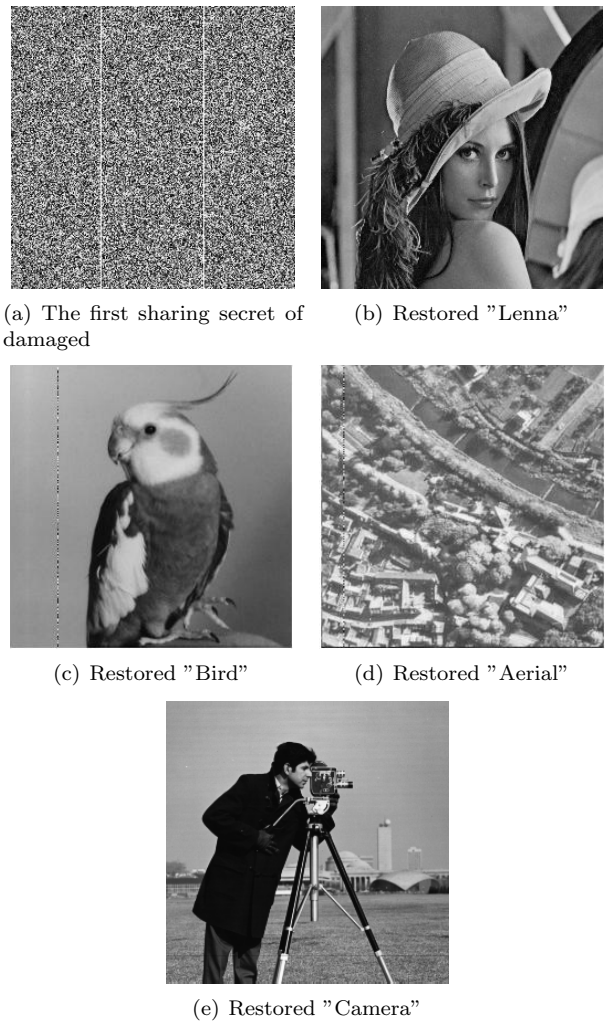
Obviously, they are different from the MAC of secret images (Table 1), so, it is regarded that the image "Bird" and "Aerial" are damaging restored.



(a) Restored "Lenna" (b) Restored "Camera"

(c) Restored "Aerial" (d) Restored "Bird"

Figure 9: Restored secret image when the second sharing secret is modified



(a) The first sharing secret of damaged (b) Restored "Lenna"

(c) Restored "Bird" (d) Restored "Aerial"

(e) Restored "Camera"

Figure 10: Restored secret image when the first sharing secret is modified

4.4 Comparison with Existing Algorithms

It can be seen from the description of the scheme, the final sharing secret is with the size of $M \times (N+1)$, the pixel expansion is $1+\frac{1}{NM}$. Obviously, the pixel expansion is mini with the larger M and N.

Table 2 gives some comparisons with existing algorithms in application type of images, size of sharing secret, property of restored secret image and verification of the scheme.

Apparently, the proposed scheme has the following advantages when it is compared with some published literatures.

1) The generated sharing secrets have only mini expansions than original secret image, if the original secret image is with the size of $M \times N$, then the expansion is only $\frac{1}{NM}$ compared with the original secret image.

2) The sharing secret is generated by One-Time Pad, which is information-theoretically secure in that the encrypted message provides no information about the original message to a cryptanalyst (except the length of the message).

3) The proposed algorithm can not only verify the integrity of the sharing secret, it can also verify the integrity of the restored secret image. If the sharing image is not modified, the restored stored image is completely the same as the original one, parts of the sharing images are modified may not means that the original secret images are lossless restored or not, it can be testified through the verification process.

4) Every right sharing secret can help to restore one part of original secret image, the more sharing secret takes part in, the better that original secret image restore, if all the sharing secret are used, the secret image can be restored completely. So the proposed scheme can progressively recover the original secret image.

5 Conclusions

A novel multi secret sharing scheme is proposed in this paper. Generation of sharing secret includes three stages in the scheme. The first stage is encryption of secret image based on hyper-chaos, the encryption algorithm rely on the secret image self. The second stage is shuffling of encryption. The third stage is generation of sharing secret. In this stage, MAC of shuffled encryption image are used for initial values of hyper-chaos, then the hyper-chaos is used to generate random grid, finally, the sharing secret is generated through XOR operation between random grid

Table 2: Some comparisons with existing scheme

The scheme	Type of secret image	Pixel expansion	Recovered image	Type of VSS	Verification
Shyu's method (2009)	B, G, C	N	Lossy	RG-based	No
Chen's method (2011)	B	N	Lossy	RG-based	No
Wu's method (2012)	B	N	Lossy	RG-based	No
Chen's method (2013)	G	N	Lossless	RG-based	No
Lin's method (2015)	B, G, C	N	Lossy	RG-based	No
Ours method	G	Mini	Lossless	RG-based	Yes

and confused image. Large numbers of experiments show that the proposed scheme can lossless restore the original secret image, and have the double verification ability for sharing secret and restored secret image. The above advantages make it especially suitable for secret sharing of important images such as medical and military images.

Acknowledgments

The work was partly supported by the Program of National Science Fund of Tianjin, China (Grant NO. 16JCYBJC15700). The authors gratefully acknowledge the anonymous reviewers for their valuable comments.

References

[1] A. A. AbdEl-Latif, X. H. Yan, L. Li, et al., "A new meaningful secret sharing scheme based on random grids, error diffusion and chaotic encryption," *Optics and Laser Technology*, vol. 54, pp. 389–400, 2013.

[2] G. R. Blakley, "Safe guarding cryptographic keys," in *Proceedings of the National Computer Conference*, pp. 313–317, 1979.

[3] T. H. Chen, K. H. Tsao, "Threshold visual secret sharing by random grids," *Journal of Systems and Software*, vol. 84, pp. 1197–1208, 2011.

[4] T. H. Chen, K. H. Tsao, "User-friendly random-grid-based visual secret sharing," *IEEE Transactions on Circuit Syst. VideoTech*, vol. 21, no. 11, pp. 1693–1703, 2011.

[5] W. K. Chen, "Image sharing method for gray-level images," *The Journal of Systems and Software*, vol. 86, pp. 581–585, 2013.

[6] Y. C. Chen, D. S. Tsai, R. B. Horng, "Visual secret sharing with cheating prevention revisited," *Digital Signal Processing*, vol. 23, pp. 1496–1504, 2013.

[7] Y. C. Chen, D. S. Tsai, R. B. Horng, "A new authentication based cheating prevention scheme in Naor-Shamir's visual cryptography," *Journal of Visual Communication and Image Representation*, vol. 23, no. 8, pp. 1225–1233, 2012.

[8] Y. C. Chen, D. S. Tsai, R. B. Horng, "Visual secret sharing with cheating prevention revisited," *Digital Signal Process*, vol. 23, pp. 1496–1504, 2013.

[9] W. P. Fang, J. C. Lin, "Progressive viewing and sharing of sensitive images," *Pattern Recognition and Image Analysis*, vol. 16, no. 4, pp. 632–636, 2006.

[10] T. Gao, Z. Chen, et al. "A hyper-chaos generated from Chen's system," *International Journal of Modern Physics C*, vol. 17, pp. 471–478, 2006.

[11] T. Gao, Z. Chen, "A new image encryption algorithm based on hyper-chaos," *Physics Letters A*, vol. 372, no. 4, pp. 394–400, 2008.

[12] G. Horng, T. H. Chen, D. S. Tsai, "Cheating in visual cryptography," *Designs, Codes and Cryptography*, vol. 38, no. 2, pp. 219–236, 2006.

[13] Y. C. Hou, Z. Y. Quan, C. F. Tsai, et al., "Block-based progressive visual secret sharing," *Information Sciences*, vol. 233, no. 1, pp. 290–304, 2013.

[14] F. J. Jeng, W. L. Huang, T. H. Chen, "Cryptanalysis and improvement of two hyper-chaos-based image encryption schemes," *Signal Processing: Image Communication*, vol. 34, pp. 45–51, 2015.

[15] O. Kafri, E. Keren, "Encryption of pictures and shapes by random grids," *Optics Letters*, vol. 12, no. 6, pp. 377–379, 1987.

[16] Y. S. Lee, T. H. Chen, "Insight into collusion attacks in random-grid-based visual secret sharing," *Signal Processing*, vol. 92, pp. 727–736, 2013.

[17] C. Q. Li, Y. S. Liu, T. Xie, et al., "Breaking a novel image encryption scheme based on improved hyperchaotic sequences," *Nonlinear Dynamics*, vol. 73, pp. 2083–2089, 2013.

[18] C. H. Lin, T. H. Chen, Y. T. Wu, et al, "Multi-factor cheating prevention in visual secret sharing by hybrid codebooks," *Journal of Visual Communication and Image Representation*, vol. 25, pp. 1453–1557, 2014.

[19] P. Y. Lin, R. Z. Wang, Y. J. Chang, et al., "Prevention of cheating in visual cryptography by using coherent patterns," *Information Sciences*, vol. 301, pp. 61–74, 2015.

[20] Y. J. Liu, C. C. Chang, "An integratable verifiable secret sharing mechanism," *International Journal of Network Security*, vol. 18, no. 4, pp. 617–624, 2016.

[21] M. Naor, A. Shamir, "Visual cryptography," in *Advances in cryptography (Eurocrypt'95)*, pp. 1–12, 1995.
[22] NIST, *Announcing The Secure Hash Standard*, Federal Information Processing Standards Publication 180-2, U.S. DoC/NIST, Aug. 2002.
[23] B. Norouzi, S. Mirzakuchaki, S. M. Seyedzadeh, et al., "A simple, sensitive and secure image encryption algorithm based on hyper-chaotic system with only one round diffusion," *Multimedia Tools and Applications*, vol. 71, no. 3, pp. 1469–1497, 2014.
[24] A. Shamir, "How to share a secret," *Communications of the ACM* , pp. 612–613, 1979.
[25] S. J. Shyu, "Image encryption by multiple random grids," *Pattern Recognition*, vol. 42, pp. 1582–1596, 2009.
[26] D. S. Tsai, T. H. Chen, G. Horng, "On generating meaningful shares in visual secret sharing scheme," *Image Science Journal*, vol. 56, pp. 49–55, 2008.
[27] X. T. Wu, W. Sun, "Random grid-based visual secret sharing for general access structures with cheat-preventing ability," *The Journal of Systems and Software*, vol. 85, pp. 1119–1134, 2012.
[28] X. Yan, S. Wang, X. Niu, et al., "Generalized random grids-based threshold visual cryptography with meaningful shares," *Signal Processing*, vol. 109, pp. 317–333, 2015.
[29] S. Yanchuk, T. Kapitaniak, "Symmetry-increasing bifurcation as a predictor of a chaos-hyperchaos transition in coupled systems," *Physical Review E*, vol. 64, pp. 056235, 2001.
[30] Y. S. Zhang, D. Xiao, W. Y. Wen, et al., "Breaking an image encryption algorithm based on hyper-chaotic system with only one round diffusion process," *Nonlinear Dynamics*, vol. 76, no. 3, pp. 1645–1650, 2014.
[31] Z. Zhou, G. R. Arce, G. D. Crescenzo, "Halftone visual cryptography," *IEEE Transactions on Image Process*, vol. 15, no. 8, pp. 2441–2453, 2006.
[32] C. Zhu, "A novel image encryption scheme based on improved hyperchaotic sequences," *Optical Commun*, vol. 285, no. 1, pp. 29–37, 2012.
[33] H. Zhu, Y. Zhang, 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.
[34] X. Zhuang, C. C. Chang, Z. H. Wang, et al., "Simple password authentication scheme based on geometric hashing function," *International Journal of Network Security*, vol. 16, no. 4, pp. 271–277, 2014.

Biography

Hang Gao was born in Tianjin City, China, in 1992. He received the B. S. degree in Software Engineering from University of Electronics Science and Technology of China, Chengdu, China, in 2015. He is currently working toward the M. S. degree in Software Engineering from Nankai University, Tianjin, China. His research interests include information security and cloud computing.

Mengting Hu was born in Shanxi Province, China, in 1993. She received the B. S. degree in Software Engineering from Tongji University, Shanghai, China, in 2015. She is currently working toward the M. S. degree in Software Engineering from Nankai University, Tianjin, China. Her research interests include cloud computing and Information Retrieval.

Tiegang Gao received Ph. D degree from Nankai University, Tianjin, China in 2005. He is a professor in college of software, Nankai University, China since 2006. His research interests include cloud computing and information security, he has published or co-authored more than 100 papers in related field.

Renhong Cheng received Ph. D degree from Nankai University, Tianjin, China. Now He is a professor in college of computer and control engineering, Nankai University, His research interests include database technique and Information Retrieval.

Coverless Text Information Hiding Method Using the Frequent Words Hash

Jianjun Zhang[1,3], Huajun Huang[2], Lucai Wang[3], Haijun Lin[3], Deng Gao[4]
(Corresponding author: Jianjun Zhang)
College of Computer Science and Electronic Engineering, Hunan University[1]
36 Lushan Rd, Yuelu Qu, Changsha Shi, Hunan 410080, China
(Email: jianjun998@163.com)
College of Computer and Information Engineering, Central South University of Forestry and Technology[2]
College of Engineering and Design, Hunan Normal University[3]
College of Software, Hunan Vocational College of Science and Technology[4]
(Received Nov. 22, 2016; revised and accepted Feb. 20 & Mar. 31, 2017)

Abstract

The attackers may discover the existence of the secret information or even get it by analyzing the cover's statistical characteristics, changes of which often occur due to the embedding. In this paper, a novel coverless text information hiding method was proposed. By using the words rank map and the frequent words hash, normal texts containing the secret information could be retrieved from the text database, and will be sent to the receiver without any modification. Because the embedding is not needed, the proposed method could be able to escape from almost all state-of-the-art steganalysis methods.

Keywords: Big Data; Coverless Information Hiding; Frequent Words Hash; Rank Map; Steganography

1 Introduction

Steganography, also known as information hiding, is a secure communication method that conveys secret messages in the form of plaintexts so that the appearances of the secret messages will not draw eavesdroppers' attention while they are being transmitted through an open channel [17]. It can be used for intellectual property protection and secret communications [9]. For example, reference [20, 25] introduced two methods of detecting illegal copies of copyrighted images. For information hiding, there are many kinds of covers, such as texts [15], images [7, 18], videos [16], etc. [1, 12, 21].

Compared to the image or other covers, text information hiding is the most difficult kind of steganography due to the lack of redundancy. Because the text is frequently used in people's daily lives, however, text information hiding has attracted many researchers' interest, and has many results [23]. Classified by the covers, text steganography could be put into three types: text format-based [3, 10, 11], generating-based and embedding-based natural language information hiding. For text format-based information hiding, the embedded information will no longer exist if the document is generated without format after extracting the text content. Generating-based natural language information hiding methods can fool the computer statistical analysis, but is relatively easy to be identified by people [2]. Embedding-based natural language information hiding methods have more robust and better concealment than text format-based information hiding, but the hiding algorithm is difficult to implement, and there are some deviations and distortions in the statistic and linguistics because of the limitation of the natural language processing [13].

Once the information hiding algorithm is public, the steganalysis methods will be appearing. So attackers will know the existence of the secret information by analyzing changes of statistical characteristics of the covers caused by the embedded information. Be there an algorithm with which the secret information could be hidden without any modification of the covers. Coverless information hiding [24, 6], firstly proposed by Xingming Sun et al., is the best answer to the above question. Reference [24] presented a coverless image steganography framework, and Reference [6] proposed a coverless text information hiding method. These two methods can directly retrieve the stego-image (stego-text) without any modification of the covers. Recently, coverless information hiding, which requires no modification on covers and could resist various steganalysis technologies, draws more and more attention from researchers [5, 22].

In this paper, a novel coverless text information hiding method is proposed. Firstly, a text database is constructed by collecting a large number of texts from the Internet. Then the word rank maps of the words will be calculated by statistically analyzing the text big data,

meanwhile, the frequent words distance of every text is calculated. When a certain information will be transmitted, a normal text containing the secret information is retrieved from the text database by using the frequent words distance and the word rank maps, and sent to the receiver without any modification.

2 Coverless Information Hiding

Coverless information hiding is a new challenging research field. In fact, "coverless" is not to say that there is no carrier, but compared with the conventional information hiding, coverless information hiding requires no other carries [6]. The idea of coverless information hiding is often used in our daily life, and the acrostic poem is a classic example. An acrostic poem is shown in Figure 1 form which we can learn that the secret information is "TREE". Coverless information hiding is essentially the disclosure of secret information in the text. Its distinctive characteristic is "no embedding", that is, a carrier cannot embed secret information by modifying it [6].

Figure 1: An acrostic poem

3 The Proposed Method

3.1 Preparation of the Text Database

We construct a natural text database by fetching the news from the normal news web sites. For each word of the vocabulary, we calculate the frequency of its occurrence, and then rank the words with the descending way (Most frequent word has rank 1, next frequent word has rank 2 ...). Figure 2 shows the ranking result of words in a text database. In order to make good use of the information of the words' occurrence in a text database (or in a text), the Word Rank Map of a text database (or a text) is defined as:

$$RM = \{(w_i, f_i)|i = 1, 2, \cdots, U\} \tag{1}$$

where U is the number of unique words in a text database (or in a text), i is the rank of a word w_i, and f_i is the frequency of w_i. Figure 2 shows the word rank map of a text database. For the example in Figure 2, we can obtain

$$\begin{aligned} RM \quad = \quad & \{(the, 124020), (and, 55654), \\ & (of, 54550), (to, 52331), \cdots\} \end{aligned} \tag{2}$$

Obviously, the top frequent words are: the, and, of, to, in, a, on, for, and etc.

total number of words: 1967696
total number of unique words: 57618

rank	word	frequency
1	the	124020
2	and	55654
3	of	54550
4	to	52331
5	in	46181
6	a	41390
7	on	19241
8	for	17883
9	that	17338
10	said	17230
11	is	16193
12	china	13951
13	with	13542
14	as	11784
15	by	10478
16	at	10236
17	it	9960
18	will	9435
19	he	9199
20	from	9079
21	has	8960
22	was	8861
23	s	7703
24	be	7389
25	have	7251
26	are	7145
27	an	7011
28	chinese	6992
29	its	6119
30	his	5587

Figure 2: Part of a text database word rank map

For each text in a text database, we can obtain its word rank map defined as Equation (1). Figure 3 shows the rank map of a text named as "2.1 million Audi cars affected by emissions cheating scandal.txt", in which there are 185 words, 113 unique words. From the rank map, we learn that the top frequent words are: the, in, software, emission, cars, etc.

For the top frequent words in the text database, we can calculate their occurrences in a text in the same collection. So, the Frequent Words Hash Function is defined as:

$$H_k(t) = \{h_1 h_2 h_3 \cdots h_k\} \tag{3}$$

where k is the number of the top frequent words chose form the vocabulary of a text database, t is a text in the text database, and h_i is defined by:

$$h_i \quad = \quad \begin{cases} 1 & \text{the i-th frequent word appears in text t} \\ 0 & \text{the i-th frequent word does not appear in text t} \end{cases} \tag{4}$$

```
total number of words: 185
total number of unique words: 113
--------------------------------------
rank    word        :    frequency:
--------------------------------------
1       the         :          19  :
2       in          :           6  :
3       software    :           4  :
4       emission    :           4  :
5       cars        :           4  :
6       car         :           4  :
7       of          :           4  :
8       on          :           4  :
9       a           :           4  :
10      audi        :           3  :
11      to          :           3  :
12      epa         :           3  :
13      cheating    :           3  :
14      million     :           3  :
15      by          :           3  :
16      said        :           2  :
17      1           :           2  :
18      total       :           2  :
19      that        :           2  :
20      were        :           2  :
21      agency      :           2  :
22      protection: :           2  :
23      emissions   :           2  :
24      environmental:          2  :
25      scandal     :           2  :
26      vehicles    :           2  :
27      s           :           2  :
28      u           :           2  :
29      volkswagen: :           2  :
30      company     :           2  :
```

Figure 3: Part of a text's word rank map

For the top 30 frequent words shown in Figure 2, we can calculate the hash value of a text named as "2.1 million Audi cars affected by emissions cheating scandal.txt". The hash value is:

$$H_k(t) = \{11111110111011111010010100100 10\} \quad (5)$$

So, we map a text into a 30 bits string. Figure 4 shows the hash values of some texts in a text database. In order to measure the occurrence of the frequent words in a text t, we define the Frequent Words Distance of a text as:

$$\begin{aligned} DFW_k(t) &= HD(H_k(t), (b_1, b_2, \cdots, b_k)) \\ b_i &= 0, \quad i = 1, 2, \cdots, k. \end{aligned} \quad (6)$$

where k is the number of the top frequent words, t is a text in the text database, and HD is Hamming Distance calculating operation.

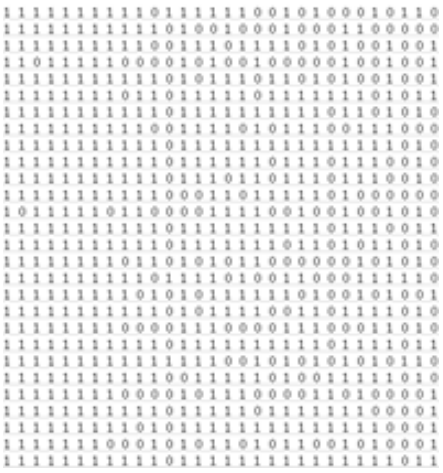


Figure 4: The hash values of some texts in a text database

By statistical analyzing the occurrence of each word in the text database, we can calculate the word rank map of each word. For each word w_i appearing in the text database, its word rank map is defined as

$$\begin{aligned} RMW_i = \{&(rw_{ij}, fw_{ij}, wt_{ij}) | i = 1, 2, \cdots, U; \\ &j = 1, 2, \cdots, N\} \end{aligned} \quad (7)$$

where rw_{ij} is the rank of the word w_i in a text wt_{ij} according to its occurrence, fw_{ij} is the frequency of w_i's occurrence in wt_{ij}, and N is the number of the texts in which w_i appears. Figure 5 shows the word rank map of "from", whose rank is 20 in the word rank map as shown in the Figure 2.

```
Word-'from'

Rank--Frequency--Source Text

46      3     'Containing China' a Japanese strateg
22      4     'Cyber Monday' sales set to hit recor
20      3     'Disturbance' only harms China, Japan Fu Yin
42      2     'Hobbit' vanished far earlier than researcher:
52      1     'In' leads by 7 percentage points ahead of Br
6       7     'Maternity tourism' in US sees chang
13      5     'Mini Messi', 5, aims to follow in star's foo
15      4     'No more survivors' in collapsed overpas
21      2     'Plane wreckage' found in Thailand fuels talk
52      3     'Slave' women held for 30 yrs rescue
41      2     'Symbol of rebirth' opens in NY
151     1     1 killed, 51 injured in Mexico candy factory
20      3     1 killed, dozens wounded at UN base in S. Sud
19      2     1 survived, 53 bodies recovered in Algeria pl
16      3     10 Afghan militants killed in army operatio
13      4     10 facts about HIVAIDS on World AIDS Da
12      2     10 killed in flash flood in eastern Indi
133     1     10 US Navy sailors detained by Iran repor
113     1     100 killed in tribal clashes in Sudan's Darfu
39      2     11 trapped miners rescued from S. African min
30      2     12 China-made helicopters delivered to Cambod
44      2     12 IS militants killed in air strike in weste
60      2     125 killed in month-long battles of Iraq's An
18      4     126 hostages rescued, 4 attackers killed as B
6       4     126 rescued, 4 attackers killed as operation
27      1     12th Elephant Festival held in Nepa
```

Figure 5: Part of the rank map of "from"

3.2 Information Hiding

The information hiding process is shown in Figure 6. Detail procedures are introduced as follows. Suppose the constructed text database is T, and the communication key is k. We can calculate the word rank map of T by using Equation (1), and get the vocabulary of T ,and let it be $W = \{w_i | i = 1, 2, \cdots, U\}$ where i is the rank of w_i, and U is the number of unique words in T. For each text t_i in T, we can calculate its word rank map by using Equation (1), and let it be RM_{t_i}.

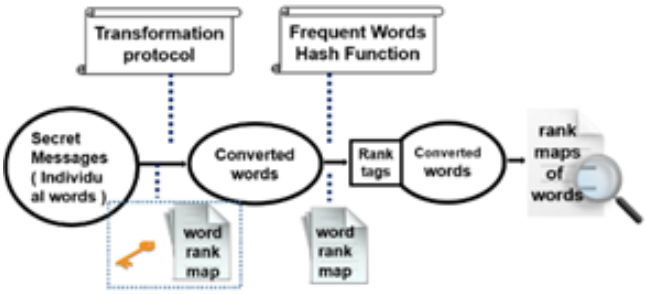


Figure 6: The process of information hiding

Because the key is k, we arrange the top frequent $k \times k$ words in W as the right part of Figure 7. Suppose the

hidden message is $M = m_1, m_2, \cdots, m_n$ where m_i is a word, and n is the number of words in the hidden message. For each word m_i in M, it is chosen from the top frequent $k \times k$ words in W. Obviously, the selection range of m_i depends on k. Therefore, both sides of communication can choose more k so that m_i has more options.

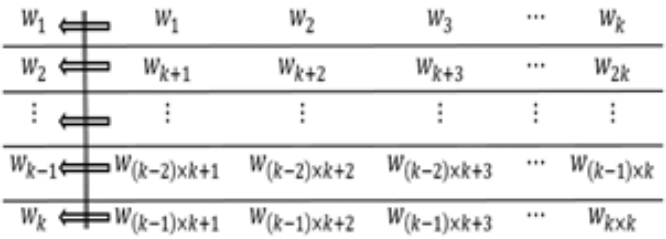


Figure 7: The words conversion table

3.2.1 Words Conversion

In order to enhance the security of the secret message, we convert each word in M into one of the top frequent k words in W before the information hiding. The conversion rule is shown in Figure 7. For each word in $\{w_1, w_2, \cdots, w_k\}$, it will be converted into w_1. For each word in $\{w_{(k+1)}, w_{(k+2)}, \cdots, w_k\}$, it will be converted into w_2. And so on.

For each word m_i in M, we can get its rank by using the word rank map of T, and let it be R_{m_i}. Then, it is located in the $((R_{m_i}-1)/k+1)$ row, $((R_{m_i}-1)\%k+1)$ column in the word conversion table shown in Figure 7, where "%" is a remainder operation. Therefore it will be converted into $m'_i = w_{((R_{m_i}-1)/k+1)}$. In this way, we can convert the secret information $M = m_1, m_2, \cdots, m_n$ into $M' = m'_1, m'_2, \cdots, m'_n$, and M' is a subset of the k top frequent words in W.

3.2.2 Searching the Stego-text

For each word $m'_i = w_{((R_{m_i}-1)/k+1)}$ in M', the stego-text is get as follows:

Firstly, for each text t in the text database, we calculate the hash value of $H_k(t)$, defined as Equation (3), where k is the communication key. Then, we can get the frequent words distance of t by using Equation (6), ant let it be $DFW_k(t)$.

Secondly, because m'_i is in the top k frequent words in the word rank map of T, we can get the rank map of m'_i by using Equation (7), and let it be $RMW_{m'_i}$.

Thirdly, by using the word rank map $RMW_{m'_i}$, we retrieve all texts containing m'_i to search a text t in which the rank of m'_i is equal to $DFW_k(t)$, and the frequency of m'_i's occurrence is equal to $((R_{m_i}-1)\%k+1)$. There may be some texts satisfying this condition, then, we can select a text from those texts as the stego-text for m'_i.

Finally, as described above, we can search a stego-text set for each m'_i in M'. These stego-texts is a normal text set that contains the converted secret message, and they can be sent to the receiver without any modifying.

3.3 Information Extraction

The process of extraction is shown in Figure 8. Suppose the stego-text is S, so S is a set of normal texts. The number of texts in S is the number of words in secret message M. Let k be the communication key. Because the text database T is open for all users, receiver can calculate the word rank map of T, and get the top k frequent words in W by using the communication key k. Certainly, receiver can get the same word conversion table shown in Figure 7. For each stego-text t in S, the details of information extraction will be introduced as follows.

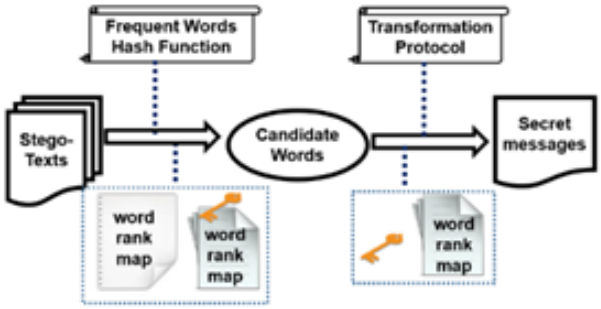


Figure 8: Secret message extracting process

3.3.1 Get the Candidate Word

Receiver can calculate the word rank map of t by using Equation (1), and the frequent words distance of t by using Equation (6) and let it be $DFW_k(t)$. By retrieving the word rank map of t, receiver can get the candidate word whose rank is equal to $DFW_k(t)$ in text t. Obviously, the candidate word is $m'_i = w_{((R_{m_i}-1)/k+1)}$.

3.3.2 Get the Secret Message

By using the word rank map of t, receiver can find the word frequency of m'_i in t, and let it be Fm'_i. Obviously, Fm'_i is equal to $((R_{m_i}-1)\%k+1)$. Receiver can find the secret message m_i that is located in the "$w_{((R_{m_i}-1)/k+1)}$" row, the $((R_{m_i}-1)\%k+1)$ column in the word conversion table shown in Figure 7.

So, receiver can get every word m_i in M. and then get the secret message $M = m_1, m_2, \cdots, m_n$.

4 Discussion

4.1 An Example

In order to clearly explain the above coverless text information hiding process, we illustrate it by a simple example. We have constructed a text database which can be expanded constantly, and it is open for all users. Suppose the communication key is 30 and the secret information M is "mutual visit". It is worth mentioning that, however, M is a subset of the top 900 frequent words in the text database. Both sides of communication may choose larger k, so that words of M have more options. The operating procedure of information hiding is introduced as follows:

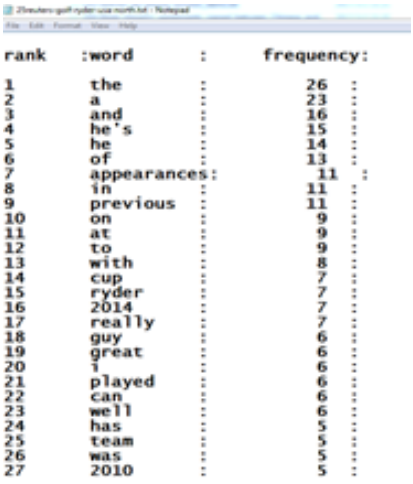
Firstly, sender computes the rank map of the text database and ones of each text in it. Because the communication key is 30, sender can obtain the top 30 frequent words set $W_{top30} = \{w_i | i = 1, 2, \cdots, 30\}$, and the W_{top30} is: $\{the, and, of, to, in, a, on, for, that, said, is, China, with, as, by, at, it, will, he, form, has, was, s, be, have, are, an, Chinese, its, his\}$.

So, sender can get the word conversion table shown in Figure 9. By retrieving the word rank map of text database, sender finds the rank of "mutual" is 605, and the rank of "visit" is 183. According to the word conversion table shown in Figure 9, therefore, "mutual" is located in 21st row, 5th column, and "visit" is located in 7th row, 3rd column. Hence, "mutual" will be converted into "has", and "visit" will be converted into "on".

the	w_1	w_2	w_3	⋯	w_{30}
and	w_{31}	w_{32}	w_{33}	⋯	w_{60}
⋮	⋮	⋮	⋮	⋮	⋮
its	w_{841}	w_{842}	w_{843}	⋯	w_{870}
his	w_{871}	w_{872}	w_{873}	⋯	w_{900}

Figure 9: The word conversion table when key is 30

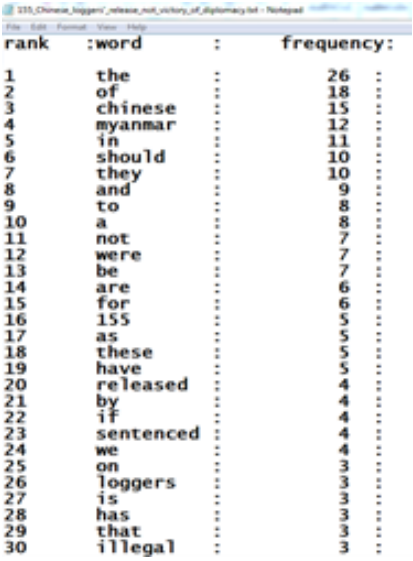
Secondly, sender calculates the word rank map of "has" and ones of "on". By retrieving the two word rank maps, sender can find a text named "25reuters-golf-ryder-usa-north.txt", and let it be t_1, and a text named "155_Chinese_loggers'_release_not_victory_of_diplomacy.txt", and let it be t_2. Their word rank maps are shown in Figure 10 and Figure 11. From Figure 10, we learn that the rank of "has" is 24 and its frequency is 5 in text t_1 whose frequent words distance is 24, and the frequent words are $\{the, a, and, of, in, on, at, to, with, has, was, is, that, as, his, for, be, an, it, by, will, he, s, have\}$ in it. From Figure 11, we learn that the rank of "on" is 25 and its frequency is 3 in text t_2 whose frequent words distance is 25, and the frequent words are $\{the, and, of, to, in, a, on, for, that, is, China, with, as, by, it, will, he, from, has, be, have, are, an, Chinese, its\}$ in it.



rank	:word	:	frequency:	
1	the	:	26	:
2	a	:	23	:
3	and	:	16	:
4	he's	:	15	:
5	he	:	14	:
6	of	:	13	:
7	appearances	:	11	:
8	in	:	11	:
9	previous	:	11	:
10	on	:	9	:
11	at	:	9	:
12	to	:	9	:
13	with	:	8	:
14	cup	:	7	:
15	ryder	:	7	:
16	2014	:	7	:
17	really	:	7	:
18	guy	:	6	:
19	great	:	6	:
20	't	:	6	:
21	played	:	6	:
22	can	:	6	:
23	well	:	6	:
24	has	:	5	:
25	team	:	5	:
26	was	:	5	:
27	2010	:	5	:

Figure 10: The word rank map of a stego-text

Finally, sender sends the two texts t_1, t_2 as the stego-texts to the receiver.



rank	:word	:	frequency:	
1	the	:	26	:
2	of	:	18	:
3	chinese	:	15	:
4	myanmar	:	12	:
5	in	:	11	:
6	should	:	10	:
7	they	:	10	:
8	and	:	9	:
9	to	:	8	:
10	a	:	8	:
11	not	:	7	:
12	were	:	7	:
13	be	:	7	:
14	are	:	6	:
15	for	:	6	:
16	155	:	5	:
17	as	:	5	:
18	these	:	5	:
19	have	:	5	:
20	released	:	4	:
21	by	:	4	:
22	if	:	4	:
23	sentenced	:	4	:
24	we	:	4	:
25	on	:	3	:
26	loggers	:	3	:
27	is	:	3	:
28	has	:	3	:
29	that	:	3	:
30	illegal	:	3	:

Figure 11: The word rank map of a stego-text

Because the text database is open to all users, receiver can calculate its word rank map, the top 30 frequent words and the word conversion table shown in Figure 9 by using the communication key $k = 30$. Then, he (or she) calculates the frequent words distance of t_1 and ones of t_2, and finds that they are 24 and 25. So, receiver retrieves the word rank maps of t_1 and t_2, and gets the candidate words "has" and "on" whose ranks are 24 and

25. From the word rank map of the stego-texts, he (or she) also learns that the candidate words' frequency are 5 and 3. Finally, receiver gets the secret message "mutual" located in "has" row, 5th column in the word conversion table shown in Figure 9, and "visit" located in "on" row, 3rd column. Hence, receiver successfully extract the secret message "mutual visit" from the stego-texts.

4.2 Security Analysis

Steganalysis is usually performed through the use of irrelevance between the embedded information and the carriers. Attackers often make steganalysis by analyzing the difference of their statistical distributions [19]. In our proposed hiding method, however, the carriers are normal pure text and the secret information is not been embedded in the carriers. The carriers can be sent to receiver without any modification. So the information hiding does not change the probability distribution of the carriers. According to the definition of the security of an information hiding system in [4], the proposed information hiding method is theoretically safe. At the same time, the proposed approach is also followed the Kerckhoffs Principle [8] in cryptography, and detail of information hiding is open. If he does not know the communication key, the attacker cannot gain any information about the hidden information [14]. Therefore, the proposed method could resist almost all kinds of current steganalysis method.

4.3 The Importance of Big Data

However, it is worth mentioning that, in order to enhance security, there are two works must be done: one is to change periodically the communication key to ensure that the secret message may be converted into different subsets of the top frequent words in the text database. The second is to establish a large text database (text big data) to increase the probability of the frequent words distance is equal to the rank of a word in a text, and so there are more choices of the stego-texts [23]. For example, in the chose text database, for the word "China", there are 3258 texts in which it appears, and their frequent words distance are shown in Figure 12. From the Figure 12, we learn that these values are not evenly distributed. There are 37 texts whose frequent words distance is 12, and there is only one text in which the rank of "China" is 12. Therefore, the text big data is necessary to ensure the smoothly implement of the proposed method.

Because the text big data is an important guarantee of the smooth implementation of the proposed method, some files should reside in the memory buffer when the big data is handling. We firstly calculate the word rank map of each text in the text big data, then the word rank map of the text database, and finally the ones of each word of the vocabulary, so the computing cost is expensive especially when computing the word rank map of the text database. In order to reduce the complexity, we will use the "inverted index" for storage optimization.

Figure 12: The distribution of frequent words distances of texts containing "China"

Because the location lab is simply designed, the capacity of the proposed method is one word per text. In order to increase the capacity of information hiding, we will design better lab location methods in the future.

5 Conclusion

This paper presented a coverless text information hiding method based on the frequent words hash. By using the words rank map and the frequent words hash, normal texts containing the secret information could be retrieved from the text database, and will be sent to the receiver without any modification. Because there is no embedding, the information hiding does not change the probability distribution of the covers. Therefore, the proposed method is theoretically safe, and could be able to escape from almost all state-of-the-art steganalysis methods.

Acknowledgments

This work is supported by National Natural Science Foundation of China (61304208), Open Fund of Demonstration Base of Internet Application Innovative Open Platform of Department of Education (KJRP1402), Open Fund of China-USA Computer Science Research Center (KJR16239), Open Fund of China-USA Computer Science Research Center (KJR16239), Hunan Province Science And Technology Plan Project Fund (2012GK3120), Scientific Research Fund of Hunan Province Education Department (13CY003, 14B106), Changsha City Science and Technology Plan Program (K1501013-11), Hunan Normal University University-Industry Cooperation Fund, and Youth Scientific Research Foundation of Central South University of Forestry & Technology (QJ2012009A)

References

[1] E. O. Blass, T. Mayberry, G. Noubir, and K. Onarlioglu, "Toward robust hidden volumes using write-only oblivious RAM," in *ACM SIGSAC Conference on Computer and Communications Security (CCS'14)*, pp. 203–214, Scottsdale, USA, 2014.

[2] S. Bo, Z. Hu, L. Wu, and H. Zhou, *Steganography of Telecommunication Information*, Beijing: National Defense University Press, 2005.

[3] J. T. Brassil, S. H. Low, and N. F. Maxemchuk, "Copyright protection for the electronic distribution of text documents," *Proceedings of the IEEE*, vol. 87, no. 7, pp. 1181–1196, 1999.

[4] C. Cachin, "An information-theoretic model for steganography," in *The Second Workshop on Information Hiding*, pp. 306–318, Oregon, USA, 1998.

[5] X. Chen, S. Chen, and Y. Wu, "Coverless information hiding method based on the chinese character encoding," *Journal of Internet Technology*, vol. 18, no. 2, pp. 91–98, 2017.

[6] X. Chen, H. Sun, Y. Tobe, Z. Zhou, and X. Sun, "Coverless information hiding method based on the chinese mathematical express," in *The First International Conference on Cloud Computing and Security (ICCCS'15)*, pp. 133–143, Nanjing, China, 2015.

[7] L. Huang, L. Tseng, and M. Hwang, "The study on data hiding in medical images," *International Journal of Network Security*, vol. 14, no. 6, pp. 301–309, 2012.

[8] S. Katzenbeisser, F. Petitcolas, *Information Hiding Techniques for Steganography and Digital Watermarking*, Artech House Publishers, 2000.

[9] T. Y. Liu, W. H. Tsai, "A new steganographic method for data hiding in microsoft word documents by a change tracking technique," *IEEE Transactions on Information Forensics and Security*, vol. 2, no. 1, pp. 24–30, 2007.

[10] S. H. Low, N. F. Maxemchuk, and J. T. Brassil, "Document marking and identification using both line and word shifting," in *IEEE International Conference on Computer Communications (Infocom'95)*, pp. 853–860, Boston, USA, 1995.

[11] S. H. Low, N. F. Maxemchuk, and A. M. Lapone, "Document identification for copyright protection using centroid detection," *IEEE Transactions on Communications*, vol. 46, no. 3, pp. 372–383, 1998.

[12] T. Mayberry, E. O. Blass, and A. H. Chan, "Efficient private file retrieval by combining ORAM and PIR," in *The Twentieth Annual Network & Distributed System Security Symposium*, pp. 1–11, San Diego, USA, 2013.

[13] P. Meng, L. Huang, Z. Chen, W. Yang, and M. Yang, "Analysis and detection of translation based steganography," *ACTA Electronica Sinica*, vol. 38, no. 8, pp. 1748–1852, 2012.

[14] F. A. Petitcolas, R. Anderson, and M. Kuhn, "Information hiding - A survey," *Proceedings of the IEEE*, vol. 87, no. 7, pp. 1062–1078, 1999.

[15] L. Y. Por, T. F. Ang, and B. Delina, "WhiteSteg: A new scheme in information hiding using text steganography," *WSEAS Transactions on Computers*, vol. 7, no. 6, pp. 735–745, 2008.

[16] S. Wang, C. Xiao, and Y. Lin, "A high bitrate information hiding algorithm for video in video," *International Journal of Computer, Electrical, Automation, Control and Information Engineering*, vol. 3, no. 11, pp. 2572–2577, 2009.

[17] N. Wu, M. Hwang, "Data hiding: Current status and key issues," *International Journal of Network Security*, vol. 4, no. 1, pp. 1–9, Jan. 2007.

[18] Z. Xia, X. Wang, X. Sun, Q. Liu, and N. Xiong, "Steganalysis of LSB matching using differences between nonadjacent pixels," *Multimedia Tools and Applications*, vol. 75, no. 4, pp. 1947–1962, 2016.

[19] Z. Xia, X. Wang, X. Sun, and B. Wang, "Steganalysis of least significant bit matching using multi-order differences," *Security and Communication Networks,* vol. 7, no. 8, pp. 1283–1291, 2014.

[20] Z. Xia, X. Wang, L. Zhang, Z. Qin, X. Sun, and K. Ren, "A privacy-preserving and copy-deterrence content-based image retrieval scheme in cloud computing," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 11, pp. 2594–2608, 2016.

[21] C. Yuan, X. Sun, and R. Lv, "Fingerprint liveness detection based on multi-scale LPQ and PCA," *China Communications*, vol. 13, no. 7, pp. 60–65, 2016.

[22] C. Yuan, Z. Xia, and X. Sun, "Coverless image steganography based on SIFT and BOF," *Journal of Internet Technology*, vol. 18, no. 2, pp. 209–216, 2017.

[23] J. Zhang, J. Shen, L. Wang, and H. Lin, "Coverless text information hiding method based on the word rank map," in *The Second International Conference on Cloud Computing and Security (ICCCS'16)*, pp. 145–155, Nanjing, China, 2016.

[24] Z. Zhou, H. Sun, R. Harit, X. Chen, and X. Sun, "Coverless image steganography without embedding," in *The First International Conference on Cloud Computing and Security (ICCCS'15)*, pp. 123–132, Nanjing, China, 2015.

[25] Z. Zhou, Y. Wang, Q.M. Jonathan Wu, C. Yang, and X. Sun, "Effective and efficient global context verification for image copy detection," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 1, pp. 48–63, 2017.

Biography

Jianjun Zhang works as an associate professor in Hunan Normal University, China, and is currently working towards the PhD degree in computer science and technology at the College of Computer Science and Electronic Engineering, in Hunan University, China. His research interests include network and information

security.

Huajun Huang is currently a faculty member in the college of Computer and Information Engineering at Central South University of Forestry & Technology. His overall research area include of Webpage information hiding and hidden information detection, XML Watermarking, Anti-phishing, Mobile Device Forensics. Dr. Huang received his Ph.D. from Hunan University in 2007, M.S. degrees from Hunan University in Software Engineering (2004), and a B.A. in Applied Physics from Yunnan University (2001).

Lucai Wang received the BE degree in Hunan University, China, in 1990, the PhD degree in Electronic Information Engineering from Hunan University, China, in 2006. He works as a professor in the College of Engineering and Design, Hunan Normal University. His research interests include intelligent information processing.

Haijun Lin received the BE degree in Hunan Normal University, China, in 2004, the PhD degree in Electronic Information Engineering from Hunan University, China, in 2009. He works as an associate professor in the College of Engineering and Design, Hunan Normal University. His research interests include intelligent information processing.

Deng Gao is currently a faculty member in the college of Software at Hunan Vocational College of Science and Technology. She received her M.S. degree from Central South University in Software Engineering in 2016. Her research interests include data mining and semantic networks.

Advanced Random Time Queue Blocking for Effective Protection of Application Servers Against Low-Rate DoS Attacks

R. Kavitha[1,2], G. Padmavathi[1]

(Corresponding author: R. Kavitha)

Avinashilingam Institute for Home Science and Higher Education for Women[1]
Department of Computer Science, Sri Krishna Arts and Science College[2]
Coimbatore, Tamilnadu, India
(Email: jayabal.kavitha@gmail.com)

(Received July 21, 2016; revised and accepted Nov. 15 & Dec. 25, 2016)

Abstract

Low-rate traffic denial-of-service (DoS) attacks are a strategy to deny services of a network by detecting the vulnerabilities in the application behaviors. The low-rate DoS attack against the application servers is considered in this paper with the motive to develop an efficient defense technique against the low-rate DoS attack. Among different defense techniques, the Improved Random Time Queue Blocking (IRTQB) performs better than other methods. IRTQB performs similar to Random Time Queue Blocking (RTQB), but it selectively chooses the blocking interval requests only from the potential attackers and discards them. However, the differentiation of the attacker requests from the legitimate users' is not always efficient as only the source IP addresses and the record timestamp are considered. This can be improved when considering more complex set of features. Hence, in this paper, the Advanced Random Time Queue Blocking (ARTQB) scheme is proposed by additionally employing Bandwidth utilization of attacker and legitimate user in IRTQB. ARTQB defines Spatial Similarity Metric (SSM) between the requests in terms of source IP addresses, the record timestamp and the bandwidth. Thus the defense of the application server against the low-rate DoS attack is be improved than IRTQB. Experimental results show that the proposed ARTQB performs better protection of Low-Rate DoS Attack against Application Servers (LoRDAS) by reducing the attack efficiency and attack impact on the server.

Keywords: ARTQB; IRTQB; Low-rate Denial-of-Service (LDoS); RTQB; Spatial Similarity Metric (SSM)

1 Introduction

Denial-of-service (DoS) is a type of attack in which the attackers attempt to prevent the legitimate users from accessing the network services. In a DoS attack, normally the attacker transmits unnecessary messages which are having invalid return addresses and requiring the network or server to authenticate requests [9]. While sending the authentication approval, the network or server has no ability to find the return address of the attacker, and causing the server to wait before closing the connection. The attacker transmits more authentication messages along with invalid return addresses while the server closes the connection. Hence, the process of authentication and server wait will begin again, keeping the network or server busy.

A network or host can be compromised with DDoS attacks using two types of traffic, namely, high-rate DoS traffic and low-rate DoS traffic [2]. DoS attacks are implemented in terms of many ways. The most common ways are the flooding the network to reduce the legitimate network traffic, disrupting the connections between the user and the server, blocking certain range of users and disrupting the state of the information of the users [10]. However the detection of DoS has become easier as it generates high inconsistent traffic rate by which the detection algorithms assures the presence of attack. Thus, low rate attacks came into real timer applications in which the DoS is achieved in low traffic scenarios.

Low-Rate DoS attacks (LRDoS) are new types of DoS attacks. In LRDoS the attacker sends a burst of well-timed packets, creating packet losses in a link and increments the retransmission timeout for only certain TCP flows. As these traffic bursts are sent during the expiration times, the overall traffic is reduced considerably thus disabling the efficiency of detection. Many techniques have been presented in the recent past to detect the LRDoS but most of the techniques performed below expectation. The introduction of new LRDoS such as Shrew and reduction of quality (RoQ) attacks increases the detection complexity.

In this paper, the DoS detection schemes such as random service time (RST), Random answer instant (RAI), Random time queue blocking (RTQB) and Improved Random Time Queue Blocking (IRTQB) are analyzed to determine the detection efficiency. The analysis results show that the IRTQB performs better than the other three methods; however the IRTQB also suffers from limitations. Particularly the differentiation between the attack requests from the legitimate users' requests is not satisfactory. Hence, bandwidth utilization is included in IRTQB to develop Advanced Random Time Queue Blocking (ARTQB) scheme for effective defense of the application servers.

The remainder of the article is organized as follows: Section 2 describes the related researches briefly. Section 3 presents the methodologies utilized in the paper. Section 4 provides the experimental results and their discussions. Section 5 concludes the research.

2 Related Works

Macia-Fernandez et al. [5] proposed evaluation method of low-rate DoS attack (LRDoS) against the iterative servers. The evaluated attack characteristics are analyzed and the potential effects of the attack are also analyzed. The iterative servers are that those servers limited to handle just a single service request compared to multiple service requests handled by the concurrent servers. The analysis provides the vulnerability details of the iterative servers and provides better possibility of forecasting the statistical metrics about the server behavior. The low-rate traffic behavior helps to subvert the provided service. It is also possible to tune the parameters of the attack in order to choose the suitable values for the efficiency and the amount of load generated in the target server. Thus, it becomes possible to bypass the intrusion detection system intended to protect the attacked server.

Macia-Fernandez et al. [6] in another work extended the analysis results to support the analysis of the low-rate DoS attacks against the concurrent servers like persistent HTTP servers. The mathematical model for the low-rate DoS attacks against the application server has been presented in an extended work [7] for the evaluation of the attack in servers with superposition among the occurrence probability functions. In another extension [8] presented four efficient methods to tackle the low-rate DoS attacks against the application server.

The efficient alternative techniques are based on the blocking the entry of the requests in the service queue of the server. Thus the efficiency of the low-rate DoS attack can be reduced effectively without any impact on the amount of time spent by the requests in the generated systems. However, there are some limitations in which the attack requests and the legitimate users' requests are not effectively differentiated in some instances.

Wang et al. [12] proposed a queuing analysis scheme for the evaluation of the DoS attacks in the computer networks. The stationary probability distribution can be determined by developing a memory-efficient algorithm and the computed probability distribution can be utilized for finding other interesting performance metrics like the connection loss probability and buffer occupancy percentages of half-open connections for regular traffic and attack traffic. Thus the impact of DoS attacks can be detected even in complicated computer networks.

Tang et al. [11] proposed a vulnerability model of feedback-control based internet services to tackle the low-rate DoS attacks. The fundamental queries, namely the impact of the LRDoS attack on the feedback-control based systems and how the systematic evaluation of LR-DoS is performed has been the center point of research. These problems are tackled by considering the target system as a switched system. Both the oscillation of steady state error and staying away from the desired state impair the system's performance and hence a novel methodology is used to analyze the impact of the attack. However the tradeoff between the effectiveness and the cost of LRDoS attack has not been explained which hinders the analysis.

Wu et al. [13] proposed an LRDoS attack detection scheme based on the network multi-fractal called as multi-fractal detrended fluctuation analysis (MF-DFA). The scheme detects the changes in terms of the multi-fractural characteristics of the network traffic, which helps in finding the LRDoS attack flows.

Adi et al. [1] introduced an analysis technique to demonstrate the impact of LRDoS attacks against the HTTP/2 servers. The resource consuming HTTP packets are transmitted along with the principle of sending requests in order to serve the full capacity of the servers. The HTTP/2 packets serve as the underlying standard and there are no computationally expensive applications due to a backend server, which is not connected to the HTTP/2 server. The server memory degrades at a certain rate indicating the presence of LRDoS.

Brynielsson et al. [4] presented a spectral analysis based detection of LRDoS attacks against the HTTP server. The weakness of the HTTP server is analyzed and the attack simulator has been developed. When the attack is present, disproportionate amounts of energy in the lower frequencies can be detected effectively. Thus the approach serves as a medium of detection with the attacker has fixed wait times or floods the server when initiating the attack. However the major drawback is that the attacker has certain approaches to reduce the disproportionate amounts of energy to some extend so that the attack detection becomes very difficult.

Bedi et al. [3] introduced the enhanced AQM technique called as Deterministic Fair Sharing (DFS) for tackling the congestion based DoS attacks. The concept of fair buffer share is dynamically determined for each competing flow to ensure optimal fairness. It is achieved in DFS by utilizing a set of data structures in combination to provide low operational overhead while maintaining limited per-flow state and offer high DoS attack identification capability. Thus the congestion based DoS attacks

can be detected effectively and DFS provides a higher degree of fairness and throughput to legitimate flows while stabilizing the router queue length and allowing the least bandwidth to the attack traffic.

Though the methods discussed in the literature are effective in the defense against the low-rate DoS attacks, there seem to be more drawbacks that reduce the overall performance. In some methods, the LoRDAS attacks become complex to detect and hence the attack efficiency becomes not possible to be minimized. The tradeoff between the effectiveness and the cost of LRDoS attack influences the attack performance, but it is not considered in the existing methods. The smart attackers have the tendency of reducing the disproportionate amounts of energy in the existing methods also affects the attack detection.

Moreover, in the existing methods, the differentiation of the legitimate user requests and the attacker requests is appropriately addressed. Similarly, only the record timestamp and the source IP address are considered for reducing the LoRDAS attack efficiency, which seems efficient; however making room for improvement.

Hence the need for a novel strategy is needed to effectively detect the attacks and also develop a response technique for reducing the attack efficiency and its impact on the application server. Many methods were introduced to detect the attacks, including the Shrew attack.

Some methods utilized the randomization of the timers in the TCP flows for avoiding the synchronization between the periodic arrival of short attack bursts and the expiration of a timer. But there are no effective solutions for the LoRDAS. The Improved Random time queue blocking (IRTQB) has been the effective solution till date but even it has relatively near-par solutions only. IRTQB employs record timestamp and the source IP address for reducing attack impact. In this paper, ARTQB is proposed with the bandwidth utilization additionally considered for the minimization of the attack efficiency.

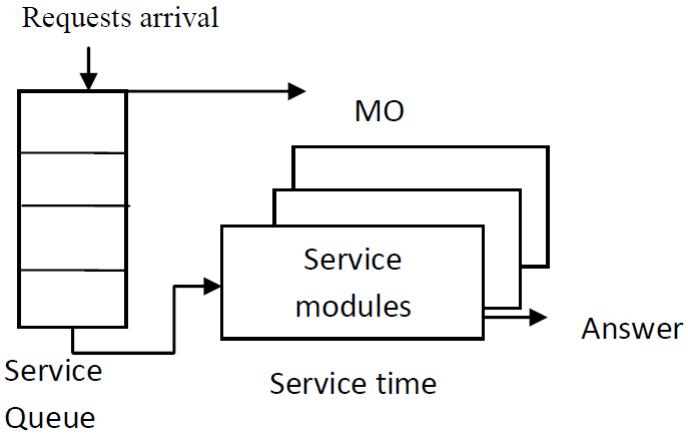


Figure 1: Application server model in LoRDAS attack

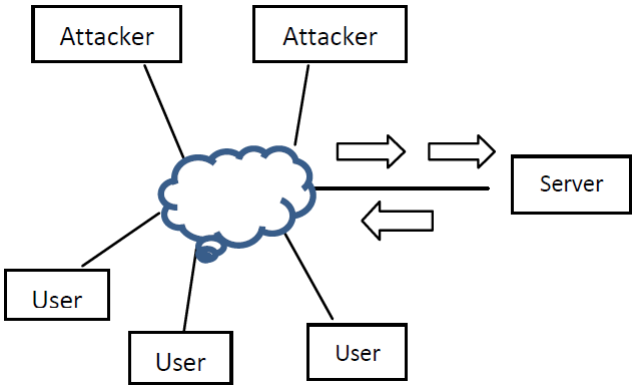


Figure 2: Study scenario

3 Methodologies

3.1 Application Server Model

Application servers are the potential victims of LoRDAS attacks shown in Figure 1. Certain conditions are required for an application to be vulnerable to this kind of attacks, and several different strategies might be followed by the attacker to deny the service. In addition, the necessary network model behind the attack is shown in Figure 2.

The application server model considered in the LoRDAS attack is composed of the following elements (1) a service queue where incoming requests are placed upon their arrival on the server, and (2) one or several service modules which are in charge of processing the requests.

The low-rate DoS attacks in the application servers depends on two major aspects of server behavior such as the presence of deterministic patterns and enabling instants concurrence with the answer instants. The defense methods are developed based on the strategy used for reducing the efficiency of the attack in the servers and without any negative impact on the normal performance of the servers. The fundamental LoRDAS attack can be understood from Figure 3.

3.2 Random Service Time (RST)

RST has been designed with the view of reducing the predictability of the server behavior, thus considerably preventing the server details to the attackers and reducing the efficiency of the attack as shown in Figure 4. In this scheme, the deterministic patterns are eliminated from its mode of operation. So whenever the server utilizes a fixed timeout feature that aims to randomizing the features to make the implementation of the attack more difficult in practice by blocking the prediction of the answer instants and enabling instants. RST can be implemented in a server such that its behavior is maintained with slight modifications that do not alter the overall performance.

When the service of a request in the queue is idle, the service module considers the request from the service queue on the basis of the popular schemes like FIFO and LIFO. Then by utilizing a processing time of the request, the request is deadline constraint. In this period, the at-

tacker manages to send a short attack burst that reaches around the estimated answer time. When the processing is finished, the service module remains locked for a random time called extra delay such that no additional position is enabled as no answers will be generated in this phase.

Either the conditions $\Delta t > \overline{\Delta t_{RST}} + B/2$ or $\Delta t < \overline{\Delta t_{RST}} - \frac{B}{2} - RTT$ must be achieved during the attack bursts so that the answer time is shifted. However condition $\Delta t < \overline{\Delta t_{RST}} - \frac{B}{2} - RTT$ cannot be fulfilled if $\overline{\Delta t_{RST}} < \frac{B}{2} + RTT$. Hence the Δt_{RST} takes values from a uniform distribution with a maximum value Δt_{max}^{RST} in order to maintain, no free positions for the legitimate users in RTT seconds.

$$\Delta t = U[0, \Delta t_{max}^{RST}], \text{ if} \overline{\Delta t_{RST}} < \frac{B}{2} + RTT \tag{1}$$

where, $\overline{\Delta t_{RST}}$ is the mean extra delay, RTT is the round trip time, B is the time period of attack burst, Δt is the variability in service time and Δt_{max}^{RST} is the maximum value of mean extra delay. When $\overline{\Delta t_{RST}} > \frac{B}{2} + RTT$, Δt_{RST} is a random variable sampled from two different uniform variables V_1 and V_2 are utilized.

$$\overline{\Delta t} = \frac{\Delta t_{max}^{RST}}{2} \tag{2}$$

The mean value of the extra delay should be $\frac{\Delta t_{max}^{RST}}{2}$ in order to appropriately shift the answer time. So Δt_{RST} is sampled from V_1 and V_2 with a probability P.

$$\Delta t_{RST} = \begin{cases} V_1 & \text{with probability } P \\ V_2 & \text{with probability } 1-P \end{cases} \tag{3}$$

where V_1 and V_2 are variables and the probability P is calculated by

$$P = \frac{\Delta t_{Max}^{RST} + B}{2(\Delta t_{Max}^{RST} + B + RTT)} \tag{4}$$

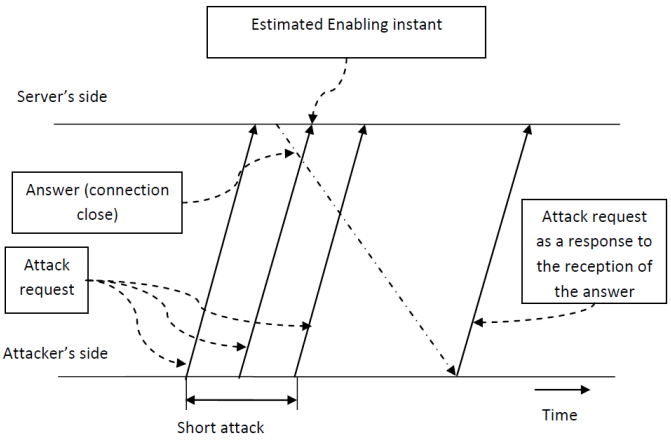


Figure 3: LoRDAS attack process

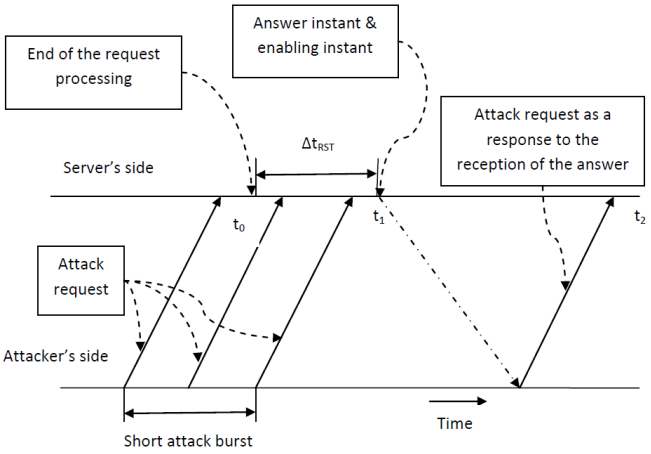


Figure 4: LoRDAS attack when RST is active

The RST initially replaces the answer instant to a newer position which does not come under the control of the attacker. If the extra delay is longer and the condition of the service queue is full of requests, then all of the attack burst traffic will be rejected by the server. The time available for the legitimate user to insert new requests in the service queue depends on the round trip time (RTT) between the server and the attacker while the reception of the attack packet will be sent as a response to the answer.

When the extra delay time expires, the answer is forwarded to the user who requested them. As the efficiency of the attack is reduced, the server performs efficiently and thus a new free position occurs at the end of transmitting the answers to the requested users. If the extra delay time is much longer, then the legitimate user can send a request at the start of the delay time while the attacker's request is assumed to be occurring at the end of the delay time, so that even the attack requests can be inserted at the free positions in the queue without affecting the user requests.

Thus the efficiency of the attack is significantly reduced; however the fact assumed that the attack requests are inserted at the end of the lengthy extra delay. But when an extra delay is added to the original service time, the attacker also perceives an increase in the estimation of the service time, so that the attack parameters can be adjusted to synchronize the attack bursts.

3.3 Random Answer Instant (RAI)

Random answer instant (RAI) differs from the RST technique by utilizing the decoupling the answer instants and the enabling instants instead of introducing variability in the server behavior as performed in RST as shown in Figure 5. The service of a request is extracted and processed during the service time as in the RST. After processing the requests from the queue, the service model waits for the extra delay similar to RST but the difference being when the first request service is processed, and then the

new request is extracted from the queue and begins processing such that the extra delay becomes non-blocking. Thus a new position in the queue is enabled at this period, enabling instant. After the completion of extra delay, the answer is sent to the corresponding user which is the answer instant.

When $\overline{T_S} + \overline{\Delta t_{RAI}}$ is the service time, a short attack burst is send at the answer instant $t_1 - t_0$. Δt is given by $\Delta t = U[\Delta t_{min}^{RAI}, \Delta t_{max}^{RAI}]$. The lower limit is given as $\Delta t_{min}^{RAI} = \frac{B}{2}$ thus ensuring the time interval duration $\overline{\Delta t_{RAI}} - B/2$ between the enabling instant and the arrival of the short attack burst.

$$\Delta t = \begin{cases} 0 & \text{if } \Delta t_{max}^{RAI} < B/2 \\ U[\Delta t_{min}^{RAI}, \Delta t_{max}^{RAI}] & \text{if } \Delta t_{max}^{RAI} > B/2 \end{cases} \quad (5)$$

where, Δt_{max}^{RAI} is the upper limit of the uniform distribution to select Δt_{RAI} and Δt_{min}^{RAI} is the lower limit of the uniform distribution to select Δt_{RAI}. The upper limit Δt_{max}^{RAI} must be high as possible in order to introduce higher variability. But by considering impact the values becomes

$$t_1^{RAI} = [Nt_s + \Delta t_{RAI}, (N+1)t_s + \Delta t_{RAI}] \quad (6)$$

In Equation (6), t_1^{RAI} is the time interval for new incoming request in RAI mechanism, t_s is the service time, Δt_{RAI} is the mean extra delay in RAI, N is the number of requests. Thus Δt_{max}^{RAI} should be configured as a trade-off between reducing the impact and increasing the variability in the server so that RAI reduces the impact on the normal behavior of the server so that the effectiveness of the attack is reduced. However there is an interval called as the tradeoff between the reduction of the impact and increase of variability in the server, which causes impact on the normal behavior of the server.

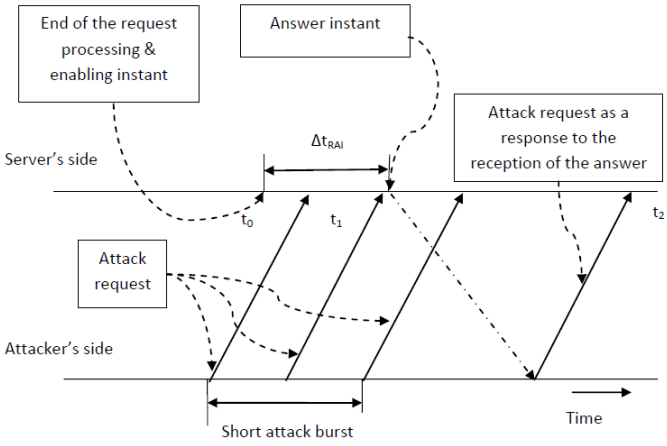


Figure 5. LoRDAS attack when RAI is active

Figure 5: LoRDAS attack when RAI is active

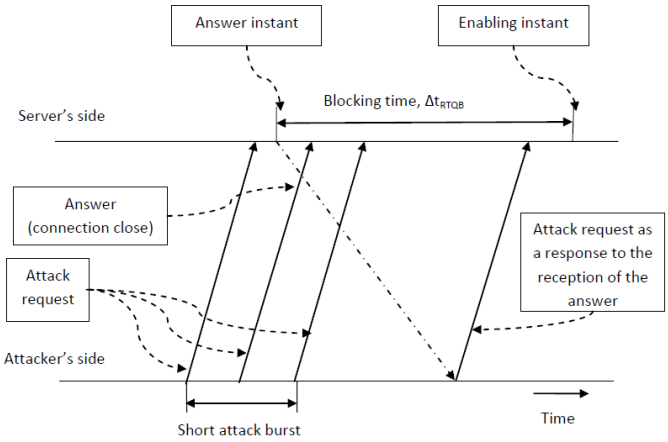


Figure 6: LoRDAS attack when RTQB is active

3.4 Random Time Queue Blocking (RTQB)

As RST and RAI have certain limitations, Random time queues blocking (RTQB) is introduced to overcome the shortcomings as shown in Figure 6. RTQB aims at reducing the attack efficiency without creating any negative impact in the server behavior. The main concept of RTQB is that when the attacker is able to accurately estimate the answer instants, then the short attack bursts will arrive with the response attack messages arriving in RTT seconds. In this situation, the legitimate users are distributing the requests while the attackers will be considering the response messages. Hence, in RTQB all the requests arriving at this time interval are blocked so that the attack efficiency and the impact are reduced.

Ideally, the value for $\Delta t_{RTQB} = RTT$, all the attack requests are projected in an interval $[-B/2, RTT]$ around the answer instant. But Δt_{RTQB} is configured as a random value taken from uniform distribution

$$\Delta t_{RTQB} = U[RTT, \Delta t_{max}^{RTQB}] \quad (7)$$

In Equation (7), Δt_{RTQB} is the mean extra delay for RTQB. There are two reasons for considering a random value. First, the attacker estimation of both the answer instants and RTT is not perfect and, attack packets will arrive even after RTT seconds from the answer instant. Second, it is recommendable to introduce certain variability into the process; the attacker might be capable of estimating the value Δt_{RTQB} and adapting the attack.

The service of a request is performed similar to the RST and RAI. After the processing of the requests, the answers are sent to the requested users in answer instant and the extraction of new requests is enabled from the queue in enabling instant. After the answer instant, all new requests are rejected during the specified interval of time. After this interval, the new requests are again accepted and the processing begins. Thus the effect of an attack can be reduced while once a request enters during the active stage of RTQB; the server behavior is also not affected. The impact of RTQB on the server behavior, it is clear that when RTQB is active and once a request

enters into the service queue, there is no difference in its process. Thus, no impact is present due to the use of RTQB. However, as RTQB blocks all the requests during the specified time interval, it fails to select the legitimate user requests.

3.5 Improved Random Time Queue Blocking (IRTQB)

In order to avoid discarding of important legitimate user requests during the time interval of RTQB, improved version called Improved Random Time Queue Blocking (IRTQB) is introduced as shown in Figure 7. IRTQB selectively chooses the requests during the time interval around the answer instants which comes from the potential attackers are only discarded. As the attacker uses only restricted spoofing mechanisms that are allowed within the same network segment, the attack is limited. The attack interval starts only at $t_0 - B/2$ where B is the length of an attack burst. Thus, when the attack length is B, the interval becomes $[t_0 - B/2, t_0 + \Delta t_{max}^{IRTQB}]$. The spatial similarity metric SSM is defined between the two requests to measure the probability that they come from the same source. In IRTQB, the SSM includes the source IP addresses only. For two generic IP addresses A_i and A_j, the similarity metric is computed as the number of consecutive bits set to '1' in the bit XNOR operation of the two addresses.

$$SSM(A_i, A_j) = \#\_consecutive\_bits_1(A_i \text{ XNOR } A_j) \quad (8)$$

A spatial similarity metric (SSM) is defined between two requests to measure the probability of the same source. SSM includes the source IP addresses of incoming requests for the analysis of the source. IRTQB maintains a record of timestamps and source IP addresses for all incoming requests with which the similarity is computed. If the similarity between the two requests is higher than a predefined threshold $SSM(A_i, A_j) > SSM_{Th}$ then, both are discarded without any notification to the users. Thus, the attackers are prevented from obtaining information about the requests.

3.6 Limitations of RST, RAI, RTQB, IRTQB

RST decreases the attack efficiency by shifting the answer time to a position that is not controlled by the attacker and also by adding a source of variability in the server behavior. However, when the extra delay included in the original service time is longer, the attacker also perceives an increase in the estimation of the service time, so that the attack parameters can be adjusted to synchronize the attack bursts. The reduction of attack effectiveness in RST is limited due to the fact that the maximum amount of time for legitimate users to seize new positions in the queue is the RTT. RAI technique performs better than RST and even provides better performance than the RTQB. However, in RAI the tradeoff between the reduction of the impact and increase of variability on the server, which causes an impact on the normal behavior of the server. As the defense technique should not cause any impact on the server, RAI is avoided and RTQB is presented. RTQB reduces the impact on server while also reducing the attack efficiency. It makes advantage of the attackers using short bursts of traffic that arrive around the answer instants and blocks all the incoming requests in a time interval. During this interval RTQB does not selectively choose the requests and blocks all requests without analyzing the request sender. This becomes a major drawback in RTQB which leads to an improved technique called IRTQB. IRTQB employs SSM and selectively blocks the requests so that the queue has at least one free space and the attack efficiency is reduced.

Though IRTQB performs better than RST, RAI and RTQB by reducing attack efficiency and no impact on the server, the SSM metric only includes the source IP address and record timestamps. By including more reliable factors, the performance can be further improved. Thus the need for more advanced defense technique with SSM considering more factors arises. This need is achieved by including the bandwidth utilization factor with source IP address and record timestamps in the proposed ARTQB defense technique.

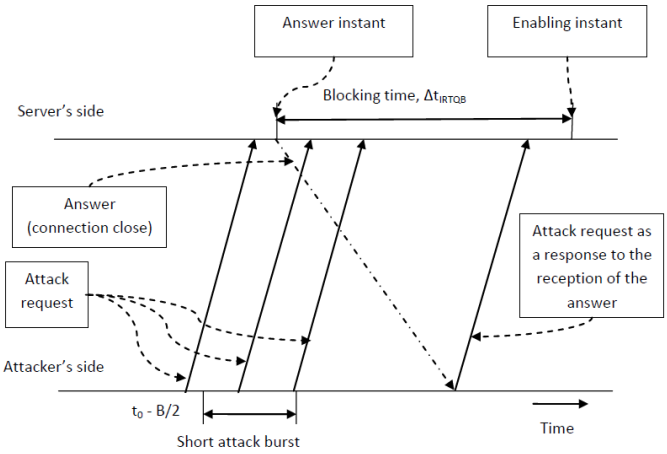


Figure 7: LoRDAS attack when IRTQB is active

3.7 Advanced Random Time Queue Blocking (ARTQB)

Though IRTQB significantly reduces the attack efficiency without impact on server behavior, still there is scope for improvement in reducing the attack efficiency. IRTQB is very efficient, but at the same time it is more costly than other methods and hence improving the performance without further increasing the cost is much more significant. Hence, a highly improved version of RTQB is proposed which is called as Advanced Random Time Queue Blocking (ARTQB) as shown in Figure 8. Consider the

utilization bandwidth of user b_u, is employed by the service in processing a given request. The service bandwidth for legitimate request and attack requests are varied for their processing. Therefore, the service bandwidth for identical requests is modeled as a random variable B_u, with the normal distribution. The mean value of B_u is denoted as $\overline{B_u}$ and the variance is denoted as $var[B_u]$.

$$B_u = N(\overline{B_u}, var[B_u]) \tag{9}$$

Where, N refers the normal distribution. Consider Δb_{ARTQB} is the extra bandwidth utilized and the value for $\Delta b_{ARTQB} = var[B_u]$, as all the attack requests are expected in an interval $[\overline{B_u}, max(var[B_u])]$ around the answer instant. Hence, Δb_{ARTQB} is configured as a random value which is taken from the uniform distribution:

$$\Delta b_{ARTQB} = U[\overline{B_u}, \Delta b_{max}^{ARTQB}] \tag{10}$$

Where, Δb_{max}^{ARTQB} is the maximum value of Δb_{ARTQB} and U refers the uniform distribution. Assume, the bandwidth utilization of the attacker's requests are around the answer instants are very high. Assume the queue contains $N-1$ requests and that they are all processed by the service module at the rate of bandwidth b_u Hz per request. The attack bandwidth b_1 for new incoming request is computed as,

$$b_1^{ARTQB} = [Nb_u + \Delta b_{ARTQB}, (N+1)b_u + \Delta b_{ARTQB}] \tag{11}$$

Therefore, the similarity metric SSM, is measured between two requests in order to identify the probability that they come from the same source including same bandwidth. Similar to Equation (8), the spatial similarity metric is computed as the number of consecutive bits set to '1' in the bit XNOR operation of the two bandwidths by considering two bandwidths B_i and B_j:

$$SSM(B_i, B_j) = \#\_consecutive\_bits_1(B_i \text{ XNOR } B_j) \tag{12}$$

Then, the simple spatial similarity metric for both bandwidth and IP addresses is computed as,

$$\begin{aligned} & SSM(A_i, A_j, B_i, B_j) \\ = \; & SSM(A_i, A_j) + SSM(B_i, B_j) \qquad (13) \\ = \; & \#\_consecutive\_bits_1(A_i \text{ XNOR } A_j) \\ & + \#\_consecutive\_bits_1(B_i \text{ XNOR } B_j) \qquad (14) \end{aligned}$$

3.8 Description

ARTQB extracts the requests from the service queue and processes them at the service time T_s. After completing the processing the answers are send to the legitimate users who requested them while on the other side called enabling instant, the new requests are started to process. These requests arriving at attack interval t_1 and attack bandwidth b_1 are selectively chosen and those from the attackers are discarded. When the attack interval expires, new requests are accepted again and the processing begins. ARTQB maintains a list of answer instants from the beginning of the server operation. Similar to IRTQB, around every answer instant, the attack interval and attack bandwidth exists indicating that the attack packets will be arriving during the interval and bandwidth. This helps in avoiding those attacks in the form of requests. The attack interval starts at the length of attack burst B which means that the interval begins at halfway through the initial time and the attack bandwidth starts that the bandwidth begins at halfway through the initial bandwidth. At this instant, no requests will enter the queue but these will be helpful in deciding the new requests as

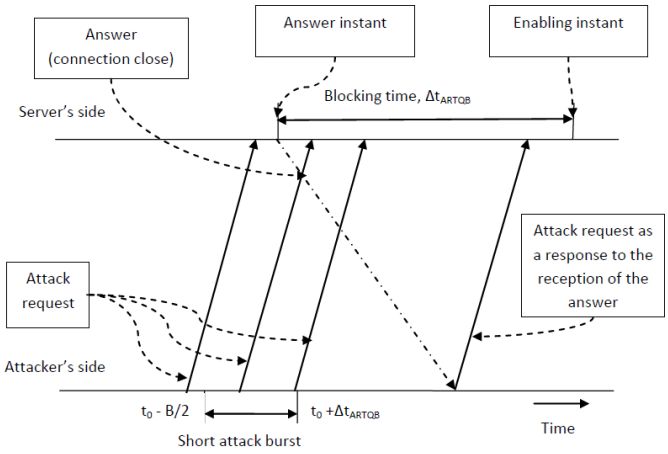


Figure 8: LoRDAS attack when ARTQB is active

Algorithm 1 ARTQB execution

```
Extract request from service queue
Processing request at service time T_s
//Answer generation
For every answer
Insert answer instant in a list L
Compute attack interval t_1
Compute attack bandwidth b_1
Send answer to users
Extract new request (enabling instant)
End For
For every incoming request R_a
Record timestamp, source IP, bandwidth utilization
Determine t_1 and b_1 for R_a
For all requests R_b in t_1 and b_1
Compute SSM
Determine a threshold for SSM, SSM_Th
if (SSM(R_a, R_b) > SSM_Th) then
    Discard R_a and R_b
else
    Insert Request in Queue
end if
End For
```

legitimate or attack requests.

ARTQB also maintains a list containing the timestamps, source IP addresses and bandwidth utilization for all incoming requests. Using the record list, for all the incoming requests the SSM is computed between the incoming request R_a and every request R_b arriving in the attack interval and bandwidth. If the SSM is higher than the defined threshold SSM_{Th}, both the requests R_a and R_b are discarded while in other case the incoming request is accepted. Thus the attack efficiency is reduced below half when the bandwidth utilization is included along with timestamps and source IP addresses in the SSM metric. It is also noted that there is no impact in the server behavior when the ARTQB scheme is active in the server. The overall flow of ARTQB is shown in Figure 9.

4 Performance Evaluations

In this section, the proposed defense techniques are evaluated experimentally in the Network Simulator-2. The performance of the low-rate DoS attack is evaluated by measuring the mean in-system time and the attack efficiency. The low-rate DoS attack is employed in the application server inorder to evaluate the attack impact in the server. The server is supplied with different defense techniques namely RST, RAI, RTQB, IRTQB and ARTQB with configuration parameters prescribed with IRTQB and ARTQB considering additional configuration parameter called the SSM threshold. Table 1 shows the configuration values for attack and server parameters in Scenario 1 (S_1), Scenario 2 (S_2) and Scenario 3 (S_3).

Table 1: Configuration values for the attack and server parameters

Parameter	Value
Duration of attack burst, B	0.4s
Time between attack packets in a burst	0.2s
Mean service time, Ts	12s
Variance of server, $var[T_s]$	$0(S_1), 0.2(S_2, S_3)$
Interval between legitimate users requests	$3s(S_1, S_2), 0.95s(S_3)$
Number of server threads	$1(S_1, S_2), 4(S_3)$
Number of positions in service queue, N	$4(S_1, S_2), 8(S_3)$
Number of attack threads	-N
Round trip time, RTT	1s
Similarity metric, ST	32

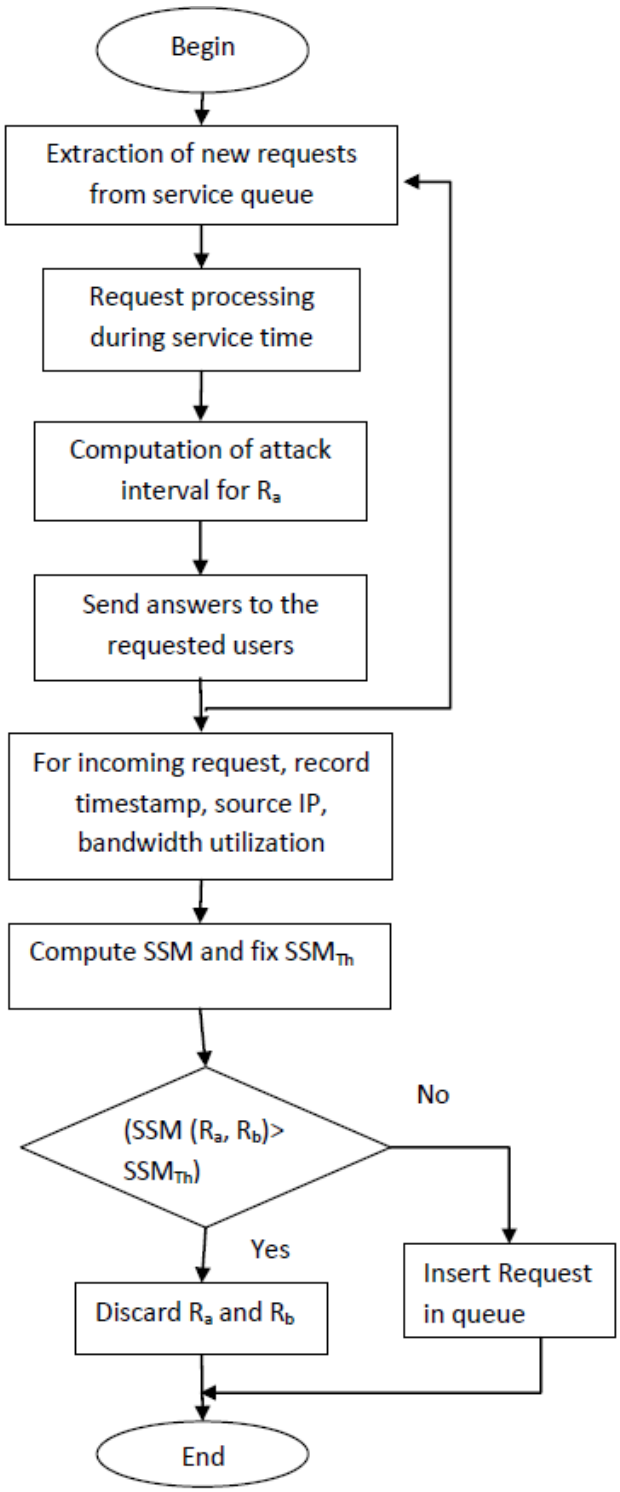


Figure 9: Overall flow of ARTQB

Scenario 1, S_1: The server is mono-threaded and the variance of the service time for attack requests, $var[T_s]$, and $var[B_s]$ is 0. Scenario 2, S_2: Server variance of $var[T_s]$, and $var[B_s]$ is modified. Scenario 3, S_3: The aim of this scenario is to check how a multithread operation in the server affects the performance of a given defense technique. The Figures 10, 11, and 12 show the comparison of RST, RAI, RTQB, IRTQB and ARTQB in terms of attack efficiency (%) in three scenarios S_1, S_2 and S_3. The Figures 13, 14, and 15 show the comparison in terms of mean in-system time (s) in scenarios S_1, S_2 and S_3.

4.1 Attack Efficiency

Attack efficiency is the percentage of service queue positions captured by the attacker over the total number of positions captured during the attack execution.

Figure 10 shows that the attack efficiency comparison of RST, RAI, RTQB, IRTQB and ARTQB during Scenario 1 (mono-threaded with zero variance). In Scenario 1, the attack efficiency decreases from 100% to an asymptotic value. The attack efficiency of ARTQB has been much reduced than the other mechanisms since, consideration of the bandwidth utilization. It shows that the 100% attack efficiency in RST decreases to 39% for ARTQB whereas the attack efficiency of other mechanisms such as RAI, RTQB and IRTQB are 90%, 82% and 60% in the time period of 2sec. When the time period is 12sec, the attack efficiency of ARTQB is 20% which is lower than the other defense mechanisms.

Figure 11 shows that the attack efficiency comparison of RST, RAI, RTQB, IRTQB and ARTQB during Scenario 2 (di- threaded with variance 0.2). In Scenario 2, the attack efficiency decreases from 85% to an asymptotic value. By considering the bandwidth utilization in ARTQB, the attack efficiency is reduced compared with other mechanisms. It shows that when the time period is 2sec, the attack efficiency of ARTQB is 54% which is smaller than the other mechanisms. Also, it is clear that the 85% attack efficiency in RST decreases to 27% for ARTQB whereas the attack efficiency of other mechanisms such as RAI, RTQB and IRTQB are 51%, 41% and 35% in the time period of 12sec.

Figure 12 shows the attack efficiency comparison of RST, RAI, RTQB, IRTQB and ARTQB during Scenario 3 (multi-threaded with variance 0.2). In Scenario 3, the attack efficiency of ARTQB is much reduced than the IRTQB technique by considering the bandwidth utilization. It shows that when the time period is 2sec, the attack efficiency value of ARTQB is 46% compared with other mechanisms. Moreover, it is clear that the 87% attack efficiency in RST decreases to 23% for ARTQB whereas the attack efficiency of other mechanisms such as RAI, RTQB and IRTQB are 55%, 44% and 35% in the time period of 12sec.

4.2 Mean In-system Time

Mean in-system time is the time from when a request enters the server to the instant at which its corresponding answer is sent.

Figure 13 shows that the mean in-system time comparison of RST, RAI, RTQB, IRTQB and ARTQB during Scenario 1 (mono-threaded with zero variance). In Scenario 1, the mean in-system time decreases from RST to ARTQB. The reduction in mean in-system time is achieved by considering the bandwidth utilization. It shows that the 87sec in RST decreases to 30sec for ARTQB whereas the mean in-system time of other mechanisms such as RAI, RTQB and IRTQB are 71sec, 49sec and 32sec in the time period of 2sec. When the time period is 12sec, the mean in-system time of ARTQB is 22sec which is lower than the other defense mechanisms.

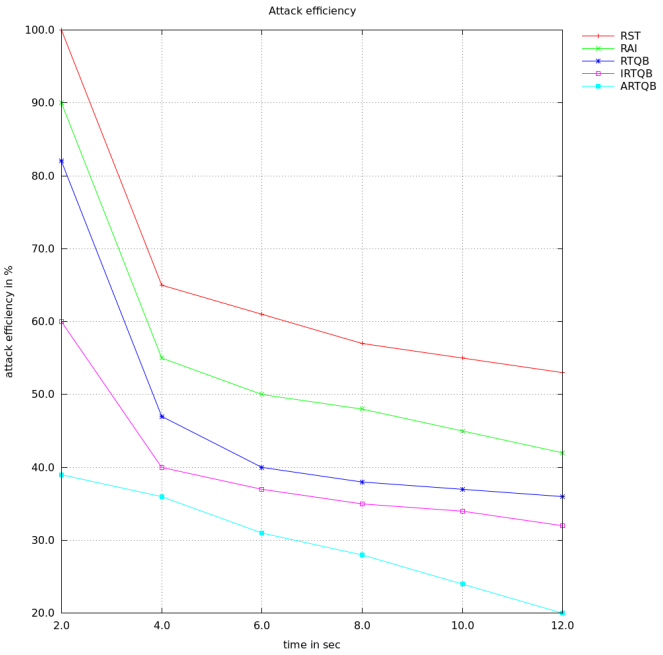


Figure 10: Scenario 1 attack efficiency (%)

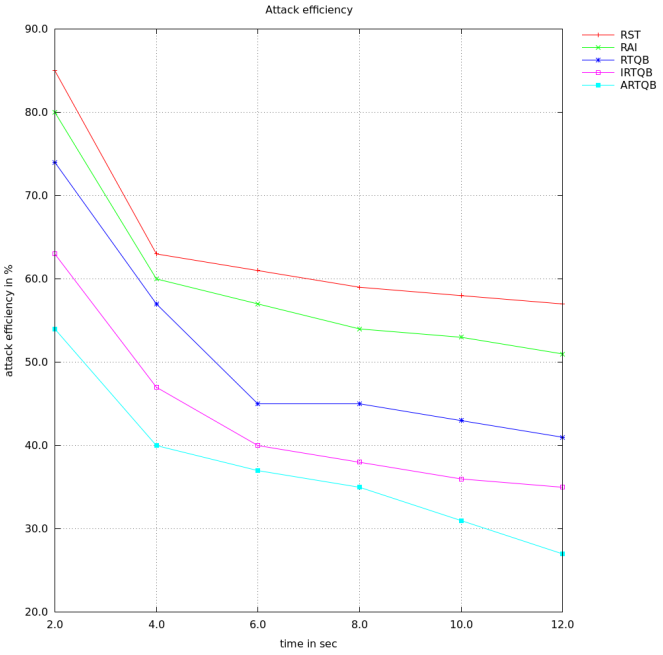


Figure 11: Scenario 2 attack efficiency (%)

Figure 14 shows that the mean in-system time comparison of RST, RAI, RTQB, IRTQB and ARTQB during Scenario 2 (di-threaded with variance 0.2). In Scenario 2, the mean in-system time decreases from RST to ARTQB.

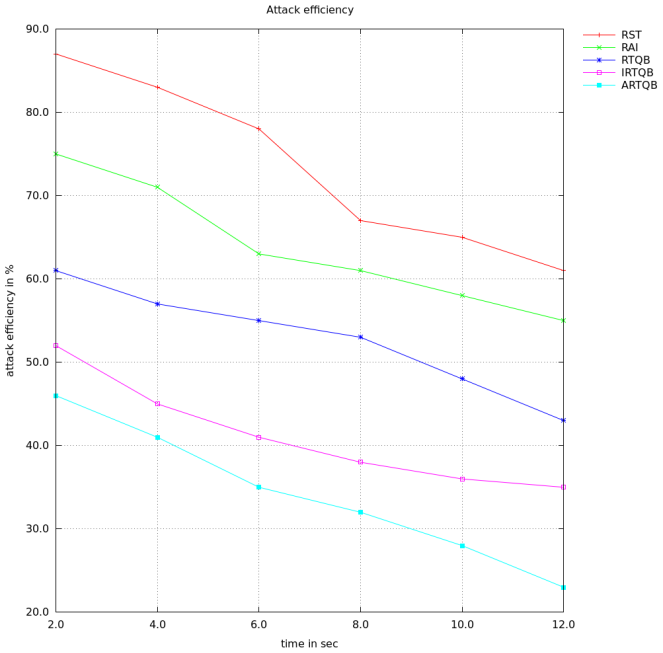


Figure 12: Scenario 3 attack efficiency (%)

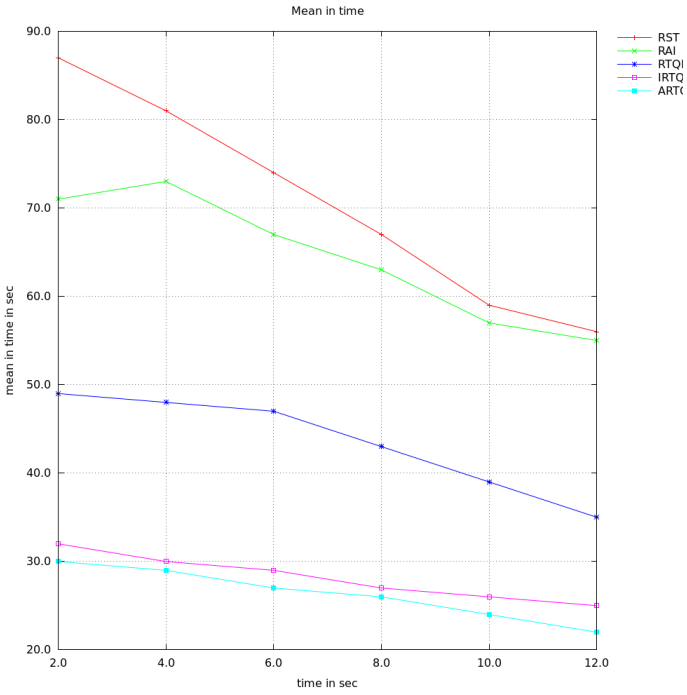


Figure 13: Scenario 1 mean in-system time (s)

The mean in-system time of ARTQB has been decreased due to considering the bandwidth utilization. It shows that when time period is 2sec, the mean in-system time of ARTQB is 41sec compared with other defense mechanisms. Moreover, it is clear that the 56sec mean in-system time in RST decreases to 25sec for ARTQB whereas the mean in-system time of other mechanisms such as RAI, RTQB and IRTQB are 55sec, 35sec and 29sec in the time period of 12sec.

Figure 15 shows that the mean in-system time comparison of RST, RAI, RTQB, IRTQB and ARTQB during Scenario 3 (multi-threaded with variance 0.2). In Scenario 3, the mean in-system time of ARTQB is much reduced than the RST technique by considering the bandwidth utilization. It shows that when the time period is 2sec, the mean in-system time of ARTQB is 40sec compared to the other mechanisms. Also, it is clear that the 38sec mean in-system time of RST decreases to 23sec for ARTQB whereas the mean in-system time of other mechanisms such as RAI, RTQB and IRTQB are 34.1sec, 28sec and 27sec in the time period of 12sec.

From the overall results the major research outcomes are that the RST is very simple defense yet not the best method as the attack efficiency is high and also there is impact of attack in the server behavior. When the extra delay is smaller, RTQB and IRTQB performs better but for higher extra delay RAI outperforms the other methods. However RAI also does not reduce the attack impact on server especially during the extra delay.

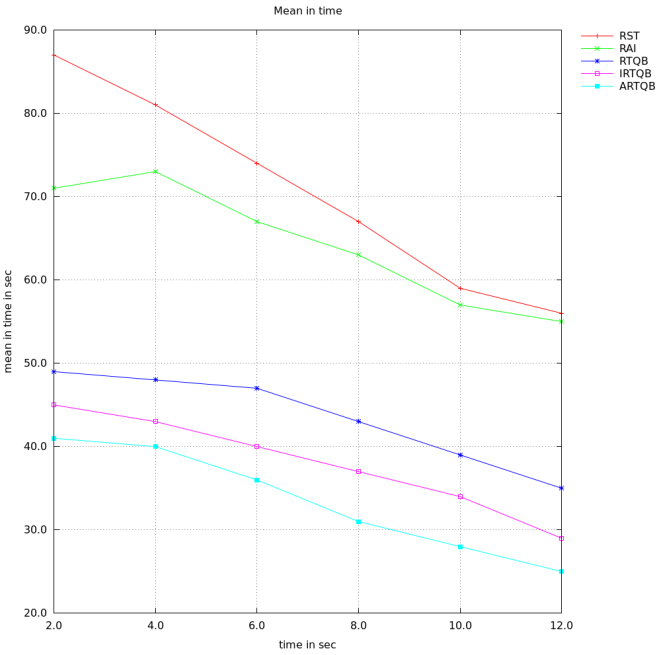


Figure 14: Scenario 2 mean in-system time (s)

Even RTQB and IRTQB have limitations. Considering these aspects the proposed ARTQB uses SSM with bandwidth utilization reducing the attack efficiency and the impact on the server more than half of the initial efficiency without further increasing the cost.

5 Conclusion

Detection of the low-rate DoS attacks is very important in ensuring the behavior of the application servers. Though

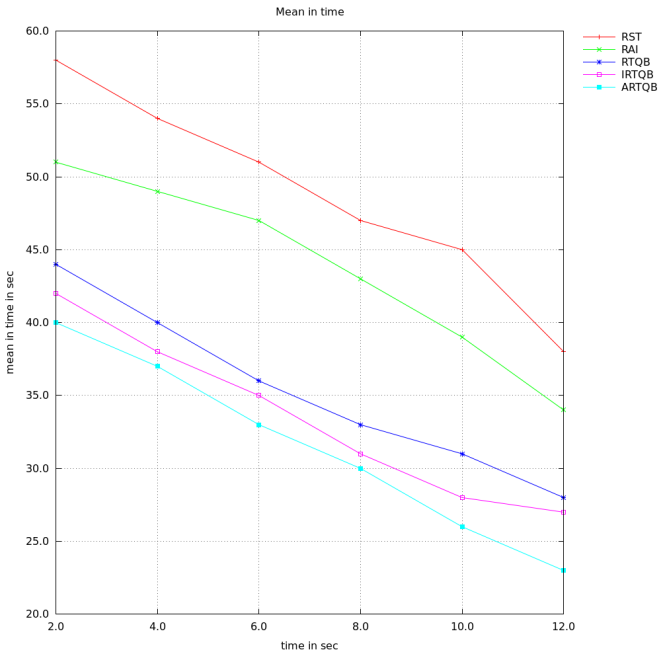


Figure 15: Scenario 3 mean in-system time (s)

the methods such as RST, RAI, RTQB and IRTQB reduces the effect of low-rate DoS attack without much impact on the server behavior, still there are limitations. Hence in this paper, ARTQB is proposed with the aim of maximal reduction of the attack efficiency on the server and minimizing the impact on server behavior. ARTQB selectively chooses the requests during answer instants. Similarly the use of SSM with the bandwidth utilization along with considering source IP addresses and the record timestamp enhances the reduction of attack efficiency. Experimental results conclude that the proposed ARTQB reduces the attack efficiency below half without any impact on the application server behavior.

References

[1] E. Adi, Z. Baig, C. P. Lam, and P. Hingston, "Low-rate denial-of-service attacks against HTTP/2 services," in *5th IEEE International Conference on IT Convergence and Security (ICITCS'15)*, pp. 1–5, 2015.

[2] A. Ain, M. H. Bhuyan, D. K. Bhattacharyya and J. K. Kalita, "Rank correlation for low-rate DDoS attack detection: An empirical evaluation," *International Journal of Network Security*, vol. 18, no. 3, pp. 474–480, 2016.

[3] H. Bedi, S. Roy, and S. Shiva, "Mitigating congestion based DoS attacks with an enhanced AQM technique," *Computer Communications*, vol. 56, pp. 60–73, 2015.

[4] J. Brynielsson, and R. Sharma, "Detectability of low-rate HTTP server DoS attacks using spectral analysis," in *Proceedings of IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, pp. 954–961, 2015.

[5] G. Maciá-Fernández, J. E. Díaz-Verdejo, and P. García-Teodoro, "Evaluation of a low-rate DoS attack against iterative servers," *Computer Networks*, vol. 51, no. 4, pp. 1013–1030, 2007.

[6] G. Maciá-Fernández, J. E. Díaz-Verdejo, and P. García-Teodoro, "Evaluation of a low-rate DoS attack against application servers," *Computers & Security*, vol. 27, no. 7, pp. 335–354, 2008.

[7] G. Maciá-Fernández, J. E. Díaz-Verdejo, and P. García-Teodoro, "Mathematical model for low-rate DoS attacks against application servers," *IEEE Transactions on Information Forensics and Security*, vol. 4, no. 3, pp. 519–529, 2009.

[8] G. Maciá-Fernández, Rafael A. Rodr?guez-G?mez, and J. E. Díaz-Verdejo, "Defense techniques for low-rate DoS attacks against application servers," *Computer Networks*, vol. 54, no. 15, pp. 2711–2727, 2010.

[9] J. Mirkovic, and P. Reiher, "A taxonomy of DDoS attack and DDoS defense mechanisms," *ACM SIGCOMM Computer Communication Review*, vol. 34, no. 2, pp. 39–53, 2004.

[10] A. Shevtekar, J. Stille, and N. Ansari, "On the impacts of low rate DoS attacks on VoIP traffic," *Security and Communication Networks*, vol. 1, no. 1, pp. 45–56, 2008.

[11] Y. Tang, X. Luo, H. Qing, and R. K. Chang, "Modeling the vulnerability of feedback-control based internet services to low-rate DoS attacks," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 3, pp. 339–353, 2014.

[12] Y. Wang, C. Lin, Q. L. Li, and Y. Fang, "A queueing analysis for the denial of service (DoS) attacks in computer networks," *Computer Networks*, vol. 51, no. 12, pp. 3564–3573, 2007.

[13] Z. J. Wu, L. Zhang, and M. Yue, "Low-rate DoS attacks detection based on network multifractal," *IEEE Transactions on Dependable and Secure Computing*, vol. 13, no. 5, pp. 559–567, 2016.

Biography

R. Kavitha is a Ph.D. research scholar of Avinashilingam Institute for Home Science and Higher Education for women, currently doing research on cyber security. She has 10 years of teaching experience in Sri Krishna Arts and Science College; Coimbatore. Her areas of interest include Low Rate Denial of Service Attack

Dr. G. Padmavathi is the Professor and Head of computer science of Avinashilingam Institute for Home Science and Higher Education for women, Coimbatore. She has 23 years of teaching experience and one year of industrial experience. Her areas of interest include Real Time Communication, Network Security and Cryptography. She has 200 publications in her research area.

Presently she is guiding M.phil researcher and PhD's Scholar. She has been profiled in various Organizations her academic contributions. She is currently the principal investigator of four projects funded by UGC and DRDO. She is the scientific mentor for one project funded by DST. She is life member of many preferred organizations of CSI, ISTE, WSEAS, AACE, and ACRS.

Analysis and Optimization of System Intrusion Tolerance Capacity Based on Markov

Zhiyong Luo[1,2], Bo You[2], Peng Wang[1], Jie Su[1], and Yi Liang[2]
(Corresponding author: Zhiyong Luo)
School of Computer Science and Technology , Harbin University of Science and Technology[1]
Harbin 150080, China
School of Mechanical Engineering, Harbin University of Science and Technology[2]
(Email: luozhiyongemail@sina.com)
(Received June 17, 2016; revised and accepted Aug. 20 & Sept. 3, 2016)

Abstract

After the occurrence of network intrusion, the system is running in a state of the lower quality. Along with the system's tolerant capacity decline, it eventually stops providing services or even shutdown. This paper developed a Markov intrusion tolerance model (SMP), aiming at difficultly evaluates and enhances the system's tolerant capacity issues. Based on formalized related security state of the model, the quantitative analysis of system's tolerant capacity is performed. Then calling the parameters solution algorithm to calculate the SMP model's average time of system fault (ATOSF) under each security state. After analyzing the variety track of ATOSF, found the system's tolerant key points. Maintenance of these key points, it can enhance the system's tolerant capacity, so as to increasing the availability of the system. The experiment results provide evidence that using the Markov to the system's tolerant capacity in the quantization process is feasible and effective.

Keywords: Capacity Analysis; Intrusion Tolerance; Markov; SMP; State Transition

1 Introduction

People use the network resources facilitate because of the Internet openness, but it also brings a lot of security threat [7]. Early network security technology focuses on solving two problems contains block the way to the invasion and repair the system security vulnerabilities. Intrusions and system vulnerabilities have unpredictability [5]. Therefore, it is impossible to repair the system in advance of all security vulnerabilities. It will certainly lead to the success of network intrusion. Researchers need to develop a mechanism to guarantee the system operating correctly under state invasion. Researchers need to develop a mechanism to guarantee the system operating correctly under state invasion, and it is called Intrusion Tolerance Technology.

In 1985, Fraga and Powell [3] have proposed intrusion tolerance technology. However, in recent years, it develops under the impetus MAFTIA and OASIS projects. It is as the current network security core technology, Intrusion tolerance technology allows weaknesses in the system. With the operation of the system, these weaknesses are likely to be captured intruder and use. Finally, it will be successful invasion. Intrusion tolerance technology is in this case to ensure the system's key features and essential services (allow degraded model) continues to run. In the study of intrusion tolerance system, we use the security attributes quantitative analysis method to accurately predict the performance of the system, and we find the weaknesses and the existence of critical points system. We raise key intrusion tolerance and achieve the purpose of increasing the system running security and long.

Currently, we use quantitative methods to analyze the intrusion tolerance system has been taken seriously by scholars. Abroad, Madan [6] uses a quantitative method SMP model to analyze the intrusion tolerance system. Denning et al. [2] built an intrusion tolerant system, which improves the tolerance by establishing the steady-state probability for each state node. Ilgun et al. [4] had established the state analysis rules of intrusion tolerance system, which can effectively improve the system tolerance. In China, Jia et al. [12] establish an intrusion tolerance public key encryption scheme in a standard model and use a probabilistic analysis for quantitative analysis model. Chen et al. [1] use the theory of Markov to quantitative analysis database security in intrusion tolerant systems, which ensure the safe operation of the database server. Xing et al. [11] putted forward the calculation method of a kind of measurement tolerance, in the case of attack is inevitable, as long as not beyond measure, the system should provide effective services to legitimate users. Through the evaluation of the results of simulation experiment, for different attack tolerance strategy to provide effective help. Wei et al. [10] built tolerating invasion

ability model to obtain stability probability of model in integrity status using the Markov chain, and they constructed multi-term index of tolerating invasion ability, such as network information machine density, integrity, system autonomy and service availability, and carried out quantitative calculation according to influence on network system of invasion and function of tolerating invasion.

In this paper, we add the Intrusion Learning State based on SITAR [9] intrusion tolerant system architecture, and build an optimized state transition model. Since the state of the model between the transfers meet the transfer of Markov, we use the Markov model to quantitative analysis. It provides a theoretical guidance to build a reliable, confidential and complete tolerance system.

2 Optimized Tolerant System State Transition Model

Tolerant system that protects objects is diversity. The framework, tolerance policy, security algorithms of each tolerant system is different. In order to abstract describe the dynamic behavior of intrusion tolerant system, we build the optimized SITAR model, and its structure is shown in Figure 1.

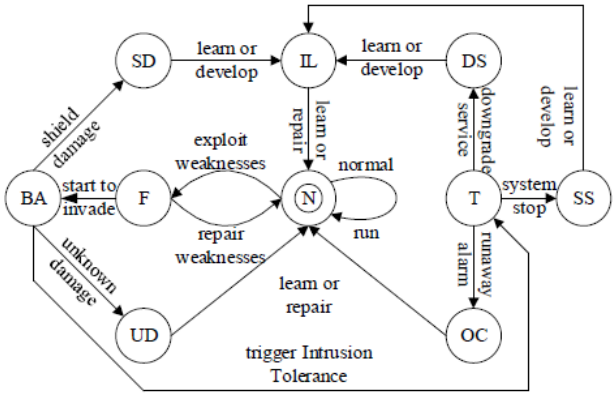


Figure 1: Tolerant system state transition model

In Figure 1, the system in the first State N (Normal State) is normal operation. The intruders detect weaknesses in the system and use these weaknesses. The system will enter the F state (Fragile State) and still run. If the system detects its own weakness in F state and is repaired successfully, the system will return to State N. If the intruders successfully exploit the system weaknesses to start to invade the system, the system will enter State BA (Being Attacked State). If the system which is under the invasion can mask the relative harm, the system will enter State SD (Shield Damage State). The system will wait to learn or improve and enters IL (Intrusion Learning State). If the system which is under the invasion cannot mask the relative harm and the tolerant system is not triggered, the intruders will make some damage to the system. The system will enter State UD (Unknown Damage State). The system will wait the improvement or repair of the administrators and return to State N. If the system which is under the invasion cannot mask the relative harm and the tolerant system is triggered, the system will enter State T (Trigger State). The system runs in State T, the tolerant system will evaluate the system's current status.

After the assessment, if we take the lower level of strategy to continue running, the system will enter State DS (Downgrade Service Status). The system will wait to learn or improve and enters IL. After the assessment, if we take the safe stop strategy, the system will enter State SS (Security Stopped States). The system will wait to learn or improve and enters IL. After the assessment, if the system is completely out of control, the system will enter State OC (Out of Control State). The system will wait the improvement or repair of the administrators and return to State N. The system is in the IL state and completes the learning, improvement or perfect. The system will return State N and start again.

After analyzing, we find that the model in each state node has certain closeness. The conversion in each state node does not affect the previous state. The characteristics of the model meet the Markov process SMP (Markov Process). Therefore, we can use Markov to analyze the model. Additionally, the model contains a number of states. Each state can take the appropriate security policy and make sure the system run normally. Therefore, the model also has a certain degree of flexibility and security.

3 Build the SMP Probability Model

We assume that the duration of each state of the node is random and has an arbitrary distributed. The conversion in each state node does not have the memory. Meanwhile, in order to simplify the analysis, we will all pay the cost of the attack are considered time cost.

3.1 DTMC Matrix

DTMC (Discrete-time Markov Chain) matrix is that we use discrete time values and combine the Markov chain technology to make each node of the state space in the process of the Markov. Each node transition probabilities in a state space form a matrix, which is called the DTMC Matrix. We analyze the Optimized SMP model and get the state space of the system, which is called $S_{space} = \{N, F, BA, SD, UD, T, DS, SS, OC, IL\}$. In addition, P_{s_l}, P_d, $P_{s_i}, P_w, P_a, P_s, P_u, P_d, P_h, P_1, P_2, P_3, P_n$, P_{on}, P_{un}, P_{in} represent a conversion between the probability of each state. The SMP probabilistic model is shown in Figure 2.

In Figure 2, the meanings of the probability symbols are as follows:

$P_{s_l} : SD \to IL$, it is a probabilistic that the system shields the intrusion harms but needs to learn or develop.

$P_{d_i} : DS \to IL$, it is a probabilistic that the system can provide the downgrade service but needs to learn or develop.

$P_{s_i} : SS \to IL$, it is a probabilistic that the system is safe to stop running and needs to learn or develop.

$P_w : N \to F$, it is a probabilistic that the system has weak and is found.

$P_a : F \to BA$, it is a probabilistic that the system successfully exploits vulnerabilities invasion.

$P_s : BA \to SD$, it is a probabilistic that the system successfully shields the intrusion.

$P_u : BA \to UD$, it is a probabilistic that the system cannot find the intrusion.

$P_d : T \to DS$, it is a probabilistic that the system finds the intrusion and provides the downgrade service.

$P_h : T \to SS$, it is a probabilistic that the system finds the intrusion and is to stop running successfully.

$P_1 = 1 - P_w - P_a : F \to N$, it is a probabilistic that the system finds the weak and repair successfully.

$P_2 = 1 - P_s - P_u : BA \to T$, it is a probabilistic that the system detects the presence of the invasion and successfully triggers intrusion tolerant systems.

$P_3 = 1 - P_d - P_h : T \to OC$, it is a probabilistic that the system eventually stops running because of the invasion caused the fault occurs.

$P_n : N \to N$, it is a probabilistic that the system is safe to run.

$P_{on} : OC \to N$, it is a probabilistic that the system is completely out of control, but improved or repaired return to normal operation.

$P_{un} : UD \to N$, it is a probabilistic that the system is not found in the invasion, and it improves or repairs to rerun after a period of time.

$P_{in} : IL \to N$, it is a probabilistic that the system develops and returns to normal after learning, improving or perfecting.

In Figure 2, transition probability matrix P describes the possibility of the system transferring between the various states. The probability value can be determined by the experience of the network management or determined through the intrusion injection mode. The transition probability matrix P of the system state transition model DTMC:

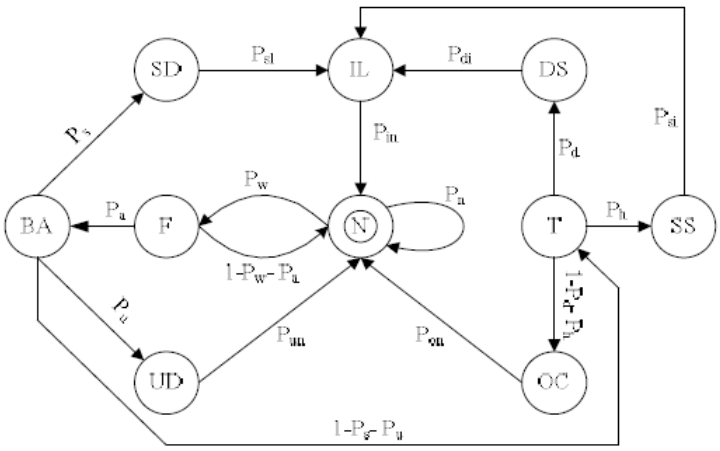


Figure 2: DTMC transition model

$$P = \begin{array}{c} \\ N \\ F \\ BA \\ SD \\ UD \\ T \\ DS \\ SS \\ OC \\ H \end{array} \begin{array}{c} \begin{array}{cccccccccc} N & F & BA & SD & UD & T & DS & SS & OC & H \end{array} \\ \begin{bmatrix} P_n & P_w & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ P_1 & 0 & P_a & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & P_s & P_u & P_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & P_{s_1} \\ P_{un} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & P_d & P_h & P_3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & P_{d_i} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & P_{s_i} \\ P_{on} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ P_{in} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

3.2 DTMC State Duration

The state duration is that each state holds time in SMP model. H represents the duration matrix of the SMP model. h_i represents some a state duration in SMP model, $i \in S_{space}$, $H = [h_N, h_F, h_{BA}, h_{SD}, h_{UD}, h_T, h_{DS}, h_{SS}, h_{OC}, h_{IL}]$. In Figure 2, the meanings of the state duration are as follows:

h_N: It is a time that the system runs normally and its weakness is not found by an intruder.

h_F: It is a time that the system is not successful invasion when the intruder finds the weakness and uses it.

h_{BA}: It is a time that the system finds invasion and triggers a successful intrusion tolerance.

h_{SD}: It is a time that the system shields invasion successfully and runs normally.

h_{UD}: It is a time that the system cannot find invasion and runs normally.

h_T: It is a time that the system is evaluated and decides what the security policies the system can use to deal with the invasion.

h_{DS}: It is a time that the system finds the invasion, but cannot block it and only provide the downgrade service.

h_{SS}: It is a time that the system is safe to stop running.

h_{OC}: It is a time that the system is completely out of control.

h_{IL}: It is a time that we learn the invasion and optimizes the system.

4 SMP Model Analysis

According to the quantitative analysis of the SMP model analysis, the network administrator can better safeguard system. In order to quantify the simplified SMP model and accurately, this paper gives the following definition.

Definition 1. *DTMC state probability, the probability of each state of the system into SMP model. In this paper, we use matrix to express V. So $V = [v_N, v_F, v_{BA}, v_{SD}, v_{UD}, v_T, v_{DS}, v_{SS}, v_{OC}, v_{IL}]$, v_i expresses State i DTMC state probability, $i \in S_{space}$.*

Definition 2. *SMP stability probabilities, the system in the state SMP model and continue to stay in the percentage of the whole model of duration. It is expressed by π, so π_i expresses State i SMP stability probability.*

Definition 3. *Average Time of System Fault (ATOSF), The system starts from a state of implementation of the SMP model in the reach system to stop running state of the average length of time due to failure caused by the invasion.*

4.1 SMP Model Security

Attributes SMP model security attributes mainly consider three main aspects. It each expresses the availability, confidentiality and integrity.

The availability expresses that the model can provide services for legitimate users, the probability of occurrence is expressed by P_{Ava}.

The confidentiality expresses that the model cannot be stolen the data by the intruder, the probability of occurrence is expressed by P_{Con}.

The integrity expresses that the model is not modified by the intruder; the probability of occurrence is expressed by P_{Int}.

The state space of node S_{space} as shown in Figure 2 is divided into two subsets: The invaders invasion behavior node space S_I and The response behavior of the system adopted in the post invasion node space S_R, $S_I = \{N, F, BA\}$; $S_R = \{SD, UD, T, DS, SS, OC, IL\}$. In State S_R, the system is in a certain state, SMP model security attributes will be lost, the state format safety damaged space S_D. Contrary, the system is in another certain state, SMP model security attributes will not be lost, and the state format safety not damaged space S_U.

Through the analysis, the probability of SMP model security attributes is related to the stability probability of each state node space S_{space}. If we use π to express the stability probability, π_i expresses each state node space S_{space}, $i \in S_{space}$.

The system in the UD, SS, OC state stops running and does not provide any service. The safety damaged space of the system $SD = \{UD, SS, OC\}$, safety not damaged space $SU = \{SD, T, DS, IL\}$. At this time, the availability probability $P_{Ava} = 1 - \pi_{UD} - \pi_{SS} - \pi_{OC}$.

The system is in the UD, OC state, when the intruder attacks the server and makes the system in unsafely stopping state. The system will make the data stolen. The safety damaged space of the system $SD = \{UD, OC\}$, The safety not damaged space of the system $SU = \{SD, T, DS, SS, IL\}$. At this time, the confidentiality probability $P_{Con} = 1 - \pi_{UD} - \pi_{OC}$.

The system in the UD, SS, OC state, the integrity will be damaged by the intruder. The safety damaged space of the system $SD = \{UD, SS, OC\}$, the safety not damaged space of the system $SU = \{SD, T, DS, IL\}$. At this time, the integrity probability $P_{Int} = 1 - \pi_{UD} - \pi_{SS} - \pi_{OC}$. Therefore, the probability of SMP model security attributes:

$$P_k = 1 - \sum_{j \in S_D} \pi_j, \quad k = Ava, Con, Int \tag{1}$$

SMP model security attributes is inversely proportional to the probability of the safety damaged space π_j.

4.2 SMP Model Parameters Algorithm

The stability probability of SMP model each state mainly consider two input parameters: each state DTMC transition probability matrix P and each state DTMC duration matrix H. Through the above analysis, SMP model parameters algorithm is shown.

Step 1: Through Equation (2), we calculate Figure 2 each state DTMC probability matrix V.

$$\begin{cases} V' & = V \cdot P \\ \sum_{i \in S_{space}} v_i & = 1 \end{cases} \tag{2}$$

The matrix P is DTMC transition probability matrix. Through Equation (2), we can calculate the relationship of each state's DTMC probability in SMP model, which is shown in Equation (3).

$$\begin{aligned} v_N &= P_n v_N + P_1 v_F + P_{un} v_{UD} + P_{on} v_{OC} + P_{in} v_{IL} \\ &= [P_n + P_1 P_w + P_{un} P_u P_a P_w + P_{on} P_3 P_2 P_a P_w \\ &\quad + P_{in}(P_{sl} P_s P_a P_w + P_{d_i} P_d P_2 P_a P_w \\ &\quad + P_{s_i} P_h P_2 P_a P_w)] v_N \\ v_F &= P_w v_N \\ v_{BA} &= P_a v_F = P_a P_w v_N \end{aligned}$$

$$\begin{aligned} v_{SD} &= P_s v_{BA} = P_s P_a P_w v_N \\ v_{UD} &= P_u v_{BA} = P_u P_a P_W v_N \\ v_T &= P_2 v_{BA} = P_2 P_a P_w v_N \\ v_{DS} &= P_d v_T = P_d P_2 P_a P_w v_N \\ v_{SS} &= P_h v_T = P_h P_2 P_a P_w v_N \\ v_{OC} &= P_3 v_T = P_3 P_2 P_a P_w v_N \\ v_{IL} &= P_{sl} v_{SD} + P_{d_i} v_{DS} + P_{s_i} v_{SS} \\ &= (P_{sl} P_s P_a P_w + P_{d_i} P_d P_2 P_a P_w \\ &\quad + P_{s_i} P_h P_2 P_a P_w) v \end{aligned} \tag{3}$$

Through $\sum_{i \in S_{Space}} v_i = 1$, we can get $v_N + v_F + v_{BA} + v_{SD} + v_{UD} + v_T + v_{DS} + v_{SS} + v_{OC} + v_{IL} = 1$. We put it into Equation (3) and get the DTMC probability of State N:

$$\begin{aligned} v_N &= 1/[P_n + P_1 P_w + PA + PB + PC] v_N \\ PA &= P_{un} P_u P_a P_w \\ PB &= P_{on} P_3 P_2 P_a P_w \\ PC &= P_{in}(P_{sl} P_s P_a P_w + P_{d_i} P_d P_2 P_a P_w \\ &\quad + P_{s_i} P_h P_2 P_a P_w) \end{aligned} \tag{4}$$

We put Equation (4) into Equation (3) and get each state DTMC probability matrix V of the SMP model.

Step 2: We put the DTMC probability matrix V and duration matrix H into Equation (5).

$$\pi_i = fracv_i h_i \sum_j v_j h_j, \quad i, j \in S_{Space} \tag{5}$$

We calculate the SMP model each state stability probability as shown:

$$\begin{aligned} Sum &= \sum_{j \in S_{space}} v_j h_j \\ &= [h_N + h_F P_w + h_{BA} P_a P_w + h_{SD} P_s P_a P_w \\ &\quad + h_{UD} P_u P_a P_w + h_T P_2 P_a P_w \\ &\quad + h_{DS} P_d P_2 P_a P_w + h_{IL}(P_{sl} P_s P_a P_w \\ &\quad + P_{d_i} P_d P_2 P_a P_w + P_{s_i} P_h P_2 P_a P_w)] v_N \\ \pi_N &= h_N v_N / Sum \\ \pi_F &= h_F P_w v_N / Sum \\ \pi_{BA} &= h_{BA} P_a P_w v_N / Sum \\ \pi_{SD} &= h_{SD} P_s P_a P_w v_N / Sum \\ \pi_{UD} &= h_{UD} P_u P_a P_w v_N / Sum \\ \pi_T &= h_T P_2 P_a P_w v_N / Sum \\ \pi_{DS} &= h_{DS} P_d P_2 P_a P_w v_N / Sum \\ \pi_{SS} &= h_{SS} P_h P_2 P_a P_w v_N / Sum \\ \pi_{OC} &= h_{OC} P_3 P_2 P_a P_w v_N / Sum \\ \pi_{IL} &= h_{IL}(P_{sl} P_s P_a P_w + P_{d_i} P_d P_2 P_a P_w \\ &\quad + P_{s_i} P_h P_2 P_a P_w) v_N / Sum \end{aligned} \tag{6}$$

Step 3: We put the SMP model state stability probability into Equation (1) and get the security attributes probability:

$$\begin{aligned} P_{Ava} &= 1 - (h_{UD} P_u P_a P_w + h_{SS} P_h P_2 P_a P_w \\ &\quad + h_{OC} P_3 P_2 P_a P_w) v_N / Sum \\ P_{Con} &= 1 - (h_{UD} P_u P_a P_w \\ &\quad + h_{OC} P_3 P_2 P_a P_w) v_N / Sum \\ P_{Int} &= 1 - (h_{UD} P_u P_a P_w + h_{SS} P_h P_2 P_a P_w \\ &\quad + h_{OC} P_3 P_2 P_a P_w) v_N / Sum \end{aligned} \tag{7}$$

According to SMP model parameters algorithm, we combine the specific parameters of the network intrusion tolerance system and can accurately quantify tolerance system and provide data basis for the future analysis.

4.3 SMP State Average Fault Time

According to Definition 3, the average fault time is a measure of an important indicator of intrusion tolerance system resistance ability. SMP model some a state average fault time is bigger and it expresses the invasion of the state of the system to stop running time is long, the cost is high, the reliability of the system is also higher. We analyze the SMP model in Figure 2 and find some state is stopping running state that the system has some problems. We repair the weakness of the system or develop the system in this state by the mode administrator manual to make the system run again. We set a model in such a state is called a stop state set SE. The collection of the rest of the state is called intermediate state set SM. According to the Trivedi algorithm [8], we can get the ATOSF:

$$ATOSF = \sum_{i \in S_M} Con_i h_i \tag{8}$$

In Equation (8), Con_i expresses total number of the system in the stop State i. h_i is the duration time of State i. The system always starts in State N and the Con_N is the key. Through Figure 2, through State N probability is divided into inflow probability P_{in} and outflow probability P_{out}, $P_{in} + P_{out} = 1$. Because Con_N is the total number that the system through State N before entering the state stopped, so $Con_N = 1/P_{in} = 1/(1 - P_{out})$. In SMP model, the factor of the effect State N coming into probability P_{in} is a problem. The paper calculates the inflow probability P_{out} to ensure Con_N.

Through the analysis, the system in State N entering the stopping state has five paths: $N \to F \to BA \to SD \to IL$, $N \to F \to BA \to UD$, $N \to F \to BA \to T \to OC$, $N \to F \to BA \to T \to DS \to IL$ and $N \to F \to BA \to T \to DS \to IL$. We analysis the five path and find the system in through State BA enters the stopping state and the stopping state enters State N again by the administrator. The outflow probability of State N $P_{out} = P_a P_w$. We get the SMP model each State Con and the

Table 1: Software configuration of network server

Server ID	Operation system	Provide services	Weaknesses ID
IP_1	Windows 2003 Server	FTP Server	CVE-2004-0575 CVE-2008-0702
IP_2	Windows 2003 Server	HTTP Server	CVE-2002-0364 CVE-2006-2379
IP_3	Windows 2000 Server	SQL Server	CVE-2007-0038 CVE-2004-0893

system ATOSF.

$$
\begin{aligned}
Con_N &= 1/(1-P_aP_w) \\
Con_F &= P_wCon_N \\
Con_{BA} &= P_aP_wCon_N \\
Con_{SD} &= P_sP_aP_wCon_N \\
Con_{UD} &= P_uP_aP_wCon_N \\
Con_T &= P_2P_aP_wCon_N \\
Con_{DS} &= P_dP_2P_aP_wCon_N \\
Con_{SS} &= P_hP_2P_aP_wCon_N \\
Con_{OC} &= P_3P_2P_aP_wCon_N \\
Con_{IL} &= (P_sP_aP_w + P_dP_2P_aP_w + P_hP_2P_aP_w)Con_N \\
ATOSF &= \sum_{i \in S_M} Con_i h_i \qquad (9)
\end{aligned}
$$

ATOSF is an important index that makes the system safe and reliable. We enlarge the ATOSF to can increase the attack price. However, the ATOSF is related to the Con_i and h_i. With the fixed intrusion tolerance system, each State Con can be ensured. We can enlarge the state duration time h to make the ATOSF.

5 Experiment Analyses and Evaluation

5.1 Experimental Environment

The topological structure of the network of the SMP model is shown in Figure 3. The server IP_1 to IP_3 forms an intrusion tolerance system in the control strategy of firewall to provide the corresponding network service with the host of users inside and outside the network. The software configuration and the weakness of its specific are shown in Table 1. The author organizes the student to simulate the intrusion tolerance system, so as to obtain test data.

5.2 Experiment Analysis and Evaluation

Through the analysis of the test data and the estimation of the statistics, the following parameters values are shown: DTMC transition probability matrix P. The system is always in normal operation. So $P_n = 1$. However, when the system is managed and run again. So $P_{s_l} = P_{s_i} = P_{d_i} = P_{in} = P_{un} = P_{on} = 1$. Each server tolerance system has a lot of weakness in Figure 3, so the weakness of the system is found and its probability is $P_w = 0.3$. When the invaders found weaknesses in the system and successfully exploited these vulnerabilities to invade the tolerance system, its probability is $P_a = 0.5$. The system detects the weakness and timely repair and its probability is $P_1 = 1 - P_w - P_a = 0.2$. The system is found to be invaded and successfully shield the intrusion and its probability is $P_s = 0.4$. The system could not find the intrusion and its probability is $P_u = 0.2$. The system detects intrusion and successful trigger intrusion tolerance system and its probability is $P_2 = 1 - P_s - P_u = 0.4$. The intrusion system continues to run but providing degraded service and its probability is $P_d = 0.5$. The system finds the invasion and succeeds to stop system and its probability is $P_h = 0.4$. The system eventually stops running because of the invasion and its probability is $P_3 = 1 - P_d - P_h = 0.1$.

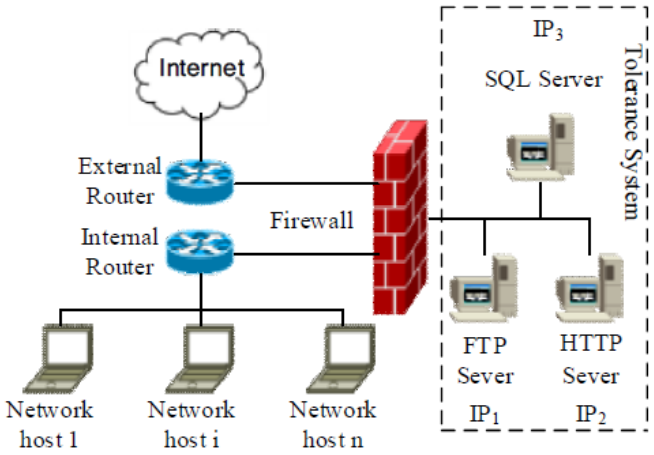


Figure 3: Topology model of the testing network

The duration time matrix H. The tests show that the system degrading service operation has most of the time. The system is the shortest in the tolerance time triggering. The test shows that the degraded service operation time of the system is the longest, the tolerance to trigger time is the shortest, the normal operation time and the no

Table 2: SMP model parameters

State i	DTMC probability v_i	SMP stability probability π_i	The visits number of each state Con_i
N	0.6098	0.6044	1.1765
F	0.1512	0.2698	0.3530
BA	0.0756	0.0299	0.1765
SD	0.0302	0.0150	0.0706
UD	0.0151	0.0150	0.0353
T	0.0302	0.0059	0.0706
DS	0.0151	0.0059	0.0353
SS	0.0121	0.0180	0.0282
OC	0.0030	0.0074	0.0071
IL	0.0575	0.0285	0.1341

found invasion to continue running time are similar, the shielding the intrusion behavior to continue running time and the learn and improve time are similar, the time of the other states are not equal. In this paper, we use the unit time measurement, and the duration of each state is set to: $h_N = 1, h_F = 1.8, h_{BA} = 0.4, h_{SD} = 0.5, h_{UD} = 1, h_T = 0.2, h_{DS} = 4, h_{SS} = 1.5, h_{OC} = 2.5, h_{IL} = 0.5$.

We put all the parameters into SMP model parameters algorithm Equation (9), and calculate the SMP related parameters the SMP parameters are shown in Table 2 We put Table 2 into Equation (7) and Equation (8). We can get the system availability probability $P_{Ava} = 0.9596$, the confidentiality probability $P_{Con} = 0.9776$, the integrity probability $P_{Int} = 0.9596$, the total probability $ATOSF = 2.2355$. We make the further analysis with Equation (8) and get SMP model each state ATOSF, which is shown in Figure 4.

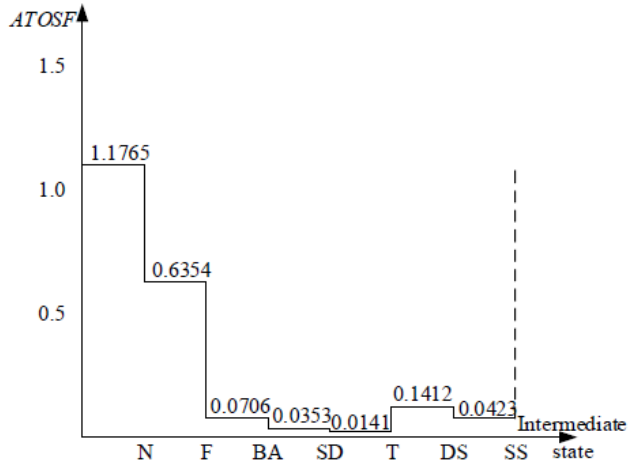


Figure 4: Change trajectory of each intermediate state's ATOSF

As can be seen from Figure 4, the duration of the intermediate state of the overall system ATOSF from big to small order: $\{N, F, DS, BA, SS, SD, T\}$. If we enlarge the intermediate state $\{N, F, DS\}$ duration time, we can get the effectively increasing the system ATOSF. At the same time, it also increases the cost of the invasion and enhances the reliability of the system.

6 Conclusion

Intrusion tolerance technology is an important technology of network security management. It is a kind of technology to ensure the operation of the network after the intrusion happened. So the research on the intrusion tolerance is a hotspot. This paper is based on the SITAR intrusion tolerance system structure and increases the attack state and puts forward to optimize the state transfer model. Due to the conversion of the state of the model meeting the semi Markov theory, the system introduces the DTMC to construct the optimized SMP model. Through the quantitative analysis of the model, we calculate the ATOSF locus of the model each state.

Finally, through the analysis of the test data, we can get the conclusion that enlarge model intermediate state $\{N, F, DS\}$ duration time to add the difficulty of intrusion. The next step for the research will be further improved the system. We increase the tolerance of online to repair system, reduce the system stop state, and improve system availability.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (Grant No.61403109). Declares: The authors declare that there is no conflict of interest regarding the publication of this manuscript.

References

[1] C. Q. Chen, X. B. Pei, H. Zhou, Y. S. Liu, "A Markov evaluation model for the survivability of real-time database with intrusion tolerance," *Chinese Journal of Computer*, vol. 34, no. 10, pp. 1907–1915, 2011.

[2] D. E. Denning, "An intrusion-detection model," *IEEE Transaction on Software Engineering*, vol. 13, no. 2, pp. 222–223, 2011.
[3] J. S. Fraga, D. Powell, "A fault and intrusion to lerant file system," in *Proceedings of the 3rd International Conference on Computer Security*, Dublin, Ireland, pp. 203–218, 1985.
[4] I. Koral, A. K. Richard, "State transition analysis:a rule-based intrusion detection approach," *IEEE Transactions on Software Engineering*, vol. 21, no. 3, pp. 181–199, 2012.
[5] Z. Luo, B. You, G. Yu, J. Su, "Research of intrusive intention self-recognition algorithm based on three-tier attack graph," *ICIC Express Letters, Part B: Applications*, vol. 6, no. 6, pp. 1575–1580, 2015.
[6] B. B. Madan et al., "A method for modeling and quantifying the security attributes of intrusion tolerant systems," *Performance Evaluation*, vol. 56, no. 14, pp. 167–186, 2004.
[7] J. Su, S. Liu, Z. Y. Luo, G. L. Sun, "Method of constructing an anonymous graph based on information loss estimation," *Tongxin Xuebao/Journal on Communications*, vol. 37, no. 6, pp. 56–64, 2016.
[8] K. S. Trivedi, *Probability and Statistics with Reliability Queuing, and Computer Science Applications*, 2nd Edition, New York: John Wiley and Sons, 2002.
[9] G. Wang, P. Wang, Z. Luo, S. Zhu, "Transfer model based on state of finite semi-markov automata intrusion tolerance," *International Journal of Security and Its Applications*, vol. 10, no. 3, pp. 183–192, 2016.
[10] K. Wei, F. Zhang, "Based on Markov network tolerate invasion ability evaluation model," *Computer Simulation*, vol. 33, no. 7, pp. 289–292, 2016.
[11] Y. F. Xing, C. Y. Luan, "A quantitative analysis and detection of intrusion tolerance system model," *Information Science*, vol. 33, no. 8, pp. 55–58, 2015.
[12] J. Yu, X. G. Cheng, F. G. Li, Z. K. Pan, F. Y. Kong, R. Hao, "Provably secure intrusion-resilient public-key encryption scheme in the standard model," *Journal of Software*, vol. 24, no. 2, pp. 266–278, 2013.

Biography

Luo Zhiyong, born in 1978, master tutor, associate professor, is currently a PhD candidate at Intelligent Machine Institute, Harbin University of Science and Technology, China. He received his bachelor degree from Harbin University of Science and Technology, China, in 2001. His research interests include network security, scientific workflow and industrial design and scheduling.

You Bo, born in 1962, doctoral tutor, professor, post doctorate, is currently working at Intelligent Machine Institute, Harbin University of Science and Technology, China.

Wang Peng, born in 1993, master, is currently studying at Harbin University of Science and Technology, China.

Su Jie, born in 1979, master tutor, associate professor, is currently working at Harbin University of Science and Technology, China.

Liang Yi, born in 1985, is currently studying at Harbin University of Science and Technology, China.

Attribute-based Encryption with Non-Monotonic Access Structures Supporting Fine-grained Attribute Revocation in M-healthcare

Yang Zhao, Pengcheng Fan, Haoting Cai, Zhiguang Qin and Hu Xiong

(Corresponding author: Hu Xiong)

School of Information and Software Engineering, University of Electronic Science and Technology of China[1]

No. 4, North Jianshe Road, Chenghua District, chengdu, Sichuan 610054, China

(Email: xionghu.uestc@gmail.com)

(Received July 21, 2016; revised and accepted Jan. 15, 2017)

Abstract

By sharing the personal health information (PHI) in the healthcare provider (HP) which is equipped with cloud servers, mobile-healthcare (m-healthcare) significantly promotes a huge revolution of medical consultation. Nonetheless there is a series of challenges such as PHI confidentiality and the attribute revocation. To deal with these problems, we propose a scheme based on the attribute-based encryption. The scheme which supports non-monotonic access structures and fine-grained attribute revocation is established over the composite order bilinear groups. By utilizing this scheme, we can well protect PHI and achieve the goal of revocation. Furthermore, the security analysis and comparison show that our scheme is more expressive despite of the lower efficiency.

Keywords: Attribute-based Encryption; Attribute Revocation; M-healthcare System; Non-monotonic Access Structures

1 Introduction

M-healthcare cloud computing system has been spread widely all over the world. Due to its high efficiency and accessibility for medical consultation, it has been increasingly adopted by world-renowned organizations such as the European Commission activities. Both patients and HP greatly benefit from its great convenience [11, 24].

M-healthcare cloud computing system can be considered as a huge social network. PHI collected by body area networks (BANs) should be securely transmitted to HP and shared among the authorized physicians. The authorized physicians may access the PHI to accomplish medical treatment [27, 36, 40].

In such situation, many issues should be considered, especially preventing the patients' PHI from being eavesdropped and tampered, and having the authorities of authorized physicians revoked.

In terms of the security aspects, access control for patients' PHI is one of the most important issues. Namely, only the authorized physicians can recover the patients' PHI . Therefore, how to share the patients' PHI and who should be shared with should be considered carefully. To solve these challenges, there were a variety of achievements [16, 23, 29, 31, 32, 37, 40].

Recently, the scheme [40] is constructed for securing the PHI along with a multi-level model which contains four entities - patient, directly authorized physician, indirectly authorized physician and unauthorized physician. The directly authorized physician can access the patient's PHI. The indirectly physician can only access the data authorized by the directly authorized physician. So the directly authorized physician owns all the privileges of patient and controls the data which the indirectly authorized physician can access. However, if the directly authorized physician is bribed, he can be capable of colluding with the indirectly physicians who do not satisfy the access control. Moreover, the directly authorized physician who is bribed can share the fake information with the indirectly authorized physicians [39]. As a result, this scheme may suffer from collusion attack and forgery attack.

In order to solve the problems mentioned above, we proposed a scheme. In this scheme, PHI confidentiality and the revocation of authorities can be achieved with high flexibility by utilizing the fine-grained attribute revocation and non-monotonic access structures. With the non-monotonic access structures [26], private keys can represent any access structures involving AND, OR, NOT, and threshold operations. To accomplish it, a set of attributes was selected as the universe. Then a set of d attributes was selected from the universe which was used to encrypt the ciphertext and the negation of remaining attributes represented the negated attributes. As for the normal Attribute-based Encryption (ABE) supporting NOT operation over access structures, there is no

choice but to add more negated attributes to the set, like "Not Nurse". Therefore, by utilizing our scheme, we can encrypt the patients' PHI flexibly with a smaller attribute set.

Moreover, we implemented fine-grained attribute revocation [34] in our scheme. By providing a revocation list for every attribute, the scheme supports attribute/user revocation. So access control of PHI confidentiality can be achieved flexibly. By making use of the above methods, the authorities of physicians can be well controlled. Our contributions are outlined below:

1) We propose a new attribute-based encryption with non-monotonic access structures supporting fine-grained attribute revocation. our scheme achieves the goals of PHI confidentiality and the revocation of authorities.

2) For the first time, we bring attribute-based encryption with non-monotonic access structures and fine-grained attribute revocation into m-healthcare cloud computing system, which can flexibly achieve PHI confidentiality.

3) The security analysis is provided in this paper and we compared it with existing works to show the advantages and disadvantages of our scheme.

2 Related Work

2.1 Attribute-based Encryption

In the identity-based encryption (IBE) system, the public key to encrypt the message is the unique identity of user [6]. In order to send the ciphertext to the user, the data owner has to know the user's identity. Biometrics are always considered as the best carrier for user identity. In the Fuzzy IBE (FIBE) [30] which was Sahai and Waters firstly proposed, the identity was characterized as a set of attributes. The ciphertext which was encrypted by the set ω could be decrypted by the attributes set ω' only if $|\omega' \bigcap \omega| \geqslant d$, where d denotes the threshold. In 2006, Goyal, Sahai and Waters et al. [13] expanded the FIBE to ABE. In ABE, the identity information of user was generalized to attributes related to user identity. According to the relation between access structure and ciphertext or private key, ABE was categorized into key policy ABE (KP-ABE) and ciphertext policy ABE (CP-ABE) [3]. Moreover, they also provided a basic security property namely collusion resistance which could prevent adversary from decrypting the ciphertext illegally by cooperation. Then Chase et al. [7] made use of a certificate authority (CA) and multi authorities to prevent single authority from being corrupted. In this scheme they utilized globally unique identifier (GUID) to defend collusion attack.

In 2008, Bethencourt, Sahai and Waters [4] proposed the first CP-ABE scheme supporting tree access structure. The scheme defended collusion attack by using different random numbers for various private keys. In the same year, Cheung and Newport [9] introduced a provably secure CP-ABE based on standard model and the decisional bilinear Diffie-Hellman (DBDH) assumption. However, the scheme only supported AND operation. Considering the adaptive security, partitioning reduction could not be well applied into ABE. Dual system encryption [35] provided a new way to solve this problem. Lewko and Waters [17] constructed the first adaptive security ABE scheme by utilizing the dual system encryption. Due to involving the composite order bilinear maps, the scheme requires an extremely large group order which also results in low efficiency. Subsequently, the first adaptive security ABE based on prime order was proposed by Okamoto and Takashima [25]. The scheme effectively implemented the dual system encryption by using dual paring vector space. Although its efficiency had been greatly improved, there was always existing a gap comparing with selective security model. To improve the efficiency, various constant-size ciphertext ABE schemes [21, 33, 38] were constructed in different ways. Furthermore, in 2015, Gorbunov et al. [12] presented an ABE scheme for circuits of any arbitrary polynomial size, which could be a new framework for constructing ABE schemes.

2.2 Attribute Revocation

Attribute revocation can be classified as direct revocation and indirect revocation. Direct revocation performs revocation directly by the encryptor who establishes and updates the revocation list. Indirect revocation performs revocation indirectly by the private key authority which publishes private keys periodically. In practical scenarios, attribute revocation is conundrum waiting to be addressed [8, 22].

In 2006, Pirretti et al. [28] introduced a scheme to implement indirect revocation by revoking the latest version of users' attributes to achieve the goal. Prior to encryption, encryptor should negotiate with the authority to confirm the validity duration of attributes. Furthermore, users and authority must accomplish key update periodically online. Then Bethencourt et al. [4] put the expiry date of attributes into ciphertext to implement revocation and solved the problem of negotiation between encryptor and authority. Afterward, binary tree was used to revoke user [5] through updating the minimum set of unrevoked users. However, the drawback was that it only supported user revocation.

In 2007, Ostrovsky et al. [26], for the first time, presented a direct revocation scheme based on the ABE. But the size of ciphertext and key was slightly large. To reduce the expense of revocation, Attrapadung et al. [2] proposed a direct revocation scheme on KP-ABE and CP-ABE in conjunction with broadcast encryption. The advantage was that revocation would not influence other users. In 2011, Asim et al. [1] constructed a direct revocation scheme by taking advantage of polynomial to share secret. In this scheme, all revoked users' secret shares were put

into ciphertext, so that only authorized users can get message when decrypting. However the complexity of pairing calculation involved the count of revoked users, which resulted in its low efficiency. Then, [20] were proposed to achieve fine-grained revocation. However, the defect was that only an attribute of user could be revoked in an encryption. Until 2012, the direct revocation [34] was firstly introduced to supporting fully fine-grained attribute revocation by specifying a revocation list for every attribute. Recently, several applications of healthcare [10, 14] based on the fine-grained attribute revocation have been constructed.

3 Preliminaries

3.1 Definitions

Definition 1 (Access Structure). Let $\{P_1, ..., P_n\}$ be a set of parties. A collection $\mathbb{A} \subseteq 2^{\{P1,...,Pn\}}$ is monotonic, for $\forall B$ and C, if $B \subseteq \mathbb{A}$ and $B \subseteq C$ then $C \subseteq \mathbb{A}$. An monotonic access structure is a monotonic collection $\mathbb{A}$ of non-empty subsets of $\{P_1, P_2, ..., P_n\}$, namely, $\mathbb{A} \subseteq 2^{\{P_1,P_2,...,P_n\}} \setminus \{\emptyset\}$. The sets in $\mathbb{A}$ are regarded as authorized sets and the sets not in $\mathbb{A}$ are unauthorized sets.

Definition 2 (Linear Secret-Sharing Schemes [26]). Call a secret-sharing scheme Π which depends on a set of parties $\mathcal{P}$ as linear (over $\mathbb{Z}_p$), if it satisfies the followings.

1) A vector over $\mathbb{Z}_p$ is composed by the shares for each party.

2) The share-generating matrix for Π is the name of A matrix M which consists of $n+1$ columns and l rows. For all $i = 1, ..., l$, mark the i'th row of M with a party $\breve{x}_i \subseteq \mathcal{P}$. The column vector $v = (s, r_1, r_2, ..., r_n)$, in which $s \subseteq \mathbb{Z}_p$ is the secret to be shared, $r_1, r_2, ..., r_n \subseteq \mathbb{Z}_p$ are chosen at random. Based on Π, l shares of which Mv is the vector make up the secret s. Correspond to $\breve{x}_i$ there exists a $(Mv)_i$.

On the basic of the above definitions there is a linear secret sharing-scheme (LSSS), which observes the definitions as follows: assume that for the access structure $\mathbb{A}$ there exists a LSSS Π. S is an authorized set that belongs to $\mathbb{A}$. Let $I = \{i : \breve{x}_i \in S\}$ and let $\{\lambda_i\}$ is a valid share of s, there is $\{\omega_i \in \mathbb{Z}_p\}_{i \in I}$, then $\sum_{i \in I} \omega_i \lambda_i = s$.

3.2 Non-Monotonic Access Structures

With the negated attributes, we can construct non-monotonic access structures based on the ABE monotonic access structures. To accomplish it, a set of attributes is selected as the universe. From the universe, a set of d attributes is selected to encrypt a ciphertext was setup by the authority. Attributes in the set were called positive attributes and the negation of remaining attributes were called negated attributes. To recover the ciphertext, the decryptor must have at least $d+1$ attributes to perform an interpolation. Moreover the decryptor has to check the rest namely the one point, if the point differs from the d attributes and it is in the negated attributes, then decryptor has the privilege to achieve the share of message. Otherwise, the decryptor can not obtain the message.

At first, there is a set of attributes $\mathcal{P}$ in which the attribute $\breve{x}$ can be positive like x or negated (the negation of attribute) like x'. $\mathcal{A}$ is a set of monotonic access structures over $\mathcal{P}$ for which we given a LSSS $\{\Pi_{\mathbb{A}}\}_{\mathbb{A} \in \mathcal{A}}$. $\tilde{\mathcal{A}}$ is a family of non-monotonic access structures over $\tilde{\mathcal{P}}$ including all positive attributes of $\mathcal{P}$. For $\forall \mathbb{A} \in \mathcal{A}$, there exists a non-monotonic access structure $\tilde{\mathbb{A}}$. Let $\tilde{S} \subset \tilde{\mathcal{P}}, N(\tilde{S}) \subset \mathcal{P}$, namely, the attributes in $\tilde{S}$ are positive, but the attributes in $N(\tilde{S})$ may be positive or negated. Then let $\tilde{S} \subset N(\tilde{S})$. For every attribute $x \in \tilde{\mathcal{P}}$ but $x \notin \tilde{S}$, we have $x' \in N(\tilde{S})$. Therefore, $N(\tilde{S})$ include all attributes in S and the other negated attributes not in S. So corresponding to the monotonic access structure $\mathbb{A}$ over $N(\tilde{S})$ there is a non-monotonic access structure $\tilde{\mathbb{A}}$ over $\tilde{S}$.

3.3 Mathematical Background

Composite Order Bilinear Maps. Let $N = p_1 p_2 p_3$ (p_1, p_2, p_3 are primes and different from each other), $\mathbb{G}, \mathbb{G}_T$ are cyclic groups of order N. Let $e : \mathbb{G} \times \mathbb{G} \to \mathbb{G}_T$ denote a bilinear map. e is a valid bilinear map from G to G_T if e satisfies the properties as follows:

1) **Bilinear**: $\forall a, b \in \mathbb{Z}_N, e(g^a, g^b) = e(g, g)^{ab}$

2) **Non-degenerate**: There exists $g \in \mathbb{G}$ that make N is the order of $e(g, g)$.

3) **Computable**: For $\forall u, v \in G, e(u, v)$ is computable.

Lagrange Coefficients. For $\forall i \in \mathbb{Z}_p$ and a set $S \in \mathbb{Z}_p$, there is $\triangle_{i,S} = \prod_{j \in S, j \neq i} \frac{x-j}{i-j}$. By utilizing the collision-resistant hash function $H : \{0,1\}^* \to \mathbb{Z}_p^*$, we can link every attribute with one and only element in $\mathbb{Z}_p^*$.

3.4 Assumption

The Decisional Bilinear Diffie-Hellman (BDH) Assumption. The decisional BDH assumption is that: Choose randomly $a, b, c, z \in \mathbb{Z}_p$, any polynomial-time adversaries can not distinguish the tuple $(A = g^a, B = g^b, C = g^c, Z = e(g, g)^{abc})$ from the tuple $(A = g^a, B = g^b, C = g^c, Z = e(g, g)^z)$.

4 System Model

HP is honest but curious, which means HP will try to find as much PHI stored in cloud servers as possible, but it will observe the rules honestly. In a sense, the HP may be corrupted by some malicious users, moreover, some users may want to achieve authories beyond theirs. Since the HP is not considered to be fully trusted, the HP must

support patient to specify the access structure to control the data decryption. The main requirements of this scheme are presented as follows:

1) Confidentiality. Since the data storage is provided by the HP, the data should not be leaked even though HP is attacked by malicious users. At the same time, the unauthorized physicians who do not satisfy the access policy can not gain the plaintext of PHI.

2) *Revocability.* While patient has updated the revocation list, the physicians who possess the attributes revoked by patient can not decrypt ciphertext successfully.

In this paper, to guarantee the correctness of the scheme, we consider the transport channel is fully secure. Before giving the concrete construction, we illustrate our framework in Figure 1. There are five entities in our framework:

- Patient: Patient is equipped with sensors in body area. Sensors can collect PHI of patient and transmit the PHI to mobile. Patient who owns the PHI, has the privilege to specify the access structure and update the revocation list.
- Mobile: Mobile is a transmitter which is used to accept the PHI and transmit ciphertext of PHI to Base Station.
- Base Station: Base station is the repeater between Mobile and HP. Via base station, data is sent by Mobile can be transmitted to HP.
- Healthcare Provider (HP): The cloud infrastructures including processors, bandwidth, storage etc are preserved by HP. We suppose that the storage space, bandwidth, computing performance of HP can be expandable, so that HP owns powerful performance. In our system, HP provides several functions as follows: data storage, key distribution, data transmission, an update of revocation list.
- Physician: Serving as the end of the system, physician is the decryptor who can achieve PHI depending on his attributes.

The basic architecture of M-healthcare includes three components: body area networks (BANs), wireless transmission and healthcare which are illustrated in Figure 1. Via wireless transmission, the ciphertext of PHI is transmitted from BANs to healthcare. The system operates as follows:

At first, the sensors in body area collect the patient's PHI and transmit PHI securely to Mobile. Mobile will encrypt the PHI based on various sets of attributes which are chosen from the universe negotiated by patient and HP. Then mobile transmits the encrypted data to HP via base station. Finally, if the attributes that physician possesses satisfy the access structure and the physician's ID is not in the revocation list, HP will generate private key for physician. The authorized physicians may gain PHI.

5 Our Construction

In this section, we will give concrete construction of ABE with non-monotonic access structures supporting fine-grained attribute revocation, which involves quadruplicate algorithm: *Setup*, *Encryption*, *Key Generation* and *Decryption*. In our system, we consider the two users patient and physician as the encryptor and the decryptor respectively.

Let $N = p_1p_2p_3(p_1,\ p_2,\ p_3$ are primes and different), $\mathbb{G}$, $\mathbb{G}_T$ are cyclic groups of order N. Let $e : \mathbb{G} \times \mathbb{G} \to \mathbb{G}_T$ denote a bilinear map where e is the generation of $\mathbb{G}_{p_1}$, and Y is the generation of $\mathbb{G}_{p_2}$.

Setup$(1^\lambda, d, n)$: The parameter d specifies the count of attributes for every ciphertext. Let positive attributes set $\tilde{S} = \{1, 2, ..., d\}$ and the user identity set $U = \{1, 2, ..., n\}$. Choose $t_i, \mu_i \in \mathbb{Z}_{p_1}$ randomly, for any attribute $i \in \tilde{S}$, compute $T_i = g^{t_i}, h_i = g^{\mu_i}$. Then, choose $c \in \mathbb{Z}_{p_1}$ randomly, for any $i \in \{1, 2, ..., n, n+2, ..., 2n\}$, compute $f_i = g^{c^i}$. Choose two secrets $\alpha, \beta \in \mathbb{Z}_{p_1}$ uniformly at random, and compute $g_1 = g^\alpha$ and $g_2 = g^\beta$. Choose two polynomials $h(x)$ and $q(x)$ of degree d randomly with the constraint is that $q(0) = \beta$. ($h(x)$ has no constraint.) Finally, choose a from $\mathbb{Z}_{p_1}$ randomly. The published public parameters are:

$$\begin{aligned} PK \quad = \quad & (N, g, g^a, g_1, g_2; g^{q(1)}, g^{q(2)}, ..., g^{q(d)}; g^{h(0)}, \\ & g^{h(1)}, ..., g^{h(d)}; \{T_i\}_{i\in\tilde{S}}, \{h_i\}_{i\in\tilde{S}}, \\ & \{f_i\}_{i\in\{1,2,...,n,n+2,...,2n\}}). \end{aligned}$$

The master key is:

$$MK = (\alpha, a, c, \{t_i, \mu_i\}_{i\in\tilde{S}}, Y).$$

The functions $T, V{:}\mathbb{Z}_{p_1}{\to}\mathbb{G}_{p_1}$ are defined by the public parameters, which are public and computable. Then, compute:

$$T(x) = g_2^{x^d} \cdot g^{h(x)}, V(x) = g^{q(x)}.$$

Encryption$(M, \tilde{S}, PK)$: Message $M \in \mathbb{G}_T$, then encrypt M(M can be PHI) under $\tilde{S}$. Then,choose $s, y \in \mathbb{Z}_{p_1}$ at random, and computes:

$$E^{(1)} = Me(g_1, g_2)^s \cdot e(f_1, f_n)^y,\ E^{(2)} = g^s,\ E^{(3)} = (g^a)^y$$

For any $x \in \tilde{S}$, then computes:

$$E_x^{(4)} = T(x)^s,\ E_x^{(5)} = V(x)^s$$

Choose a d degree polynomial $l(x)$ randomly with the constraint is $l(0) = y$. For any $x{\in}\tilde{S}$, S_x is the non-revocation list, R_x is the revocation list, let $S_x = U - R_x$ ($S_x \neq \varnothing$), then computes:

$$E_x^{(6)} = g^{l(x)},\ E_x^{(7)} = T_x^{l(x)}$$

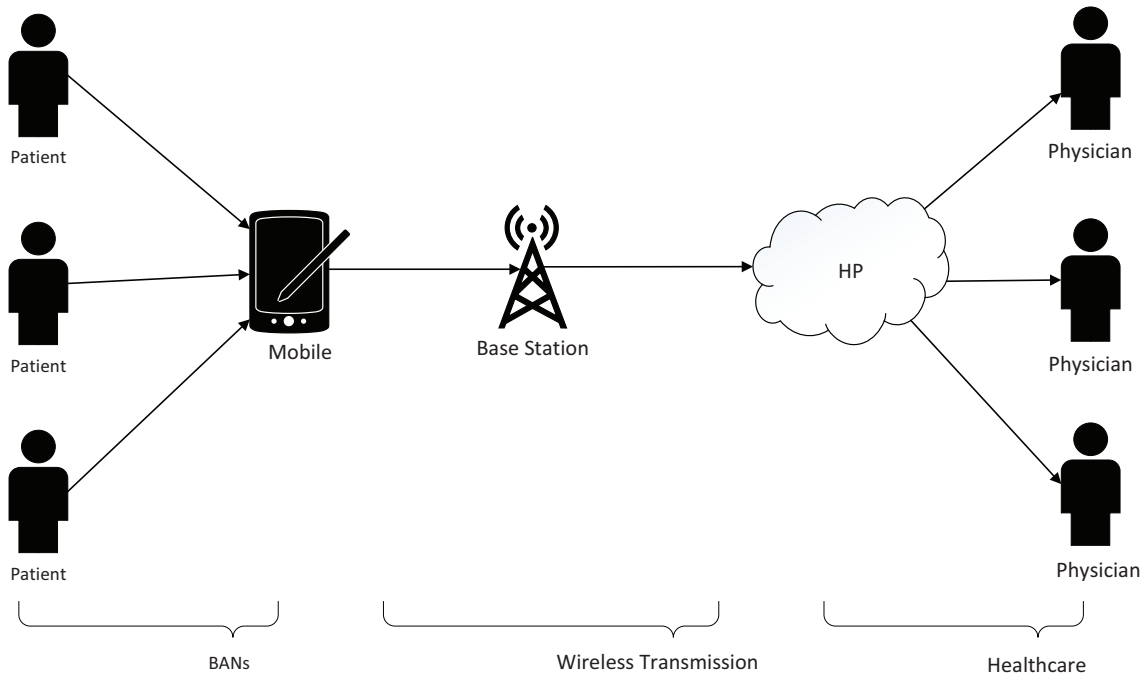


Figure 1: A basic architecture of M-healthcare

If $S_x \neq U$ namely $R_x \neq \varnothing$, choose η_x, s_x from $\mathbb{Z}_{p_1}$ at random and computes:

$$\begin{aligned} E_x^{(8)} &= g^{\eta_x}(h_x \prod_{j\in S_x} f_{n+1-j})^{l(x)}, \\ E_x^{(9)} &= g^{s_x}, \\ E_x^{(10)} &= g^{\eta_x}(\prod_{j\in R_x} f_{n+1-j})^{s_x} \end{aligned}$$

Remark: $\eta_x, s_x, E_x^{(9)}, E_x^{(10)}$ are used to randomize $E_x^{(8)}$ which is used for revocation to prevent $e(g_1, g_n)^{l(x)}$ from being computed by the potential adversary. If $S_x = U$ namely $R_x = \varnothing$, then computes:

$$E_x^{(8)} = (h_x \prod_{j\in S_x} f_{n+1-j})^{l(x)},\ E_x^{(9)} = E_x^{(10)} = 1$$

Namely:

$$\eta_x = s_x = 0$$

Then output the ciphertext as:

$$\begin{aligned} E = (\gamma, E^{(1)}, E^{(2)}, \{E_x^{(3)}, E_x^{(4)}, E_x^{(5)}, E_x^{(6)}, E_x^{(7)}, E_x^{(8)}, \\ E_x^{(9)}, E_x^{(10)}\}_{x\in\tilde{S}}). \end{aligned}$$

Key Generation$(\tilde{\mathbb{A}}, MK, PK)$: Except for the attributes which are in $\tilde{\mathbb{A}}$(suppose that can be checked efficiently), if the negation of remaining attributes are not the negated attributes of $N(\tilde{S})$, this algorithm will generate the private components for user with which the users can decrypt the ciphertext and obtain the data. By utilizing the LSSS to obtain the shares $\{\lambda_i\}$ of the secret α. We also select a $r_i \in \mathbb{Z}_{p_1}$ for each i.

For each $i, \tilde{x}_i$ is positive, we have:

$$D_i^{(1)} = g_2^{\lambda_i} \cdot T(x_i)^{r_i},\ D_i^{(2)} = g^{r_i}$$

The fine-grained attribute revocation performs under the set of positive attributes. At first, choose t, ξ_i from $\mathbb{Z}_{p_1}$ and Y_0 from $\mathbb{G}_{p_3}$ randomly. Then compute:

$$D_i^{(3)} = g^t Y_0,\ D_i^{(4)} = g^{at+c^{ID}\mu_i+t_i\xi_i} Y_{i,1},\ D_i^{(5)} = g^{\xi_i} Y_{i,2}$$

Then we can achieve the key component for positive attribute x:

$$D_i = (D_i^{(1)}, D_i^{(2)}, D_i^{(3)}, D_i^{(4)}, D_i^{(5)})$$

For each $i, \tilde{x}_i$ is negated, we have:

$$D_i^{(6)} = g_2^{\lambda_i+r_i},\ D_i^{(7)} = V(x_i)^{r_i},\ D_i^{(8)} = g^{r_i}$$

Then we can get the key component for negated attribute x':

$$D_i = (D_i^{(1)}, D_i^{(2)}, D_i^{(6)}, D_i^{(7)}, D_i^{(8)})$$

For all of the shares i, the private key D for decryptor to decrypt the ciphertext is made up of D_i.

Decryption(E, D): E and D are given as a ciphertext and private key. The decryption perform as follows: Let $I = \{i : \tilde{x} \in N(\tilde{S})\}$. An efficient process related to the LSSS can generate a set of coefficients $\Omega = \{\omega_i\}_{i\in I}$ which satisfy $\Sigma_{i\in I}\omega_i\lambda_i = \alpha$ (the λ_i, α is unknown to the decryption).

For each i, $\tilde{x} \in N(\tilde{S})$ and $x_i \in \tilde{S}$, namely the attribute is positive, we have:

$$\begin{aligned} Z_i &= e(D_i^{(1)}, E^{(2)})/e(D_i^{(2)}, E_i^{(2)}) \\ &= e(g_2^{\lambda_i} \cdot T(x_i)^{r_i}, g^s)/e(g^{r_i}, T(x)^s) \\ &= e(g_2, g)^{s\lambda_i} \end{aligned}$$

For each i, $\tilde{x} \in N(\tilde{S})$ and $x_i' \notin \tilde{S}$, namely the attribute is negated. Let $\tilde{S}_i = \tilde{S} \cup \{x_i'\}$, then $|\tilde{S}_i| = d+1$. Based on the function $V(x)$ and $\tilde{S}_i$, compute lagrangian coefficients $\{\sigma_x\}_{x\in\tilde{S}_i}$ which satisfy $\Sigma_{x\in\tilde{S}_i}\sigma_x qx = q(0) = \beta$, then compute:

$$\begin{aligned} Z_i &= \frac{e(D_i^{(6)}, E^{(2)})}{e(D_i^{(8)}, \prod_{x\in\tilde{S}}(E_x^{(5)})^{\sigma_x}) \cdot e(D_i^{(7)}, E^{(2)})^{\sigma_{x_i}}} \\ &= \frac{e(g_2^{\lambda_i + r_i}, g^s)}{e(g^{r_i}, \prod_{x\in\tilde{S}}(V(x)^s)^{\sigma_x}) \cdot e(V(x_i)^{r_i}, g^s)^{\sigma_{x_i}}} \\ &= \frac{e(g_2^{\lambda_i}, g^s) \cdot e(g_2^{r_i}, g^s)}{e(g^{r_i}, g^{s\sum_{x\in\tilde{S}}\sigma_x q(x)}) \cdot e(g^{r_i\sigma_{x_i}q(x_i)}, g^s)} \\ &= \frac{e(g_2, g)^{s\lambda_i} \cdot e(g, g)^{r_i s\beta}}{e(g, g)^{r_i s\sum_{x\in N(\tilde{S})}\sigma_x q(x)}} \\ &= e(g_2, g)^{s\lambda_i} \end{aligned}$$

Then compute the revocation component. At first, let $L = \{x_i | x_i \in \tilde{S}, ID \notin R_x\}$. For each $x \in L$, then compute:

$$\begin{aligned} X_i &= \frac{e(D_i^{(4)}, E_x^{(6)})e(E_x^{(6)}, \prod_{j\in S_x, j\neq ID} f_{n+1-j+ID})e(f_{ID}, E_x^{(10)})}{e(D_i^{(5)}, E_x^{(7)})e(f_{ID}, E_x^{(8)})e(E_x^{(9)}, \prod_{j\in R_x} f_{n+1-j+ID})} \\ &= \frac{e(g^{at}, g^{l(x)})}{e(f_1, f_n)^{l(x)}} \end{aligned}$$

Finally, let $A = \{i : x_i \in \tilde{S}\}$, the message is achieved by decryption as follows:

$$\begin{aligned} \frac{E^{(1)}}{\prod_{i\in I} Z_i} \cdot \frac{\prod_{i\in A} X_i}{e(D_i^{(3)}, E^{(3)})} &= \frac{Me(g_1, g_2)^s e(f_1, f_n)^y \cdot e(g^{at}, g^y)}{e(g_2, g)^{s\alpha} \cdot e(f_1, f_n)^y e(g^t, g^{ay})} \\ &= M \end{aligned}$$

Discussion. By utilizing our scheme, the patient can encrypt the PHI by specifying a set of attributes. The fine-grained attribute revocation supports that the patient revokes the attribute of physicians, such as revoking the physicians who possess the attribute "Nursing Care". So the physicians who hold the attribute can not access the PHI. According to the non-monotonic access structures, not only the encryption of the PHI was using less attributes but also NOT operation over the access structure can be achieved. The goal of PHI confidentiality can be attained flexibly.

6 Analysis

6.1 Security Analysis

In our scheme, a physician who can recover the data from ciphertext, must be an unrevoked user with valid authority. Therefore, we will analyse the security from two aspects that revocation and decryption.

At first, from the aspect of revocation, the adversary is an revoked user. If he wants to recover the data from $E^{(1)}$, what he must compute is $e(f_1, f_n)^y$ which is for revocation. But for y, there is a d degree polynomial $l(x)$ and the constraint is $l(0) = y$. Moreover, $y, l(x)$ are chosen randomly and the scope of $y, l(x)$ is the encryption algorithm. Therefore the y can not be computed. The only way to compute $e(f_1, f_n)^y$ is relying on the X_i. Note that if an attribute of the set of adversary is revoked, then he can only compute the result $e(g^{at}, g^{l(x)}) \cdot e(f_1, f_n)^{s_x}$. Due to the s_x is a random value, that is to say, the adversary can not compute $e(g^{at}, g^{l(x)})$. Namely the adversary can not compute $e(f_1, f_n)^y$.

Then, from the aspect of decryption, the adversary is an unrevoked user without valid authority, so he can gain the $e(f_1, f_n)^y$ legitimately. To prove he can attack the scheme, he should recover M from $Me(g_1, g_2)^s$. According to the decisional BDH assumption, think of $A = g_1 = g^\alpha, B = g_2 = g^\beta, C = g^s, Z = e(g_1, g_2)^s = e(g, g)^{\alpha\beta s}$, namely $a = \alpha, b = \beta, c = s$. Considering the M' is the message which is achieved by the adversary, if $M' \neq M$, then the adversary can not recover the message, otherwise $M' = M$, where $z = \alpha\beta s$, namely the adversary can distinguish the tuple $(A = g^a, B = g^b, C = g^c, Z = e(g, g)^{abc})$ from the tuple $(A = g^a, B = g^b, C = g^c, Z = e(g, g)^z)$. In other words, the adversary can solve the decisional BDH assumption. As a result, the semantic security of our scheme is that the construction can be easily broken by the adversary who can solve the decisional BDH assumption. However, the decisional BDH assumption is proven to be a hard problem to be solved.

6.2 Comparison

In this section, we compare our scheme with some existing works which are similar to our scheme in attribute-based encryption and healthcare.

There are three schemes [15, 18, 19] to be compared with our scheme. The first scheme [15] is a ciphertext-policy attribute-based encryption (CP-ABE) scheme, which applies direct revocation to revoke user or attribute. The difference of revocation between this paper and our scheme is that this paper makes use of a mediator who holds a revocation list to implement revocation. In the revocation list, there is a set of user identities which are respectively related to a set of attributes. The second scheme [18] is a multi-authority attribute-based encryption (MA-ABE) scheme and the third scheme [19] is a key-policy attribute-based encryption (KP-ABE) scheme, which all take advantage of the indirect revocation. Based on the indirect revocation, the second and third scheme have to rely on the authority to enforce revocation, which means the ciphertext and users' private key must be updated. Moreover the second scheme does not support revoking attributes. Further, the three schemes do not support non-monotonic access structures which means they can not support NOT operation over access structure. If a patient wants to encrypt the PHI with negated attributes

Table 1: Property comparison

Schemes	[15]	[18]	[19]	Ours
Encryption type	CP-ABE	MA-ABE	KP-ABE	KP-ABE
Revocation type	Direct	Indirect	Indirect	Direct
User revocation	√	√	√	√
Attribute revocation	√	×	√	√
Non-monotonic	×	×	×	√

Table 2: Computation comparison

Schemes	[15]	[18]	[19]	Ours
Key generation	$(2n+1)e$	$(n+1)e$	$(n+1)e$	$6ne$
Encryption	$(n+1)e+e_T$	$e_T+(n+1)e$	$e_T+(n+1)e$	$2e_T+(3n+2)e$
Decryption	$(3n+1)e_T+2np$	$(n+1)(p+e_T)$	$(n+1)(p+e_T)$	$(3n+5k)p+(n-k)e_T+(n-k)e$

by utilizing the above schemes, he has to input more attributes, for example, "Nurse" and "Not Nurse" and so on. Certainly, to accomplish the non-monotonic access structures in our scheme, we must sacrifice efficiency.

In addition, we present the theoretical comparison with the above schemes in Table 1 and Table 2 respectively for property and computation. The explanation of notations defined by us in tables are as follows: p, e_T and e represent the computation cost of a bilinear pairing, an exponentiation in $\mathbb{G}_T$ and an exponentiation in $\mathbb{G}$, respectively. Due to positive attributes and negated attributes in our scheme, k denotes the count of positive attributes.

7 Conclusions

In the paper, for the first time, we proposed an ABE scheme with non-monotonic access structures supporting fine-grained attribute revocation in m-healthcare. The advantage is that we provide a flexible solution for access control of m-healthcare. However, there exists some problems such as the slightly large size of ciphertext and the lower efficiency. The next step is to reduce the size of ciphertext and improve the efficiency while ensuring the properties of the scheme.

Acknowledgments

This work was supported in part by the National Science Foundation of China (No. 61370026), the National High Technology Research and Development Program of China (No. 2015AA016007), the Sichuan Key Technology Support Program (No. 2014GZ0106), Science Technology Project of Guangdong Province (No. 2016A010101002) and Fundamental Research Funds for the Central Universities under Grant ZYGX2016J091.

References

[1] M. Asim, L. Ibraimi, and M. Petković, "Ciphertext-policy attribute-based broadcast encryption scheme," in *Communications and Multimedia Security*, pp. 244–246, 2011.

[2] N. Attrapadung and H. Imai, "Conjunctive broadcast and attribute-based encryption," in *Pairing-Based Cryptography (Pairing'09)*, pp. 248–265, 2009.

[3] M. Bayat, M. Aref, "An attribute based key agreement protocol resilient to KCI attack," *International Journal of Electronics and Information Engineering*, vol. 2, no. 1, pp. 10–20, 2015.

[4] J. Bethencourt, A. Sahai, and B. Waters, "Ciphertext-policy attribute-based encryption," in *IEEE Symposium on Security and Privacy*, pp. 321–334, 2007.

[5] A. Boldyreva, V. Goyal, and V. Kumar, "Identity-based encryption with efficient revocation," in *Proceedings of the 15th ACM Conference on Computer and Communications Security*, pp. 417–426, 2008.

[6] D. Boneh and M. Franklin, "Identity-based encryption from the weil pairing," in *Advances in Cryptology (CRYPTO'01)*, pp. 213–229, 2001.

[7] M. Chase, "Multi-authority attribute based encryption," in *Theory of cryptography*, pp. 515–534, 2007.

[8] 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.

[9] L. Cheung and C. Newport, "Provably secure ciphertext policy ABE," in *Proceedings of the 14th ACM Conference on Computer and communications security*, pp. 456–465, 2007.

[10] M. K. Debnath, S. Samet, and K. Vidyasankar, "A secure revocable personal health record system with policy-based fine-grained access control," in *13th Annual Conference on Privacy, Security and Trust (PST'15)*, pp. 109–116, 2015.

[11] L. Gatzoulis and I. Iakovidis, "Wearable and portable ehealth systems," *Engineering in Medicine and Biology Magazine*, vol. 26, no. 5, pp. 51–56, 2007.

[12] S. Gorbunov, V. Vaikuntanathan, and H. Wee, "Attribute-based encryption for circuits," *Journal of the ACM*, vol. 62, no. 6, pp. 45–78, 2015.

[13] V. Goyal, O. Pandey, A. Sahai, and B. Waters, "Attribute-based encryption for fine-grained access control of encrypted data," in *Proceedings of the 13th ACM Conference on Computer and Communications Security*, pp. 89–98, 2006.

[14] C. Guo, R. Zhuang, Y. Jie, Y. Ren, T. Wu, and K. R. Choo, "Fine-grained database field search using attribute-based encryption for e-healthcare clouds," *Journal of medical systems*, vol. 40, no. 11, p. 235, 2016.

[15] L. Ibraimi, M. Petkovic, S. Nikova, P. Hartel, and W. Jonker, "Mediated ciphertext-policy attribute-based encryption and its application," in *Information security applications*, pp. 309–323, 2009.

[16] H. Jung and K. Chung, "Phr based life health index mobile service using decision support model," *Wireless Personal Communications*, vol. 86, no. 1, pp. 315–332, 2016.

[17] A. Lewko, T. Okamoto, A. Sahai, K. Takashima, and B. Waters, "Fully secure functional encryption: Attribute-based encryption and (hierarchical) inner product encryption," in *Advances in Cryptology (EUROCRYPT'10)*, pp. 62–91, 2010.

[18] M. Li, S. Yu, K. Ren, and W. Lou, "Securing personal health records in cloud computing: Patient-centric and fine-grained data access control in multi-owner settings," in *International Conference on Security and Privacy in Communication Systems*, pp. 89–106, 2010.

[19] M. Li, S. Yu, Y. Zheng, K. Ren, and W. Lou, "Scalable and secure sharing of personal health records in cloud computing using attribute-based encryption," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 1, pp. 131–143, 2013.

[20] Q. Li, D. Feng, and L. Zhang, "An attribute based encryption scheme with fine-grained attribute revocation," in *Global Communications Conference (GLOBECOM'12)*, pp. 885–890, 2012.

[21] Q. Li, H. Xiong, F. Zhang, and S. Zeng, "An expressive decentralizing kp-abe scheme with constant-size ciphertext.," *International Journal of Network Security*, vol. 15, no. 3, pp. 161–170, 2013.

[22] C. Liu, W. Hsien, C. Yang, and M. 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.

[23] J. Liu, X. Huang, and J. K. Liu, "Secure sharing of personal health records in cloud computing: ciphertext-policy attribute-based signcryption," *Future Generation Computer Systems*, vol. 52, pp. 67–76, 2015.

[24] O. G. Morchon and H. Baldus, "Efficient distributed security for wireless medical sensor networks," in *International Conference on Intelligent Sensors, Sensor Networks and Information Processing*, pp. 249–254, 2008.

[25] T. Okamoto and K. Takashima, "Fully secure functional encryption with general relations from the decisional linear assumption," in *Advances in Cryptology (CRYPTO'10)*, pp. 191–208, 2010.

[26] R. Ostrovsky, A. Sahai, and B. Waters, "Attribute-based encryption with non-monotonic access structures," in *Proceedings of the 14th ACM Conference on Computer and Communications Security*, pp. 195–203, 2007.

[27] M. B. Parish and P. Yellowlees, "The rise of person-centered healthcare and the influence of health informatics and social network applications on mental health care," in *Mental Health Informatics*, pp. 17–39, 2014.

[28] M. Pirretti, P. Traynor, P. McDaniel, and B. Waters, "Secure attribute-based systems," *Journal of Computer Security*, vol. 18, no. 5, pp. 799–837, 2010.

[29] H. Qian, J. Li, Y. Zhang, and J. Han, "Privacy-preserving personal health record using multi-authority attribute-based encryption with revocation," *International Journal of Information Security*, vol. 14, no. 6, pp. 487–497, 2015.

[30] A. Sahai and B. Waters, "Fuzzy identity-based encryption," in *Advances in Cryptology (EUROCRYPT'05)*, pp. 457–473, 2005.

[31] S. Sarpong, C. Xu, and X. Zhang, "An authenticated privacy-preserving attribute matchmaking protocol for mobile social networks," *International Journal of Network Security*, vol. 17, no. 3, pp. 357–364, 2015.

[32] H. Shi and R. Guo, "Provably-secure certificateless key encapsulation mechanism for e-healthcare system," *International Journal of Network Security*, vol. 17, no. 5, pp. 548–557, 2015.

[33] K. Takashima, "Expressive attribute-based encryption with constant-size ciphertexts from the decisional linear assumption," in *International Conference on Security and Cryptography for Networks*, pp. 298–317, 2014.

[34] S. Tu, S. Niu, H. Li, Y. Xiao-ming, and M. Li, "Fine-grained access control and revocation for sharing data on clouds," in *26th International Parallel and Distributed Processing Symposium Workshops & PhD Forum (IPDPSW'12)*, pp. 2146–2155, 2012.

[35] B. Waters. "Dual system encryption: Realizing fully secure ibe and hibe under simple assumptions,". in *Advances in Cryptology (CRYPTO'09)*, pp. 619–636. 2009.

[36] N. Xiao, R. Sharman, H. R. Rao, and S. Upadhyaya, "Factors influencing online health information search: An empirical analysis of a national cancer-related survey," *Decision Support Systems*, vol. 57, pp. 417–427, 2014.

[37] H. Xiong, J. Tao, and C. Yuan, "Enabling telecare medical information systems with strong authentication and anonymity," *IEEE Access*, 2017.

[38] Y. Zhang, D. Zheng, X. Chen, J. Li, and H. Li, "Computationally efficient ciphertext-policy attribute-based encryption with constant-size ciphertexts," in *International Conference on Provable Security*, pp. 259–273, 2014.

[39] Y. Zhao, F. Yue, S. Wu, H. Xiong, and Z. Qin, "Analysis and improvement of patient self-controllable multi-level privacy-preserving cooperative authentication scheme," *International Journal of Network Security*, vol. 17, no. 6, pp. 779–786, 2015.

[40] J. Zhou, X. Lin, X. Dong, and Z. Cao, "Psmpa: Patient self-controllable and multi-level privacy-preserving cooperative authentication in distributedm-healthcare cloud computing system," *IEEE Transactions on Parallel and Distributed Systems*, vol. 26, no. 6, pp. 1693–1703, 2015.

Biography

Yang Zhao is a Ph.D. Candidate at the School of Computer Science and Engineering, University of Electronic Science and Technology of China. His research interests are in the area of network security and e-commerce protocol.

Pengcheng Fan received his B.S. degree from Hebei University of Science and Technology of China (HEBUST) in 2014. He is currently pursuing his M.S. degree in the School of Information and Software Engineering, UESTC. His research interests include cryptography and network security.

Haoting Cai received his B.S. degree from University of Electronic Science and Technology of China (UESTC) in 2014. He is currently pursuing his M.S. degree in the School of Information and Software Engineering, UESTC. His research interests include cryptography and network security.

Zhiguang Qin is the dean and professor in the School of Computer Science and Engineering, University of Electronic Science and Technology of China (UESTC). He received his PH.D. degree from UESTC in 1996. His research interests include: information security and computer network.

Hu Xiong is an associate Professor at University of Electronic Science and Technology of China (UESTC). He received his Ph.D. degree from UESTC in 2009. His research interests include cryptographic protocols and network security.

Improvement of Efficient and Secure Smart Card Based Password Authentication Scheme

Jongho Moon[1], Donghoon Lee[1], Jaewook Jung[1], and Dongho Won[2]
(Corresponding author: Dongho Won)

Department of Electrical and Computer Engineering, Sungkyunkwan University[1]
Department of Computer Engineering, Sungkyunkwan University[2]
Suwon 16419, Korea
(Email: dhwon@securiry.re.kr)
(Received Aug. 31, 2016; revised and accepted Nov. 15 & Dec. 25, 2016)

Abstract

Remote user authentication scheme is one of the most convenient authentication schemes to deal with secret data over public communication channel. In order to satisfy the security requirements, the smart card has become an essential device, one that is widely used. This is because its low computational cost and expedient portability. Recently, Liu et al. pointed out some security weaknesses in Li et al.'s scheme, such as man-in-the middle attack and insider attack. They hence claimed that their scheme is more secure and practical remote user authentication scheme. However, we find that Liu et al.'s scheme is still insecure against outsider attack and off-line password guessing attack. To overcome these security vulnerabilities, we propose a new authentication and key agreement scheme using smart card. In addition, we demonstrate that proposed authentication scheme has strong resistance to the various attacks. Finally, we compare the performance and functionality of the proposed scheme with other related schemes.

Keywords: Authentication; Biometrics; Elliptic Curve Cryptosystem; Smart Card

1 Introduction

Since Lamport [14] proposed the first password-based authentication scheme over insecure communication in 1981, password-based authentication schemes [1, 9, 10, 22, 25] have been extensively investigated. However, a main problem of password-based remote user authentication scheme is that a server must maintain a password table for verifying the legitimacy of a remote user. Therefore, the server requires additional memory space for storing the password table for verifying user identity. Furthermore, password is generally simple and can be easily broken or forgotten. For this reason, many researchers has proposed a new remote user authentication scheme by using biological characteristics of persons such as fingerprint, iris and so on. The main property of using biometric is its uniqueness. In the view of the fact that many remote user authentication schemes using biological characteristics [3, 16, 19, 20] have been proposed. In 2009, Xu et al. [21] proposed a novel user authentication and claimed that their scheme is secure against various attacks. However, Song [23] and Sood et al. [24] found that Xu et al.'s scheme has some weaknesses, and Sood et al. proposed an improved schemes. Subsequently, Chen et al. [2] pointed out that there are vulnerabilities on Song and Sood et al.'s schemes. Chen et al. then presented an enhanced version to solve the weaknesses. Recently, Li et al. [15] claimed that Chen et al.'s scheme is still insecure and proposed a modified smart card based remote user password authentication scheme. Unfortunately, Liu et al. [17] found that there are weaknesses in Li et al.'s scheme, such as the man-in-the-middle attack and insider attack, and proposed a novel scheme to defend against these security weaknesses. However, we found that Liu et al.'s scheme is still insecure against the outsider attack and off-line password guessing attack.

In this paper, we find that the security weaknesses of the two-factor authentication scheme by Liu et al. After careful analysis, we demonstrated their scheme does not actually resist off-line password guessing and user impersonation attacks. To overcome these security vulnerabilities, we propose a new biometrics-based authentication and key agreement scheme using smart card. In addition, we demonstrate that the proposed authentication scheme has strong resistance to various attacks, and compare the performance and functionality with other related schemes.

The rest of this paper is organized as follows. In Section 2, we briefly introduce some cryptographic definitions. In Section 3, we briefly review Liu et al.'s smart card-based password authentication scheme, and Section 4 analyzes its weaknesses. In Section 5, we propose new authentication scheme. Section 6 and 7 gives security and performance analysis of the proposed scheme.

Finally, we present the conclusion in Section 8.

2 Preliminaries

In this section, we briefly introduce the Elliptic curve cryptosystem, threat assumptions and fuzzy-extractor.

2.1 Elliptic Curve Cryptosystem

The elliptic curve cryptosystem (ECC) was first proposed by Koblitz [11] and Miller [18] to design public key cryptosystem, and presently it is widely used in several cryptographic schemes to provide desired level of security and performance [13]. For example, 160-bit ECC and 1024-bit RSA have the same security level in practice [8]. An elliptic curve E_K defined over a field K of the characteristic $\neq 2$ or 3 is the set of solutions $(x, y) \in K^2$ to the equation:

$$y^2 = x^3 + ax + b, \quad a, b \in K, 4a^3+27b^2 \neq 0.$$

Cryptosystems based on GF$(q)^*$ can be translated to systems using the group E, where E is an elliptic curve defined over GF(q). The point multiplications $kP = (P + P + ... + P, k$ times) that means k times addition of point P. Given an elliptic curve E defined over GF(q) and two points $P, Q \in E$, find an integer x such that $Q = xP$ if such x exists. This problem proved to be more intractable than the typically discrete logarithm problem. The details of the ECC definitions can be found in [11].

2.2 Threat Assumptions

We introduce the Dolev-Yao threat model [7] and consider the risk of side-channel attack [12] to construct the threat assumptions which are described as follows:.

1) An adversary $\mathcal{A}$ can be either a user or a server. A registered user can act as an adversary.

2) An adversary $\mathcal{A}$ can eavesdrop every communication in public channels. He/she can capture any message exchanged between user and server.

3) An adversary $\mathcal{A}$ has the ability to alter, delete or reroute the captured message.

4) Information can be extracted from the smart card by examining the power consumption of the card.

2.3 Fuzzy Extractor

We describe the basis of a biometric-based fuzzy extractor, which converts biometric information data into a random value. Based on Refs. [4, 5, 26], the fuzzy extractor is given by two procedures (Gen, Rep). The fuzzy extractor consists of two procedures (Gen, Rep).

- $Gen(BIO) \rightarrow \langle R, P \rangle$.
- $Rep(BIO^*, P) = R$ if BIO^* is reasonable close to BIO.

The function Gen is a probabilistic generation procedure, which on biometric input BIO outputs an "extracted" string $R \in \{0, 1\}^l$ and an auxiliary string $P \in \{0, 1\}^*$ and the function Rep is a deterministic reproduction procedure, which allows to recover R from the corresponding auxiliary string P and any vector BIO^* close to BIO. The detailed information about fuzzy extractor can be founded in literature [6].

3 Review of Liu et al.'s Scheme

In this section, we demonstrate Liu et al.'s smart card-based password authentication scheme [17] before demonstrating its weaknesses. Their scheme is an improvement of Li et al.'s scheme [15]. The notations used in Liu et al.'s scheme are listed in Table 1. Their scheme involves two entities, i.e., the user U_i and the server S, to communicate with each other to perform the following four phases: (1) The registration phase; (2) the login phase; (3) the authentication phase; and (4) the password change phase.

Table 1: The notations used in Liu et al.'s scheme

Term	Description
U_i	The i^{th} user
ID_i, PW_i	The identity and password of the user i
S	The server
x	The naster secret key stored in the S
T_i	The timestamp of the U_i
T_i'	The time of receiving the login request message
T_s	The timestamp of the server S
T_s'	The time of receiving the mutual authentication message
$h(\cdot)$	A secure hash function
$\oplus$	Exclusive-or operation
$\|$	Concatenation operation
sk	The shared session key

3.1 Registration Phase

At the beginning of the proposed scheme, the server S selects the master secret key x and a collision-free one-way hash function $h(\cdot)$. The user U_i then registers to the server S by the way below:

S1. The user U_i first selects his/her identity ID_i, password PW_i and a random number r, and then computes $h(r||PW_i)$. The U_i then submits $\{ID_i, h(r||PW_i)\}$ to the server S for registration over a secure channel.

S2. The server S computes the following parameters:

$$\begin{aligned} A_i &= h(ID_i \oplus x)||h(x), \\ B_i &= A_i \oplus h(r||PW_i), \\ C_i &= h(A_i||ID_i||h(r||PW_i)). \end{aligned}$$

S3. The server S stores the data $\{B_i, C_i, h(\cdot)\}$ on a new smart card and issues the smart card to the user U_i over a secure channel.

S4. The user U_i stores the random number r into the smart card.

3.2 Login Phase

This phase is invoked whenever the user U_i wants to login to the server S. These steps of the login phase are conducted as follows:

S1. The user U_i inserts his/her smart card into a card reader, and inputs his/her identity ID_i and password PW_i.

S2. The smart card first computes two parameters: $A_i' = B_i \oplus h(r||PW_i)$ and $C_i' = h(A_i'||ID_i||h(r||PW_i))$. Then, the smart card checks whether C_i' is equal to the stored C_i. If this holds, the smart card continues to perform the next step; otherwise, the smart card terminates this session.

S3. The smart card randomly generates a number α, and computes the following parameters:

$$\begin{aligned} D_i &= h(ID_i \oplus \alpha), \\ E_i &= A_i' \oplus \alpha \oplus T_i, \end{aligned}$$

where T_i is the current timestamp of the user U_i.

S4. The smart card sends the login request message $\{ID_i, D_i, E_i, T_i\}$ to the server S.

3.3 Authentication Phase

After completing this phase, the user U_i and the server S can mutually authenticate each other and establish a shared session key for the subsequent secret communication. These steps of the authentication phase are shown as follows:

S1. The server S verifies whether ID_i is valid and $T_i' - T_i \leq \triangle T$, where T_i' is the time of receiving the login request message and $\triangle T$ is a valid time threshold. If both conditions are true, the server S continues to execute Step 2; otherwise, the server S rejects the login request.

S2. The server S then computes the following parameters:

$$\begin{aligned} A_i' &= h(ID_i \oplus x)||h(x), \\ \alpha' &= E_i \oplus A_i' \oplus T_i, \\ D_i' &= h(ID_i \oplus \alpha'). \end{aligned}$$

Then, the server S checks whether D_i' is equal to the received D_i. If this holds, the server S confirms that the user U_i is valid and the login request is accepted; otherwise, the login request is rejected.

S3. The server S randomly generates a number β, and computes the following parameters:

$$\begin{aligned} F_i &= h(ID_i \oplus \beta), \\ G_i &= A_i' \oplus \beta \oplus T_s, \end{aligned}$$

where T_s is the current timestamp of the server S.

S4. The server S sends the mutual authentication message $\{F_i, G_i, T_s\}$ to the user U_i.

S5. Upon receiving the message $\{F_i, G_i, T_s\}$ from the S, the user U_i checks the validity of the T_s. If $T_s' - T_s \leq \triangle T$, where T_s' is the time of receiving the mutual authentication message, the user U_i continues to perform Step 6; otherwise, the user U_i terminates this connection.

S6. The user U_i computes $\beta' = G_i \oplus A_i' \oplus T_s$ and $F_i' = h(ID_i \oplus \beta')$, then checks whether F_i' equals to the received F_i. If they are equal, the validity of the server S is authenticated; otherwise, the session is terminated.

S7. The user U_i and the server S construct a shared session key $sk = h(\alpha'||\beta||h(A_i \oplus ID_i))$ to ensure the secret communication.

3.4 Password Change Phase

S1. The user U_i inserts his/her smart card into a card reader, enters his/her old identity ID_i and password PW_i, and requests to change the password.

S2. The smart card computes $A_i^* = B_i \oplus h(r||PW_i)$ and $C_i^* = h(A_i^*||ID_i||h(r||PW_i))$, and then checks whether C_i^* is equal to C_i that is stored in the smart card. If the equation holds, the user U_i submits the new password PW_i^{new}; otherwise, the smart card rejects the password change request.

S3. The smart card computes $B_i^{new} = A_i^* \oplus h(r||PW_i^{new})$ and $C_i^{new} = h(A_i^*||ID_i||h(r||PW_i^{new}))$. Then, the smart card replaces B_i andC_i with B_i^{new} and C_i^{new}, respectively.

4 Security Analysis of Liu et al.'s Scheme

Liu et al. claimed that their scheme could resist the various attacks. However, we find that their scheme is still insecure against the outsider and off-line password guessing attacks. The following attacks are based on the threat assumptions that a malicious adversary $\mathcal{A}$ was completely monitored through the communication channel connecting the U_i and S in the login and authentication phases, and obtained the information saved in their own smart card [7]. Thus, the $\mathcal{A}$ can eavesdrop, modify, insert, or delete any message transmitted via a public channel [14]. We now reveal the details of these problems.

4.1 Outsider Attack

The outsider is the person who has registered with the server S, not the person who is not the user of the system. In the registration phase, the server S stores $\{B_i, C_i, h(\cdot)\}$ on a smart card, and submits them to the user U_i. After receiving the smart card, the U_i stores the random number r into the smart card. Let $\mathcal{A}$ be an active adversary who is a legal user and owns a smart card to extract information $\{B_{\mathcal{A}}, C_{\mathcal{A}}, r_{\mathcal{A}}, h(\cdot)\}$. The $\mathcal{A}$ then can easily obtain $h(x)$ which is the same for each legal user belonging to the S.

$$h(ID_{\mathcal{A}} \oplus x)||h(x) = B_{\mathcal{A}} \oplus h(r_{\mathcal{A}}||PW_{\mathcal{A}}).$$

4.2 Off-line Password Guessing Attack

Suppose that an adversary $\mathcal{A}$ intercepts the communication messages $\{ID_i, D_i, E_i, T_i, F_i, G_i, T_s\}$ between the user U_i and the S, and steals the smart card of the U_i after login and authentication phase. The $\mathcal{A}$ then can extract the data $\{B_i, C_i, r, h(\cdot)\}$, and now perform an off-line password guessing to obtain the current password of the user U_i.

S1. The $\mathcal{A}$ selects a random password PW_i^*, calculates $h(ID_i \oplus E_i \oplus B_i \oplus h(r||PW_i^*) \oplus T_i)$, and compares it with D_i. If this holds, the adversary $\mathcal{A}$ infers that PW_i^* is the user U_i's password; otherwise, the $\mathcal{A}$ selects another password nominee, and performs the same processes, until he/she locates the valid password.

4.3 Not Support User Anonymity

Suppose an adversary $\mathcal{A}$ intercepts the communication messages $\{ID_i, D_i, E_i, T_i, F_i, G_i, T_s\}$, and then he/she can easily obtain the identity ID_i of the user U_i. We therefore concluded that Liu et al.'s scheme cannot provide user anonymity.

4.4 Not Support Perfect Forward Secrecy

Liu et al. claimed that even if an adversary obtained the server S's master secret key x, he/she cannot derive the previous session key sk because α and β are encrypted into the ciphertext D_i and F_i, respectively. Therefore, the $\mathcal{A}$ cannot obtain α and β. However, if the $\mathcal{A}$ obtain the server S's master secret key x, then he/she can easily obtain α and β. Suppose that an adversary $\mathcal{A}$ intercepts the communication messages $\{ID_i, D_i, E_i, T_i, F_i, G_i, T_s\}$. The $\mathcal{A}$ can then compute $\alpha = E_i \oplus (h(ID_i \oplus x)||h(x)) \oplus T_i$ and $\beta = G_i \oplus (h(ID_i \oplus x)||h(x)) \oplus T_s$.

5 The Proposed Scheme

In this section, we propose a new biometric-based password authentication scheme using smart card. In the proposed scheme, there are also two participants, the user U_i and the server S. The proposed scheme consists of four phases: registration, login, authentication, password changing phase. For convenience, some notations used in the proposed scheme are described in Table 2.

Table 2: The notations used in the proposed scheme

Term	Description
U_i	The i^{th} user
ID_i, PW_i	The identity and password of the user i
S	The server
x	The master secret key stored in the S
P	The base point of the elliptic curve E
rP	The point multiplication defined as $rP = \underbrace{P + P + \cdots + P}_{r \ times}$
T_i	The timestamp of the user U_i
T_i'	The time of receiving the login request message
T_s	The timestamp of the S
T_s'	The time of receiving the mutual authentication message
R_i, P_i	The U_i's nearly random binary string and auxiliary binary string
$h(\cdot)$	A collision-resistant hash function
$\oplus$	Exclusive-or operation
$\|\|$	Concatenation operation
sk	The shared session key

5.1 Registration Phase

At the beginning of the proposed scheme, the server S selects the master secret key x, the base point P of the elliptic curve E and a collision-resistant one-way hash function $h(\cdot)$. Then, the user U_i registers to the server S by the way below:

S1. The U_i imprints the personal biometric information BIO_i at the sensor. The sensor then scans the BIO_i, extracts (R_i, P_i) from $Gen(BIO_i) \rightarrow (R_i, P_i)$, and stores P_i in the memory. Next, the U_i selects the identity ID_i and password PW_i, and computes $RPW_i = h(PW_i||R_i)$. Lastly, the U_i sends the registration request message $\{ID_i, RPW_i\}$ to the S over a secure channel.

S2. After receiving the registration request message from the U_i, the server S verifies whether ID_i is valid, and computes the following parameters:

$$\begin{aligned} A_i &= h(ID_i \oplus x), \\ B_i &= h(A_i) \oplus RPW_i, \\ C_i &= h(ID_i||RPW_i), \\ D_i &= x \oplus A_i \oplus h(x). \end{aligned}$$

S3. The server S stores the data $\{B_i, C_i, D_i, h(\cdot), P\}$ on a new smart card and issues the smart card to the user U_i over a secure channel.

S4. The user U_i stores the random string P_i into the smart card.

5.2 Login Phase

This phase is invoked whenever the user U_i wants to login to the server S. The steps of this phase are conducted as follows.

S1. The U_i inserts his/her smart card into the card reader and enters the identity ID_i and password PW_i, and imprints the biometrics BIO_i^* at the sensor. The sensor then sketches BIO_i^* and recovers R_i from $Rep(BIO_i^*, P_i) \rightarrow R_i$.

S2. The smart card first computes two parameters: $RPW_i = h(PW_i||R_i)$ and $C_i' = h(ID_i||RPW_i)$. The smart card then examines whether C_i' is equal to the stored C_i. If this holds, the smart card continues to perform Step 3; otherwise, the smart card terminates this session.

S3. The smart card randomly generates a number α and n_i, and computes the following parameters:

$$\begin{aligned} h(A_i) &= B_i \oplus RPW_i, \\ AID_i &= ID_i \oplus h(A_i), \\ E_i &= \alpha P, \\ F_i &= h(ID_i||h(A_i)||E_i||T_i), \end{aligned}$$

where T_i is the current timestamp of the user U_i.

S4. The smart card sends the login request message $\{AID_i, D_i, E_i, F_i, T_i\}$ to the server S.

5.3 Authentication Phase

Upon completing this phase, the user U_i and the server S can mutually authenticate each other and establish a shared session key for the subsequent secret communication. These steps of the authentication phase are shown as follows:

S1. The server S verifies whether $T_i' - T_i \leq \triangle T$, where T_i' is the time of receiving the login request message and $\triangle T$ is a valid time threshold. If both conditions are true, the server S continues to execute Step 2; otherwise, the server S rejects the login request.

S2. The server S computes the following parameters:

$$\begin{aligned} A_i' &= D_i \oplus x \oplus h(x), \\ ID_i' &= AID_i \oplus h(A_i'), \\ F_i' &= h(ID_i'||h(A_i')||E_i||T_i). \end{aligned}$$

The server S then compares whether F_i' is equals F_i. If this holds, the server S confirms that the user U_i is valid and the login request is accepted; otherwise, the server S rejects the login request.

S3. Next, the server S randomly generates a number β and computes the following parameters:

$$\begin{aligned} F_i &= \beta P, \\ G_i &= h(ID_i'||h(A_i')||F_i||T_s), \end{aligned}$$

where T_s is the current timestamp of the server S.

S4. The server S sends the mutual authentication message $\{F_i, G_i, T_s\}$ to the user U_i.

S5. Upon receiving the message $\{F_i, G_i, T_s\}$ from the S, the user U_i checks the validity of the T_s. If $T_s' - T_s \leq \triangle T$, where T_s' is the time of receiving the mutual authentication message, the user U_i continues to perform Step 6; otherwise, the user U_i terminates this connection.

S6. The user U_i computes $G_i' = h(ID_i||h(A_i)||F_i||T_s)$, then checks whether G_i' is equal to the received G_i. If this holds, the validity of the server S is authenticated; otherwise, the session is terminated.

S7. Finally, the user U_i and the server S construct a shared session key $sk = \alpha\beta P$ to ensure the secret communication.

5.4 Password Change Phase

During the password change phase, U_i updates the password without any assistance from server S_j. This phase consists of the following steps:

S1. The U_i enters the identity ID_i and password PW_i, and imprints the biometrics BIO_i^* at the sensor. The sensor then scans BIO_i^*, and recovers R_i from $Rep(BIO_i^*, P_i) \rightarrow R_i$.

S2. Next, the SC_i calculates $RPW_i = h(PW_i||R_i)$, and checks whether $h(ID_i||RPW_i)$ is equal to the stored C_i. If this holds, the smart card asks the U_i for a new password; otherwise, the SC_i immediately terminates the password change phase.

S3. The U_i inputs new password PW_i^{new}, and the smart card further computes $RPW_i^{new} = h(PW_i^{new}||R_i)$, $B_i^{new} = B_i \oplus RPW_i \oplus RPW_i^{new}$ and $C_i^{new} = C_i \oplus RPW_i \oplus RPW_i^{new}$.

S4. Finally, the smart card replaces B_i with B_i^{new} and C_i with C_i^{new} in memory.

6 Security Analysis of The Proposed Scheme

In this section, we demonstrate that the proposed scheme, which retains the merits of Liu et al.'s scheme, can withstand several types of possible attacks; and we also show that the proposed scheme supports several security properties. The security analysis of the proposed scheme was conducted with the threat assumptions made in the Section 2.

6.1 Resisting Outsider Attack

The outsider is the person who has registered with the server S, not the person who is not the user of the system. Suppose an outsider adversary $\mathcal{A}$ extracts all of the information $\{B_{\mathcal{A}}, C_{\mathcal{A}}, D_{\mathcal{A}}, h(\cdot), P, P_{\mathcal{A}}\}$ from own smart card by side channel attack [12]. However, he/she cannot obtain any secret information of the S. The $\mathcal{A}$ can compute $h(A_{\mathcal{A}}) = B_{\mathcal{A}} \oplus RPW_{\mathcal{A}}$. However, the value x is a master secret key stored in the server S. Therefore, $\mathcal{A}$ does not know, and the proposed scheme can resist outsider attack.

6.2 Resisting Insider Attack

The insider is the authorized users of the server who access data or resources. In the registration phase, the user conceals the password in a ciphertext from the server S to resist an insider attack. More specifically, the user U_i first selects their password PW_i, and then submits $RPW_i = h(PW_i||R_i)$ to the server S for the registration over a secure channel. The S cannot retrieve the password PW_i or biometrics BIO_i from $RPW_i = h(PW_i||R_i)$. In addition, the S does not store RPW_i in the database. The proposed scheme therefore can resist insider attack.

6.3 Resisting User Impersonation Attack

Suppose an adversary $\mathcal{A}$ intercepts all of the message $\{D_i, E_i, F_i, T_i, F_i, G_i, T_s\}$ transmitted in public channel between the U_i and S, and steals the smart card of the U_i, then extracts the all of the information $\{B_i, C_i, D_i, h(\cdot), P, P_i\}$. However, the $\mathcal{A}$ cannot generate the legal login request message $\{AID_i, D_i, E_i, F_i, T_i\}$, where $AID_i = B_i \oplus h(A_i)$, $D_i = x \oplus A_i \oplus h(x)$, $E_i = \alpha P$, and $F_i = h(ID_i||h(A_i)||E_i||T_i)$. This is because the value $h(A_i)$ is protected by $RPW_i = h(PW_i||R_i)$, and the A_i is protected by server S's master key x. The user U_i's password is protected by collision resistant one-way hash function, such as, $h(PW_i||R_i)$, where R_i possesses high entropy. Moreover, there is no the same biometric templates between any two people. The $\mathcal{A}$ cannot generate the mutual authentication message $\{F_i, G_i, T_s\}$ without the value $h(A_i)$. The proposed scheme can therefore resist user impersonation attack.

6.4 Providing Perfect Forward Secrecy

The perfect forward secrecy means that if one of long-term keys is compromised, a session key which is derived from these long-term keys will not be compromised in the future [27]. In the proposed scheme, a session key between the user U_i and server S is calculated as follows:

$$\begin{aligned} E_i &= \alpha P, \\ F_i &= \beta P, \\ sk &= \alpha\beta P. \end{aligned}$$

Even if the server S's long-term key x is compromised, the adversary $\mathcal{A}$ cannot retrieve α and β to generate the session keys between the U_i and S. The session key of the proposed scheme is based on elliptic curve discrete logarithm problem. The adversary $\mathcal{A}$ cannot obtain $\alpha\beta P$ from the αP and βP. Above all, the scheme achieves the perfect forward secrecy.

7 Performance Analysis

In this section, we compare the functionality between the proposed scheme and the other recent schemes [2, 10, 15, 16, 17, 23, 25]. From Table 4, we can see that all of other existing schemes involve some time-consuming operations, such as modulus exponential operations, symmetric encryption/decryption operations, multiplication/division operations and scalar multiplication. From this comparison, we can see that the hash operation costs of the proposed scheme are slightly lower than the authentication scheme by Liu et al., and the proposed scheme performs four further scalar multiplication functions and four fuzzy-extraction functions than Liu et al.'s scheme to accomplish mutual authentication and key agreement; however, the proposed scheme archives perfect forward secrecy.

Table 3: Functionality comparison of the proposed scheme and other related schemes

	F1	F2	F3	F4	F5	F6	F7	F8	F9
Juang et al. [10]	○	○	○	○	○	○	×	○	○
Sun et al. [25]	○	○	×	○	○	○	○	○	○
Li et al. [16]	○	○	○	○	○	○	○	○	○
Song [23]	○	○	○	○	×	○	×	○	○
Chen et al. [2]	○	○	○	○	×	○	×	○	○
Li et al. [15]	○	○	○	×	×	○	○	○	○
Liu et al. [17]	○	○	○	○	○	○	×	○	×
The proposed	○	○	○	○	○	○	○	○	○

F1: Mutual authentication; F2: Session key agreement; F3: Freely chosen and exchanged password; F4: Withstanding man in the middle attack; F5: Withstanding insider attack; F6: Withstanding replay attack; F7: Providing perfect forward secrecy; F8: Satisfying known-key security; F9: User impersonation attack.

Table 4: Computational cost comparison of the proposed scheme and other related schemes

	C1	C2	C3	C4	C5	C6	Total
Juang et al. [10]	1H	2H+3S	3H+3S	4H+6S+1M	1H+5S	1H+5S	12H+22S+1M
Sun et al. [25]	-	2H+1S	4H+2M	4H+1S+2M	2H	-	12H+2S+4M
Li et al. [16]	1H	2H+3S	8H+4S	10H+10S+1M	1H+6S	1H+9S	23H+32S+1M
Song [23]	-	2H+1E	3H+1S	3H+1S+1E	-	-	8H+2S+2E
Chen et al. [2]	-	1H+1E	2H+2M+4E	1H+1M+4E	3H+2M+2E	3H+2M+3E	10H+14E+7M
Li et al. [15]	-	2H+2E	4H+1M+4E	3H+3E	3H+2M+4E	-	12H+3M+13E
Liu et al. [17]	1H	3H	6H	6H	4H	-	20H
The proposed	1H+1F	4H	4H+1F+2P	3H+1F+2P	3H+1F	-	15H+4F+4P

C1: Computational cost of the user in registration phase; C2: Computational cost of the server in registration phase; C3: Computational cost of the user in login and authentication phases; C4: Computational cost of the server in login and authentication phases; C5: Computational cost of the user in password change phase; C6: Computational cost of the server in password change phase; H: Hashing operation; E: Modulus exponential operation; S: Symmetric encryption/decryption operation; M: Multiplication/division operation; Null: P: Scalar multiplication; F: Fuzzy extraction; Null: Cannot provide this functionality.

8 Conclusions

In this paper, we proposed a biometrics-based user authentication scheme using smart card to overcome the security weaknesses of Liu et al.'s scheme. The proposed scheme can achieve mutual authentication and perfect forward secrecy, and users can freely choose and change their password. We prove that the proposed scheme can resist various attacks, such as the outsider attack and user impersonation attack.

Acknowledgments

This work was supported by Institute for Information & communications Technology Promotion(IITP) grant funded by the Korea government(MSIP) (No.R0126-15-1111, The Development of Risk-based Authentication·Access Control Platform and Compliance Technique for Cloud Security)

References

[1] N. Anwar, I. Riadi, A. Luthfi, "Forensic SIM card cloning using authentication algorithm," *International Journal of Electronics and Information Engineering*, vol. 4, no. 2, pp. 71–81, 2016.

[2] B. L. Chen, W. C. Kuo and L. C. Wuu, "Robust smart-card-based remote user password authentication scheme," *International Journal of Communication Systems,* vol. 27, no. 2, pp. 377–389, 2014.

[3] Y. Choi, Y. Lee and D. Won, "Security improvement on biometric based authentication scheme for wireless sensor networks using fuzzy extraction," *International Journal of Distributed Sensor Networks,* vol. 2016, pp. 1–16, 2016.

[4] A. Das, "A secure and effective biometric-based user authentication scheme for wireless sensor networks using smart card and fuzzy extractor," *International Journal of Communication Systems,* vol. 30, no. 1, pp. 1–25, 2017.

[5] Y. Dodis, B. Kanukurthi, J. Katz, L. Reyzin and A. Smith, "Robust fuzzy extractors and authenticated key agreement from close secrets," *IEEE Transactions on Information Theory,* vol. 58, no. 9, pp. 6207–6222, 2012.

[6] Y. Dodis, L. Reyzin and A. Smith, "Fuzzy extractors:How to generate strong keys from biometrics and other noisy data," *Advances in Cryptology (EUROCRYPT'04),* LNCS 3027, pp. 523–540, Springer, 2004.

[7] D. Dolev and A. Yao, "On the security of public key protocols," *IEEE Transactions on Information Theory,* vol. 29, no. 2, pp. 198–208, 1983.

[8] D. Hankerson, A. Menezes, and S. Vanstone, *Guide to Elliptic Curve Cryptography*, Springer, Berlin, 2004.

[9] W. Jeon, J. Kim, J. Nam, Y. Lee and D. Won, "Two-round password-only authenticated key exchange in the three-party setting," *IEICE Transactions on Communications,* vol. 95, no. 5, pp. 1819–1821, 2012.

[10] W. S. Juang, S. T. Chen and H. T. Liaw, "Robust and efficient password-authenticated key agreement using smart card," *IEEE Transactions on Industrial Electronics,* vol. 55, no. 6, pp. 2551–2556, 2008.

[11] N. Koblitz, "Elliptic curve cryptosystems," *Mathematics of Computation*, vol. 48, no. 117, pp. 203–209, 1987.

[12] P. Kocher, J. Jaffe, B. Jun and P. Rohatgi, "Introduction to differential power analysis," *Journal of Cryptographic Engineering,* vol. 1, no. 1, pp. 1–23, 2011.

[13] A. V. N. Krishna, A. H. Narayana, K. M. Vani, "Window method based cubic spline curve public key cryptography," *International Journal of Electronics and Information Engineering*, vol. 4, no. 2, pp. 94–102, 2016.

[14] L. Lamport, "Password authentication with insecure communication," *Communications of the ACM,* vol. 24, no. 11, pp. 770–772, 1981.

[15] X. Li, J. Niu, M. K. Khan, and J. Liao, "An enhanced smart card based remote user password authentication scheme," *Journal of Network and Computer Applications,* vol. 36, no. 5, pp. 1365–1371, 2013.

[16] X. Li, W. D. Qiu, D. Zheng, K. F. Chen, and J. H. Li, "Anonymity enhancement on robust and efficient password-authenticated key agreement using smart cards," *IEEE Transactions on Industrial Electronics,* vol. 57, no. 2, pp. 793–800, 2010.

[17] Y. Lin, 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.

[18] V. Miller, "Use of elliptic curves in cryptography," in *Advances in Cryptology (CRYPTO'85)*, vol. 218, pp. 417–426, 1985.

[19] J. Moon, Y. Choi, J. Jung and D. Won, "An improvement of robust biometrics-based authentication and key agreement scheme for multi-server environments using smart cards," *Plos One,*, vol. 10, no. 12, pp. 1–15, 2015.

[20] J. Moon, Y. Choi, J. Kim and D. Won. "An improvement of robust and efficient biometrics based password authentication scheme for telecare medicine information systems using extended chaotic maps," *Journal of Medical Systems*, vol. 40, no. 3, pp. 1–11, 2016.

[21] J. Moon, J. Kim and D. Won, "An improvement of user authentication framework for cloud computing," *Journal of Conputers,* vol. 11, no. 6, pp. 446–454, 2016.

[22] J. Nam, K. K. R. Choo, S. Han, J. Paik and D. Won, "Two-round password-only authenticated key exchange in the three-party setting," *Symmetry,* vol. 7, no. 1, pp. 105–124, 2015.

[23] R. Song, "Advanced smart card based password authentication protocol," *Computer Standards and Interfaces*, vol. 32, no. 5, pp. 321–325, 2010.

[24] S. K. Sood, A. K. Sarje, and K. Singh, "An improvement of Xu et al.'s authentication scheme using smart cards," in *Proceedings of the Third Annual ACM Bangalore Conference,* pp. 17–22, Bangalore, Karnataka, India, 2010.

[25] D. Z. Sun, J. P. Huai, J. Z. Sun, J. X. Li, J. W. Zhang, and Z. Y. Feng, "Improvements of Juang et al.'s password-authenticated key agreement scheme using smart cards," *IEEE Transactions on Industrial Electronics*, vol. 56, no. 6, pp. 2284–2291, 2009.

[26] C. Wang, X. Zhang and Z. Zheng, "Cryptanalysis and improvement of a biometric-based multi-server authentication and key agreement scheme," *Plos One,* vol. 11, no. 2, pp. 1–25, 2016.

[27] H. Zhu and X. Hao, "A provable authenticated key agreement protocol with privacy protection using smart card based on chaotic maps," *Nonlinear Dynamics,* vol. 81, no. 1, pp. 311–321, 2015.

Biography

Jongho Moon received the B.S. degree in electrical and computer engineering from Sungkyunkwan University, Suwon, Korea, in 2012 and the M.S. degree in electrical and computer engineering from Sungkyunkwan University, Suwon, Korea, in 2014. He also worked as a malware analyzer in SECUI between 2014 and 2015. He is currently pursuing the Ph.D. degree in electrical and computer engineering at Sungkyunkwan University,

Suwon, Korea. His current research interest includes cryptography, malware, forensic, and authentication or key management protocols.

Donghoon Lee received the B.S. degree in Computer Science from National Institute for Lifelong Education(NILE), Korea, in 2009 and the M.S. degree in Information Security Engineering from Sungkyunkwan University, Korea, in 2011. He is currently undertaking a Ph.D. course on Electrical and Computer Engineering in Sungkyunkwan University. His current research interest is in the area of software security, cryptography, authentication protocol, and network security.

Jaewook Jung received the B.S. degree in Electrical and Computer Engineering from Korea Aerospace University, Goyang, Korea, in 2010 and the M.S. degree in Electrical and Computer Engineering from Sungkyunkwan University, Suwon, Korea, in 2012. He is currently undertaking a Ph.D. course on Electrical and Computer Engineering in Sungkyunkwan University, Suwon, Korea. His current research interest is in the area of cryptography, forensic, authentication protocol, and mobile security.

Dongho Won received B.S., M.S. and Ph.D. in Electronic Engineering from Sungkyunkwan University, Suwon, Korea. After working in Electronics and Telecommunication Research Institute for two years, he joined Sungkyunkwan University, where he is currently a leader professor at Information and Communication Engineering. He also served as a President of Korea Institute of Information Security and Cryptography. His research interests are cryptology and information security.

Clustering Based K-anonymity Algorithm for Privacy Preservation

Sang Ni[1], Mengbo Xie[1], Quan Qian[1,2]
(Corresponding author: Quan Qian)

School of Computer Engineering & Science, Shanghai University[1]
Materials Genome Institute of Shanghai University[2]
No. 99, Shangda Road, Shanghai, China
(Email: qqian@shu.edu.cn)

(Received July 31, 2016; revised and accepted Nov. 5, 2016 & Jan. 15, 2017)

Abstract

K-anonymity is an effective model for protecting privacy while publishing data, which can be implemented by different ways. Among them, local generalization are popular because of its low information loss. But such algorithms are generally computation expensive making it difficult to perform well in the case of large amount of data. In order to solve this problem, this paper proposes a clustering based K-anonymity algorithm and optimizes it with parallelization. The experimental result shows that the algorithm performs better in information loss and performance compared with the existing KACA and Incognito algorithms.

Keywords: Clustering based K-anonymity; Information Loss; Privacy Preservation

1 Introduction

Along with the development of computer network, distributed computing, data mining and big data, huge amounts of data can be collected and analyzed efficiently. But when we explore the potential value of large amounts of data, privacy and privacy protection would be the focusing point. According to Manish Sharma [21], the secondary use of data is a source of privacy disclosure, which is the use of data for some purpose other than the purpose for which the data was collected initially. Jisha also noted that privacy is being violated mainly through three types of attack, such as linking attack, homogeneity attack and background knowledge attack [16]. Therefore, it is a very important issue to pay equal attention to data secondary using, data misusing, data mining and privacy preservation.

Privacy preservation is tightly associated with database security. Database security is usually achieved by means of access control, security management and database encryption [24]. Access control is a selective policy for restricting unauthorized users to access a resource through the user permissions. Security management refers to what kind of security management mechanisms are used to distribute database management authorities. Centralized control and decentralized control are 2 typical modes. Database encryption mainly includes three aspects: record encryption, database structure encryption and hardware encryption. These measures can protect the security of the database to a certain extent, for example, the direct disclosure of sensitive information such as, identification card number, home address, health information etc. But they are unable to prevent those indirect accesses to private data through federation reasoning. In [5], it shows that through joining voter registration table and medical information table (individual identification is hidden), by attributes of Zip code, Sex, Date of Birth, etc., more than 85% of American citizens can be uniquely identified. In addition, encryption and access control, to some extent, limits the sharing of data.

For such reasons, data anonymity is an effective means to achieve privacy preservation. The basic idea is to transform some part of the original data, for instance, through generalization, compression, etc., and let the transformed data cannot be combined with other information to reason about any personal privacy information. Specifically, the implementation of privacy preservation mainly concentrates on two aspects: (1) How to ensure that the data been used without privacy disclosure? (2) How to make the data to be better utilized? Therefore, a better trade off between privacy preservation and data utilization is a problem that the academia and industry need to be solved urgently.

K-anonymity was first proposed in 1998 by Sweeney et al. [23]. K-anonymity depends on anonymizing the original data set to satisfy the anonymization requirements, which can be used for data publishing. The common anonymization techniques are generalization and hidden. The basic idea of K-anonymity is anonymizing the publishing data to meet the requirement that at least K

tuples cannot be distinguished by each other. Namely, for each tuple there exists at least K tuples with equal value of quasi-identifiers. Researchers have proved that the complexity of K-anonymity is NP-hard [20].

Currently, there are many algorithms to implement K-anonymity [17]. From the point of generalization, can be categorized as recoding mode (global recoding, local recoding), data grouping strategy (classification, clustering, and Apriori algorithm). From the perspective of the data characteristics, they are static data set and dynamic data set (incremental data, stream data, and uncertain data).

The rest of the paper is organized as follows: Sections 2 and 3 discuss the related work and some prerequisite knowledge of k-anonymity. Our main contributions, the clustering based k-anonymity algorithm *GCCG* and its parallel optimization are described in Sections 4 and 5. The experimental results are shown in Section 6. Section 7 provides some final conclusions and directions for future work.

2 Related Work

Privacy preservation was firstly concerned in the field of statistics and then extended to various areas. The main research directions are as shown in Table 1.

Table 1: Main research directions of privacy preservation [17]

Research direction	Relevant techniques
General privacy protection technique	Data perturbation, randomization, data exchange, data encryption, etc.
Privacy protection technology for data mining	association rules mining, classification, clustering etc.
Privacy preservation based data publishing principles	K-anonymity, M-invariant, T-closeness, etc.

2.1 Different Kinds Of Anonymity Algorithms

Anonymity methods mainly include generalization, classification and clustering etc. [17].

Generalization and suppression based anonymity. The main idea of generalization based anonymity is to increase equivalence class size of the table by reducing the data precision of the quasi-identifier attributes. Generally, quasi-identifier can be divided into two kinds: numeric attribute and category attribute. Numeric attributes are usually generalized to interval, for instance, the age 16 can be generalized to interval $[10-20]$. And for category attributes, the original concrete values will be replaced by more general ones, according to a priori established VGH(value generalization hierarchies). For example, a nationality attribute whose value is "China", then it can be generalized to "Asia". Suppression can be viewed as an extreme form of generalization, in which the generalized attributes cannot be further generalized [6]. Kameya et al. proposed a cell-suppression based k-anonymization method which aims to preserve the MAR(Missing at random) condition uses the Kullback-Leibler (KL) divergence as a utility measure [28]. Besides, He et al. proposed a novel linking-based anonymity model, which can resist the attack incurred by homogeneous generalization [7].

Dividing and grouping based anonymity. According to the different ways of dividing, it can be divided into micro aggregation, condensation, anatomy and permutation, etc.

- Micro aggregation divides each class according to the data similarity, making the tuple size of a class at least k, and then using the class centroid to generalize all the data of a class. Mortazavi et al. proposed a Fast Data-oriented micro aggregation algorithm (FDM)in [18] that efficiently anonymizes large multivariate numerical datasets for multiple successive values of k and proposed a disclosure-aware aggregation model in [19], where published values are computed in a given distance from the original ones to obtain a more protected and useful published dataset.
- Condensation[1, 2, 3] is a new kind of method similar to micro aggregation, which was proposed by Aggarwal et al. in 2004. The basic idea is to divide the original data into different groups, and then process each group with condensation technique. Condensated data will be reconstructed by general reconstruction algorithm, which would not reveal any privacy information of the original tuples.
- Anatomy method was firstly put forward by Xiao et al. [25]. It publishes the sensitive attributes and quasi-identifier attributes separately to reduce the correlation degree between them. Permutation depends on disturbing the order of sensitive attributes after grouping to reduce the correlation between quasi-identifiers and numeric sensitive attributes. In [30], Yu et at. proposed a novel anonymization method based on anatomy and reconstruction in LBS privacy preservation.

Clustering based anonymity. Anonymity can also be implemented by clustering, which is the most commonly used method. The basic idea is to produce at

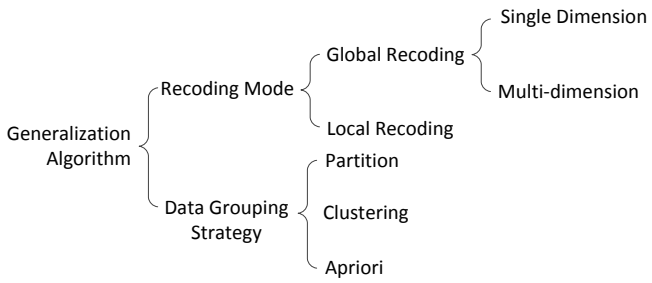


Figure 1: Generalization based anonymity algorithm [5]

least k records of a class as the equivalence class. The tuples in a same class need to be as similar as possible to make the information loss minimum after generalization. Liu et al. proposed a privacy-preserving data publishing method, namely MNSACM, which uses the ideas of clustering and Multi-Sensitive Bucketization (MSB) to publish microdata with multiple numerical sensitive attributes [14]. Bhaladhare et al. proposed two approaches for minimizing the disclosure risk and preserving the privacy by using systematic clustering algorithm [4]. Xin et al. proposed a trajectory privacy preserving method based on the adaptive clustering, and designed a 2-stage clustering method for trajectory k-anonymity [26].

2.2 Generalization Based Anonymity

Currently, there are many algorithms to implement k-anonymity, and most of them use the generalization and suppression as shown in Figure 1.

From the perspective of generalization methods, it can be divided into 2 categories: global generalization and local generalization.

Global generalization. This kind of algorithms allows the whole domain of identifier attributes mapped to a generalization domain, that is to say, a value in a table would only have one generalization value. In general, global generalization algorithm is simple and efficient. But it needs to set generalization level in advance and has problems of over generalization causing high information loss. Representative global generalization algorithms include: *u-Argus* algorithm [25], Datafly [22], Incognito [10], etc.

u-Argus algorithm proposed by De.Wall et al. [8], which the published data includes all tuples and all properties of the initial data, except very few data will be lost. But if there exists many attribute combinations, it can not provide enough protection for released data. Datafly proposed by Sweeney uses suppression and heuristic generalization, which is efficient but has much distortion. LeFevre proposed Incognito, which adopts the Generalization Graph for data generalization with the bottom-up approach. It prunes redundant branches of the graph to narrow the search scope. However, the information distortion of Incognito is relatively high.

Local generalization. This kind of algorithm usually maps attribute value to generalization value based on the grouping , that is to say, even the same attribute values can be generalized to different values if they are in the different groups. Grouping data usually adopts some heuristic principles, such as division, clustering and so on. Information loss of this kind of algorithm is less than the global generalization algorithms, but its complexity usually is higher. In the case of a large amount of data, the performance is a problem to be concerned. Representative algorithms are: GA [9], Mondrian Multidimensional algorithm [11], $KACA$ algorithm [12] as follows.

Iyengar proposed GA (Genetic algorithm), which can meet the requirements of K-anonymity, but when processing large amount of data, it will spend a few hours. Mondrian multidimensional algorithm proposed by $Le.Fevre$ can partition continuous attributes but not for discrete attributes. Li put forward $KACA$ algorithm, through merging the nearest equivalence class to form a bigger cluster. Although, $KACA$ has low information loss, the performance is poor because of massive distance computations.

To sum up, it concludes that, generally, the global generalization algorithms are efficient, but the information quality is low. On the contrary, local generalization algorithms can greatly improve the information quality, but it is often inefficient. So, in this paper our motivation is to propose an efficient local generalization algorithm which has great information quality and great performance simultaneously.

3 Prerequisite Knowledge

To use K-anonymity, supposing the original data are stored in database with the form of structured table. We assume that data publishers have the raw data table T, each row in the table is corresponding to a specific entity, such as student id, name, gender, birth place, etc. As shown in Table 2, each row in the table is so called a tuple.

Definition 1. *Quasi-identifier. Quasi-identifier is some form of attributes combination, which can determine some individuals in table T by joining some external information.*

Theoretically, in a table, all attributes except identifiers, can be quasi-identifiers.

Definition 2. *Sensitive identifier. Sensitive identifiers are those attributes concerning sensitive privacy information, such as salary, health information, etc.*

Table 2: An example of data to be anonymized

Education	Workclass	Race	Sex	Age
Bachelors	State-gov	White	Male	39
Bachelors	Self-emp-not-inc	White	Male	50
HS-grad	Private	White	Male	38
11th	Private	Black	Male	53
Bachelors	Private	Black	Female	28
Masters	Private	White	Female	37
9th	Private	Black	Female	49
HS-grad	Self-emp-not-inc	White	Male	52

Since attribute sensitivity is context-dependent, so it is not invariable and should be configured manually according to actual situations. In this paper, taking Table 2 as an example, we set attribute *workclass* as a sensitive identifier.

Definition 3. *Equivalence class. The Equivalence class of table T on attribute $A_i, \cdots, A_j$ is the set of tuples that all values of these attributes are identical.*

For example, in Table 1, the top 2 rows: $\{Bachelors, State-gov, White, Male, 39\}$ and $\{Bachelors, Self-emp-not-inc, White, Male, 50\}$ are the same equivalence class on attribute $\{Sex, Education, Race\}$.

Definition 4. *K-anonymity Property. Generate several equivalence classes on quasi-identifiers. If each size of the equivalence class is no less than K, we can say the equivalence class partition has k-anonymity property.*

That is to say, according to the quasi-identifiers, each record has at least $(k-1)$ other same records to make them unable to be identifier by each other. Therefore, we can say the Table 2 is a 2-anonymity result of Table 1.

K-anonymity adopts generalization and suppression to preserve privacy. So there will be a certain degree of information loss inevitably. In order to describe quantitatively, it is necessary to introduce a corresponding measurement for information loss. There are many different measurement models for information loss, including Prec, DM, NE, etc. In this paper we use the measurement method in [27]. In order to compare the result under different amount of data, the sum of the original formula is modified as the mean value. Such a change won't change the relationship of size between results.

Definition 5. *Information loss. Supposing that a numeric attribute in a tuple, the original value x is generalized to $[x_{min}, x_{max}]$, where x_{min} is the minimum of the equivalence class and x_{max} is the maximum of the equivalence class. Max and Min is the maximum and minimum value of the attribute in the whole domain. Then the information loss (IL) of the tuple on the numeric attribute is defined as Equation (1) .*

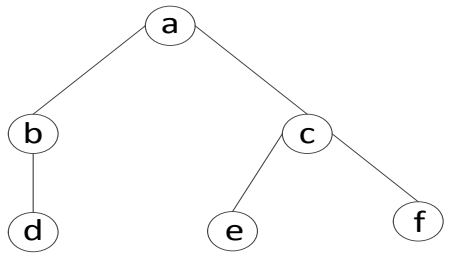


Figure 2: An classification tree example for a category attribute

$$IL = \frac{x_{max} - x_{min}}{Max - Min}. \quad (1)$$

For a category attribute, we usually need to build a classification tree at first. As shown in Figure 2, it is a classification tree example for a category attribute. Supposing that the value of a tuple is generalized from e to c. Then the information loss of the tuple on the attribute is defined as Equation (2).

$$IL = \frac{size(c)}{Size}. \quad (2)$$

"size(c)" is the number of its descendant leaf nodes and $Size$ is the number of all leaf nodes. Therefore, in Figure 2, the information loss is 2/3, in that the total number of leaf nodes is 3, and node c's descendant leaf is 2. For all the attributes of a tuple, its information loss is defined as Equation (3), where m is the number of all attributes.

$$IL = \frac{\sum_{i=1}^{m} IL_i}{m}. \quad (3)$$

Finally, the average IL of all the tuples is the information loss of the whole data set after generalization.

4 GCCG: Clustering Based K-anonymity

The key point of K-anonymity is to produce a number of equivalence classes whose size is at least k and each equivalence class has the same form on quasi-identifiers. This idea is very similar to clustering. Each equivalence class can be regarded as a cluster, and at the same time, the centroid of a cluster can be seem as a generalization form of a equivalence class. Next, we will take the data set in Table 2 as a 2-anonymity example to explain the GCCG algorithm in detail.

Table 3: An example of anonymized data after k-anonymity (k=2)

Education	Workclass	Race	Sex	Age
Bachelors	State-gov	White	Male	[39 - 50]
Bachelors	Self-emp-not-inc	White	Male	[39 - 50]
HS-grad	Private	White	Male	[38 - 52]
HS-grad	Self-emp-not-inc	White	Male	[38 - 52]
Low	Private	Black	*	[49 - 53]
Low	Private	Black	*	[49 - 53]
High	Private	*	Female	[28 - 37]
High	Private	*	Female	[28 - 37]

4.1 Algorithm Overview

There are four main steps of our clustering based K-anonymity algorithm: *Grading*, *Centering*, *Clustering*, and *Generalization*, abbreviated by *GCCG*. The pseudo code of *GCCG* is as Algorithm 1, and we will explain each step in detail.

Algorithm 1 GCCG Algorithm

1: Input: Dataset D (with n records), Anonymity constant K, Classification tree for each attribute.
2: Output: D's k-anonymity result
3: Begin
4: **for** $i = 1$ to n **do**
5: Grade each tuple by the cluster centroid;
6: **end for**
7: Sort D by the centroid grading score;
8: **for** $i = 1$ to $n/k - 1$ **do**
9: Choose the first tuple to be the clustering centroid;
10: Choose the centroid and the nearest $(k-1)$ tuples to make up a new equivalence class;
11: Remove the k tuples from D;
12: **end for**
13: Make the rest tuples to be the last equivalence class;
14: **for** each equivalence class **do**
15: Generalize the tuples using the class centroid;
16: **end for**
17: End

4.2 Grading Tuples Using Cluster Centroid

Since the cluster centroid will be used for equivalence class generalization. So, the clustering quality will affect anonymity greatly. If the tuples in a cluster are much scattered, then we need a more general value to generalize them, thus resulting in much information loss. So in this paper, we propose an evaluation method to find an appropriate clustering centroid. The method can be divided into two kinds according to different attributes: category or numeric.

Table 4: Original data table

Education	Workclass	Race	Sex	Age
Bachelors	State-gov	White	Male	39
Bachelors	Self-emp-not-inc	White	Male	50
HS-grad	Private	White	Male	38
11th	Private	Black	Male	53
Bachelors	Private	Black	Female	28
Masters	Private	White	Female	37
9th	Private	Black	Female	49
HS-grad	Self-emp-not-inc	White	Male	52
Masters	Private	White	Female	31
Bachelors	Private	White	Male	42

For a category attribute, the score is the ratio of the count of the attribute value to the whole attribute values in dataset. Assuming that the ratio is P1, then P1 is regared as the attribute score of the tuple. That means more frequent a value, more possibility for it to be a center. For a numeric attribute, we take the proportion of the ratio of the value to K as the score. Finally, the sum of all attributes is the score of the tuple on center grading. Table 5 is result of sorted data set of Table 4 by the score.

4.3 Tuples Distance Definition And Calculation

After selecting the cluster centroid, distance computation among tuples is another key problem for clustering. Not as KACA (another typical clustering based generalization algorithm for k-anonymity), it uses iterative generalization and its efficiency is low [29]. So, in our algorithm, we just calculate the distances among tuples, which simplifies the distance calculation. In this paper, the distance

Table 5: Dataset sorted by score

Education	Workclass	Race	Sex	Age	Score
Bachelors	State-gov	White	Male	39	1.8
Bachelors	Self-emp-not-inc	White	Male	50	1.8
Bachelors	Private	White	Male	42	1.8
HS-grad	Private	White	Male	38	1.6
HS-grad	Self-emp-not-inc	White	Male	52	1.6
Masters	Private	White	Female	37	1.4
Masters	Private	White	Female	31	1.4
Bachelors	Private	Black	Female	28	1.2
11th	Private	Black	Male	53	1.1
9th	Private	Black	Female	49	0.9

between tuples is defined as follows:

For a numeric attribute A, supposing a_1, a_2 are the values of two tuples, then the distance on attribute A is:

$$dist = \frac{|a_1 - a_2|}{range}. \quad (4)$$

Where "range" is the difference between the maximum and minimum value on attribute A.

For a category attribute B, supposing b_1, b_2 are values of two tuples, then the distance on attribute B is:

$$dist = \frac{Parent\_depth(b_1, b_2)}{Depth}. \quad (5)$$

Where "Parent_depth(...)" is the subtree depth whose root is the nearest common ancestor of b_1 and b_2. "Depth" is the depth of the entire classification tree. For example, in Figure 2, the distance between e and f is 1/2.

According the distance definition above, Table 6 is the result after clustering. Each two rows is regarded as an equivalence class.

4.4 Generalization Procedure

Since *GCCG* algorithm is based on clustering, we generalize the tuples after being clustered. That is, use the cluster centroid to generalize them. Specific ways are as follows:

For a numeric attribute A, we use $[a_{min}, a_{max}]$ to generalize. a_{min} is the minimum value of A in the equivalence class and a_{max} is the maximum value. For a category attribute, we use the value of the nearest common ancestor in the classification tree to generalize. For both type of attributes, "*" is the most generic form, that is to say, all attribute information is removed. According to the method, the generalized result of Table 6 is shown as Table 7.

Table 6: Dataset afer clustering

Education	Workclass	Race	Sex	Age
Bachelors	State-gov	White	Male	39
Bachelors	Private	White	Male	42
Bachelors	Self-emp-not-inc	White	Male	50
HS-grad	Self-emp-not-inc	White	Male	52
HS-grad	Private	White	Male	38
Masters	Private	White	Female	37
Masters	Private	White	Female	31
Bachelors	Private	Black	Female	28
11th	Private	Black	Male	53
9th	Private	Black	Female	49

5 GCCG Parallel Optimization

To enhance the performance of *GCCG* algorithm, *GCCG* parallelization is necessary. Observing the whole algorithm, we can find that most of the operations are focusing on centroid selection and distance computation. So, if this part of operations can be parallelized, the performance of the whole algorithm will be improved.

In this paper, we use distances to divide the data set into a few small sub-datasets. That is, based on the original center selection method, the original data set is divided into n clusters (sub-datasets). Then all the sub-datasets do anonymity at the same time by GCCG with multithreading. During the clustering based data partition, we choose similar tuples into a same sub-dataset that reduces the information loss. Moreover, as the original times of distance computing is an arithmetic progression, after data partition, the distance computation

Table 7: Dataset after being generalized

Education	Workclass	Race	Sex	Age
Bachelors	State-gov	White	Male	[39 - 42]
Bachelors	Private	White	Male	[39 - 42]
*	Self-emp-not-inc	White	Male	[50 - 52]
*	Self-emp-not-inc	White	Male	[50 - 52]
*	Private	White	*	[37 - 38]
*	Private	White	*	[37 - 38]
High	Private	*	Female	[28 - 31]
High	Private	*	Female	[28 - 31]
Low	Private	Black	*	[49 - 53]
Low	Private	Black	*	[49 - 53]

Table 8: Result using parallel anonymity

Education	Workclass	Race	Sex	Age
Bachelors	State-gov	White	Male	[39 - 42]
Bachelors	Private	White	Male	[39 - 42]
*	Self-emp-not-inc	White	*	[37 - 50]
*	Private	White	*	[37 - 50]
*	Private	White	*	[37 - 50]
Senior high	Self-emp-not-inc	*	Male	[52 - 53]
Senior high	Private	*	Male	[52 - 53]
*	Private	*	Female	[28 - 49]
*	Private	*	Female	[28 - 49]
*	Private	*	Female	[28 - 49]

Algorithm 2 Dataset partition in parallel mode

```
Input: Dataset D (n records), anonymity constant k, parallel constant c;
Output: sub-dataset D'[c];
Begin
subsize=n/c;
while exists dataset's size > subsize do
  for each dataset with size more than subsize do
    Choose the first tuple to be the cluster center;
    Choose the nearest size/2-1 tuples to make up a sub-dataset with the new center;
    Make the rest tuples to be another sub-dataset;
  end for
end while
End
```

complexity can be reduced by the square of the number of sub-datasets. Table 8 is the result using 2-threads to do the anonymity for the data in Table 4.

6 Experimental Evaluation

The hardware used in the experiment is: Intel Xeon E5504 @ 2.00 GHz, 4G DDR3 Memory. Program implementation is Java7 and use Java7 Fork/Join multi-threading framework to do parallel programming [15]. Database: MySQL 5.6.18. The experimental data are Adult Database from UCI Machine Learning Repository [13]. After preprocessing, the dataset contains 30,661 tuples and 5 attributes. Table 9 provides a brief description of the dataset including 4 quasi- identifiers and 1 sensitive attribute.

6.1 Information Loss

The experiments in this paper are all implemented by JAVA under the same hardware environment. We choose Incognito and KACA as the algorithm to compare.

Figure 3 describes the information loss of the three algorithms when K changes from 3 to 10. It shows that the two local generalization algorithms have lower information loss than the global generalization algorithm. More-

Table 9: Adult dataset description

Attribute	Distinct values	Type	CTree Height
Age	72	Numeric	5
Sex	2	Category	2
Workclass	8	Sensitive	-
Race	5	Category	2
Education	16	Category	3

Table 10: Average size of equivalence classes

K	Incognito	KACA	GCCG
3	851.69	3.08	3.00
4	851.69	4.14	4.00
5	1533.05	5.12	5.00
6	1533.05	6.09	6.00
7	1533.05	7.24	7.00
8	1533.05	8.20	8.00
9	1533.05	9.22	9.00
10	1533.05	10.30	10.00

over, concerning the information loss, *GCCG* performs the best, about one third of the Incognito.

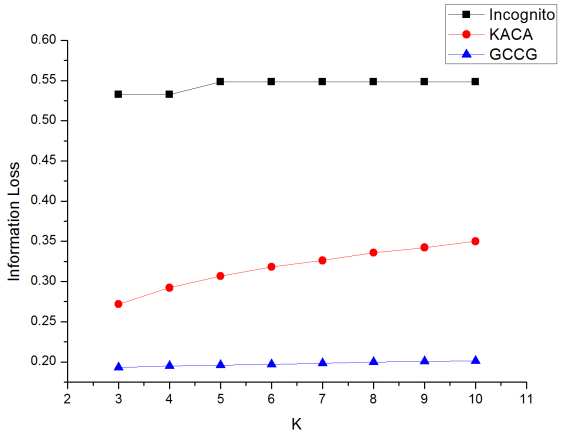


Figure 3: Information loss comparisons among different generalizations

Table 10 describes the average size of the equivalence classes of the three algorithms when K changes from 3 to 10. The result shows that the information loss is related to the size of equivalence class. When increasing the size of equivalence class, the information loss will increase appropriately. Therefore, controlling the size of the equivalences class is a effective way to control the information loss.

6.2 Execution Performance

Figure 4 describes the running time of the three algorithms when K changes from 3 to 10. It says that, as a global generalization algorithm, Incognito performs the best and $KACA$ consumes much time because of its frequent distance computation and cluster merging. $GCCG$, also belongs to local generalization algorithm, performs about 10 times faster than that of $KACA$ and is much close to Incognito.

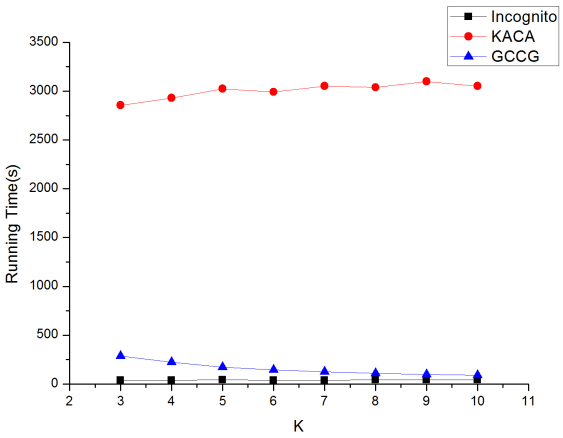


Figure 4: Execution performance comparison among different algorithms

6.3 Comparison Between Serial and Parallel Algorithm

Figure 5 shows the running time of serial $GCCG$ and parallel $GCCG$ when K changes from 3 to 10. From Figure 5, it shows that the performance of 2-threads parallelization improved about 3.5 times faster than that of the serial one, and 4-threads improved about 10.5 times. Moreover, be different from the serial algorithm, the running time of the parallel algorithm is not relevant to K, that is to say, the performance keeps relatively stable.

Figure 6 describes the information loss of serial $GCCG$ and parallel $GCCG$ when K changes from 3 to 10. From the result we can see that parallel $GCCG$ will have more information loss. But in the case of large amount of data, the growth is quite few.

7 Conclusions

In this paper, we proposed an clustering based local generalization algorithm $GCCG$ for K-anonymity. Experimental results show that $GCCG$ has lower information loss and better performance compared with the classical local generalization algorithms,$KACA$. More specifically, comparing with classical local generalization algorithm $KACA$, the information loss of $GCCG$ is about half of $KACA$, but with 10 times of performance improvement. Comparing with global generalization algo-

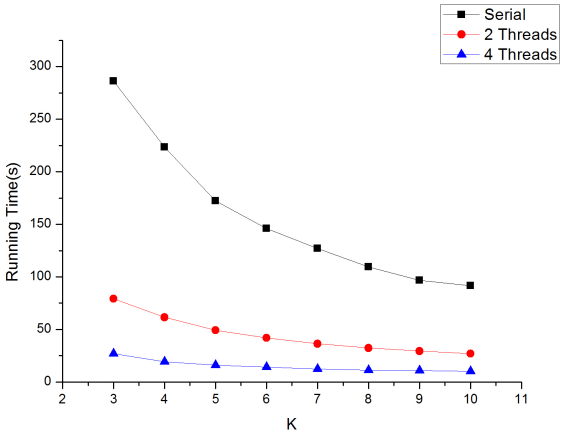


Figure 5: Performance comparison between serial and parallel *GCCG*

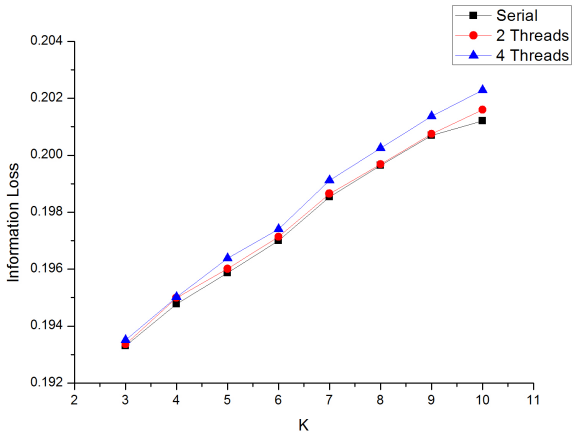


Figure 6: Information loss comparison between serial and parallel *GCCG*

rithm *Incognito*, the information loss of *GCCG* is about one third but with almost equal performance. Besides that, the parallel *GCCG* shows great performance improvement, when using 4-threads parallelization, there are 10 times acceleration with little information loss in the case of a large amount of data.

With the arrival of the era of big data, compared with the traditional data model, it is more likely to become the target of network attacks. Due to the system fault, hacker intrusion, internal leakage and other reasons, data leakage may occur at any time, resulting in unquantifiable losses. Therefore, aiming at big data oriented privacy protection issues deserves a enough attention with broad prospects.

Acknowledgments

This work is partially sponsored by National Key Research and Development Program of China(2016YFB0700504), Shanghai Municipal Science and Technology Commission(15DZ2260301), Natural Science Foundation of Shanghai(16ZR1411200). The authors gratefully appreciate the anonymous reviewers for their valuable comments.

References

[1] C. C. Aggarwal and P. S. Yu, "A condensation approach to privacy preserving data mining," in *Advances in Database Technology (EDBT'04)*, pp. 183–199, Heraklion, Crete, Greece, Mar. 2004.

[2] C. C. Aggarwal and P. S. Yu, "A framework for condensation-based anonymization of string data," *Data Mining & Knowledge Discovery*, vol. 16, no. 3, pp. 251–275, 2008.

[3] C. C. Aggarwal and P. S. Yu, "On static and dynamic methods for condensation-based privacy-preserving data mining," *ACM Transactions on Database Systems*, vol. 33, no. 1, pp. 41–79, 2008.

[4] P. R. Bhaladhare and D. C. Jinwala, "Novel approaches for privacy preserving data mining in k-anonymity model," *Journal of Information Science and Engineering*, vol. 32, no. 1, pp. 63–78, 2016.

[5] C. Clifton, M. Kantarcioglu, and J. Vaidya, "Defining privacy for data mining," in *National Science Foundation Workshop on Next Generation Data Mining*, pp. 126–133, Baltimore, MD, Nov. 2002.

[6] K. Dhivya and L. Prabhu, "Privacy preserving updates using generalization-based and suppression-based k-anonymity," *International Journal of Emerging Technology in Computer Science & Electronics*, vol. 8, no. 1, pp. 98–103, 2014.

[7] X. M. He, X. Y. Wang, D. Li, and Y. N. Hao, "Semi-homogenous generalization: Improving homogenous generalization for privacy preservation in cloud computing," *Journal of Computer Science and Technology*, vol. 31, no. 6, pp. 1124–1135, 2016.

[8] A. Hundepool and L. Willenborg, "Argus,software for statistical disclosure control," in *Proceedings of 13th Symposium on Computational Statistics*, pp. 341–345, Bristol, Great Britain, Aug. 1998.

[9] V. S. Iyengar, "Transforming data to satisfy privacy constraints," in *Proceedings of the 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'02)*, pp. 279–288, Edmonton, AB, Canada, July 2002.

[10] K. LeFevre, D. J. DeWitt, and R. Ramakrishnan, "Incognito: Efficient full-domain k-anonymity," in *Proceedings of the 2005 ACM SIGMOD International Conference on Management of Data (Sigmod'05)*, pp. 49–60, Baltimore, Maryland, USA, June 2005.

[11] K. LeFevre, D. J. DeWitt, and R. Ramakrishnan, "Mondrian multidimensional k-anonymity," in *22nd International Conference on Data Engineering (ICDE'06)*, pp. 25, Atlanta, Georgia, USA, Apr. 2006.

[12] J. Y. Li, R. C. Wong, A. W. Fu, and J. Pei, "Achieving k-anonymity by clustering in attribute

hierarchical structures," in *8th International Conference on Data Warehousing and Knowledge Discovery*, pp. 405–416, Krakow, Poland, Sept. 2006.

[13] M. Lichman, *UCI Machine Learning Repository*, 2013. (`http://archive.ics.uci.edu/ml`)

[14] Q. H. Liu, H. Shen, and Y. p. Sang, "Privacy-preserving data publishing for multiple numerical sensitive attributes," *Tsinghua Science and Technology*, vol. 20, no. 3, pp. 246–254, 2015.

[15] K. L. Nitin and S. Sangeetha, *Java 7 Fork/Join Framework*, 2012. (`http://www.developer.com/java-7-forkjoin-framework.html`)

[16] J. J. Panackal, A. S. Pillai, and V. N. Krishnachandran, "Disclosure risk of individuals: a k-anonymity study on health care data related to indian population," in *International Conference on Data Science & Engineering (ICDSE'14)*, pp. 200–205, Kochi, India, Aug. 2014.

[17] X. M. Ren, "Research for privacy protection method based on k-anonymity(in chinese)," Master Thesis, Harbin Engineering University, 2012.

[18] M. Reza and J. Saeed, "Fast data-oriented microaggregation algorithm for large numerical datasets," *Knowledge-Based Systems*, vol. 67, pp. 195–205, 2014.

[19] M. Reza and J. Saeed, "Enhancing aggregation phase of microaggregation methods for interval disclosure risk minimization," *Data Mining and Knowledge Discovery*, vol. 30, no. 3, pp. 605–639, 2016.

[20] P. Samarati and L. Sweeney, "Generalizing data to provide anonymity when disclosing information," in *Proceedings of the 17th ACM SIGACT-SIGMOD-SIGART symposium on Principles of database systems*, p. 188, Seattle, WA, USA, June 1998.

[21] M. Sharma, A. Chaudhary, M. Mathuria, and S. Chaudhary, "A review study on the privacy preserving data mining techniques and approaches," *International journal of computer science and telecommunications*, vol. 4, no. 9, pp. 42–46, 2013.

[22] L. Sweeney, "Computational disclosure control - a primer on data privacy protection," Ph.d Thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, May 2001.

[23] P. F. Wu and Y. Q. Zhang, "K-anonymity: a model for protecting privacy," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 10, no. 5, pp. 557–570, 2002.

[24] P. F. Wu and Y. Q. Zhang, "Summary of database security," *Computer Engineering (in Chinese)*, vol. 32, no. 12, pp. 85–88, 2006.

[25] X. K. Xiao and Y. F. Tao, "Anatomy: Simple and effective privacy preservation," in *Proceedings of the 32nd International Conference on Very Large Data Bases (VLDB'06)*, pp. 139–150, Seoul, Korea, Sept. 2006.

[26] Y. Xin, Z. Q. Xie, and J. Yang, "The privacy preserving method for dynamic trajectory releasing based on adaptive clustering," *Information Sciences*, vol. 378, pp. 131–143, 2017.

[27] J. Xu, "Data anonymization based on the data availability(in chinese)," Master Thesis, Fudan University, Shanghai, China, June 2008.

[28] K. Yoshitaka and H. Kentaro, "Bottom-up cell suppression that preserves the missing-at-random condition," in *International Conference on Trust and Privacy in Digital Business*, pp. 65–78, Porto, Portugal, Sept. 2016.

[29] J. Yu, J. M. Han, and J. M. Chen, "Topdown-kaca: an efficient local-recoding algorithm for k-anonymity," in *The 2009 IEEE International Conference on Granular Computing (GrC'09)*, pp. 727–732, Nanchang, China, Aug. 2009.

[30] L. Yu, J. M. Han, Y. U. Juan, J. Jia, and H. B. Zhan, "A novel anonymization method based on anatomy and reconstruction in lbs privacy preservation," *Advances in Differential Equations*, vol. 19, no. 11, pp. 544–553, 2014.

Biography

Sang Ni is a master degree student in the school of computer science, Shanghai University. His research interests include cloud computing, big data analysis, computer and network security.

Mengbo Xie is a master degree student in the school of computer science, Shanghai University. His research interests include privacy protection, big data analysis, computer and network security.

Quan Qian is a Professor in Shanghai University, China. His main research interests concerns computer network and network security, especially in cloud computing, big data analysis and wide scale distributed network environments. He received his computer science Ph.D. degree from University of Science and Technology of China (USTC) in 2003 and conducted postdoc research in USTC from 2003 to 2005. After that, he joined Shanghai University and now he is the lab director of network and information security.

An Efficient Code Based Digital Signature Algorithm

Fang Ren[1], Dong Zheng[1], WeiJing Wang[1]
(Corresponding author: Fang Ren)
School of Telecommunication and Information Engineering, Xi'an University of Posts and Telecommunications[1]
Xi'an 710121, China
(Email: renfang_81@163.com)
(Received Aug. 4, 2016; revised and accepted Nov. 15, 2016 & Feb. 19, 2017)

Abstract

In the case of most current digital signature algorithm can be attacked by quantum algorithm, code based digital signature algorithm, which represents the Post-Quantum Cryptography, has become the hotspot of current research. CFS algorithms proposed in 2001 is one of the most important code based digital signature algorithm, but its signature efficiency is very low. In this paper, an improved CFS algorithm is proposed by means of code based hash function. The output of this hash function is a syndrome of a regular word whose weight is no more than error correcting capacity t of the code. By using this hash function instead of the random hash function, the decoding algorithm can avoid the time-consuming syndrome decoding attempts. The signing time of the improved algorithm reduces $t!$ times than the original. At the same time, the signature efficiency is no longer restricted to error correcting capacity of the code. Furthermore, the securities of these two algorithms both rely on the equivalent NP complete problems.

Keywords: Digital Signature; Hash Function; Quantum Attack; Syndrome

1 Introduction

Public-key cryptography has obtained a lot of valuable results since it was developed over 30 years ago. No matter in the field of individual privacy, commercial confidentiality, or even national security etc., it has played a key role. Under the threat of Quantum algorithm [13, 24], most of the widely using public-key algorithms based on number theoretic difficult problems nowadays are no longer secure. Currently, code based public key cryptography technique is regarded as a method which can resist Quantum attack [20, 23]. Because of this, it has become one of the mainstream of public key cryptography in future development.

McEliece proposed the first code-based public-key encryption algorithm based on the irreducible binary Goppa codes [14]. In this algorithm, the encryption process is equivalent to adding a random wrong vector to the plaintext; while the decryption process is corresponding to decoding. Another important algorithm named Niederreiter's algorithm [18] realizes encryption and decryption process through syndrome decoding. It has been proved that its security is equivalent to the McEliece algorithm: that means their security can be reduced to two NP-complete problems: the random binary codes decoding problem and the Goppa code distinguishing problem [3, 9]. Since code based cryptography technology was proposed, there are a number of research achievements over the past 30 years, including encryption, digital signature [6, 22], identification [5], hash function [1], stream ciphers [12] and so on, almost throughout all the fields of cryptography. And in the midst of digital signature field, Courtois-Finiasz-Sendrier (CFS) signature algorithm [6], which was proposed in 2001, has been viewed as a classic algorithm.

CFS algorithm, which is the first secure signature algorithm based on binary Goppa codes, is constructed on the basis of the Niederreiter encryption algorithm. Many comprehensive discussions about the security of the CFS have been made in the past decade [7, 11]. There are also a variety of improved algorithms such as mCFS [7], parallel CFS [10], etc. In addition, other special-purpose signature, such as ring signature [15, 25], blind signatures [19], etc. also can be constructed on the basis of CFS. Similar as CFS, the core of these algorithms is that the hash value of the message has to be transformed to a syndrome of Goppa code through preprocessing during signature. The signing process is decoding syndrome by using the secret decoding algorithm, which regards the codeword as signature value; to verify the validity of signature, the syndrome of the codeword is calculated and compared with the hash value of the message.

Although signature algorithms based on CFS could provide relatively high security, its shortcoming is still evident, namely, the efficiency of signature is rather low.

The original CFS algorithm, for example, in order to get a decodable syndrome, it has to execute $t!$ attempts averagely, where t is error correcting capacity of the Goppa codes. Apparently, if the parameter t increases, signatures times will grow exponentially rapidly, while low security defects can be brought by smaller t value. Meanwhile, this contradicts the basic aim of high error correcting capacity of error correcting codes, which, to some extent, hampers the application of CFS series algorithms.

In this paper, we mainly study the flaw mentioned above of CFS algorithms. Through analyzing algorithm implement details and identifying the main causes of the inefficiency of signature, signing process could be improved. We propose an efficient code based digital signature algorithm, whose signing time does not grow rapidly with the parameter t. Without reducing security, the efficiency of algorithms could be effectively improved. It is a type of more practical digital signature algorithm.

2 Preliminaries

2.1 Error Correcting Codes

Definition 1. *A (n,k) linear code C over a finite field F_q is a linear subspace with dimension k of the vector space F_q^n. The elements of F_q^n are called words, while the elements of C are called codewords. Number n is called the length of C and k is called the rank of it.*

Definition 2. *The matrix $\mathbf{G} \in F_q^{k\times n}$ is a generator matrix for the (n,k) linear code C over F_q, if the row of $\mathbf{G}$ span C over F_q.*

The generator matrix $\boldsymbol{G}$ for linear code C is not unique, but the different generator matrixes can mutually convert by elementary row transformation, namely, if $\boldsymbol{G}$ is a generator matrix for C, and $\boldsymbol{P}$ is an elementary matrix, $\boldsymbol{PG}$ is also a generator matrix for C.

Definition 3. *The parity check matrix $\mathbf{H} \in F_q^{(n-k)\times n}$ of (n,k) linear code C is defined by $\mathbf{H}\cdot x^T = 0, \forall x \in C$.*

The parity check matrix of the linear code is also not unique, which is similar to generator matrix. And different parity check matrix can also mutually convert by elementary row transformation. Vector c of length n is a codeword of C is equivalent to $\boldsymbol{H}c^T = 0$. For any word c, $\boldsymbol{H}c^T$ is called the syndrome of c.

Definition 4. *The Hamming distance $d(u,v)$ is defined as the number of different components of u and v, of which $u = (u_1, u_2, \ldots, u_n)$ and $v = (v_1, v_2, \ldots, v_n)$ are two codewords of the linear codes respectively, i.e. $d(u,v) = |\{i|u_i \neq v_i\}|$. Hamming weight $w(u)$ of codeword u is defined as the Hamming distance between u and all zero codeword, i.e.$w(u) = d(u,0)$, the minimum Hamming weight of non all zero codewords of code C is called the minimum distance of code C, generally sign as d_{min}.*

The error correcting capacity of the code is determined by the minimum distance, in general, error correcting capacity t of linear codes with the minimum distance d_{min} meets the condition $t \leq \lfloor \frac{d_{min}-1}{2} \rfloor$.

Goppa code is a kind of special linear code [2], whose parameters used in the McEliece encryption algorithm have the following form: $n = 2^m, k = n - mt$. The foundation of the efficient decoding is the specific structure of the generator polynomial of Goppa codes, which is also the basis of constructing code based cryptographic algorithms. That is by regarding the structure information of Goppa codes and corresponding decoding algorithm as secret trapdoor information or decryption private keys, a one-way trapdoor function can be used to construct public-key encryption algorithm and signature algorithm.

In this paper, linear codes and Goppa codes are over the binary field F_2.

2.2 Difficult Problems

Public-key cryptography is always founded upon some difficult problems, such as the security of RSA relies on the difficult problem of factoring big integer problem. The following is a summary of some difficult problems on which code based public key algorithms mainly rely. All of them have been proved that are NP-complete problems and can effectively resist known quantum attacks.

Problem 1. *Syndrome Decoding**(SD)** Problem,*

Input: *A finite field F_q, randomly select a matrix $\mathbf{H} \in F_q^{(n-k)\times n}$ and vector $s \in F_q^{n-k}$, integer $k > 0$.*

Output: *A word $x \in F_q^n$, its weight $w(x) \leq k$, and meets $\mathbf{H}x^T = s$.*

Problem 2. *Goppa Codes Distinguishing**(GD)** Problem,*

Input: *A finite field F_q, randomly select a matrix $\mathbf{H} \in F_q^{(n-k)\times n}$.*

Output: *Judge whether $\mathbf{H}$ is a (n,k) Goppa parity check matrix or a (n,k) random code parity check matrix?*

3 CFS Signature Algorithm

3.1 Principle and Realization

Digital signature is an important cryptographic techniques used to realize non-repudiation and authentication. There are generally three different ways to build code based digital signature algorithm: (1) Building an algorithm whose procedure is just the inverse process of the code based public-key encryption algorithm; (2) Using zero-knowledge identification algorithm together with the Fiat-Shamir paradigm to develop a signature algorithm; (3) Constructing a special subset of the syndrome space as the foundation of digital signature algorithm.

CFS signature algorithm belongs to the first category, which is a kind of signature algorithm based on classic

Niederreiter encryption algorithm. Such algorithms digital signature process can be concluded as follows:

- Calculate the hash value of the message m by using a public hash function;
- Regard the hash value as the cipher text and use the signers private key to decrypt it;
- Attach the proper forms of the decryption results behind a message m as a signature value.

For code based signature algorithm, however, it's pretty hard to accomplish the second step. The main reason is the output of cipher text by Niederreiter algorithm should be a syndrome with low weight error vectors. But the message m may not be transferred to a required syndrome, which is the cause of ineffectively decoding. Only the syndrome of the error vector whose weight does not exceed the decoding capacity t of the selected Goppa codes can be decoded successfully. Therefore, in effect, CFS algorithm is a probabilistic signature algorithm, which could not pause transforming the hash value of the message repeatedly until a valid syndrome has been found.

Basic CFS signature algorithm uses an increment counter to tag the number of decoding attempts. In order to avoid the security risks of this counter, Dallot developed a mCFS algorithm [7] which based on CFS signature algorithms but much secure. mCFS includes three phases: ***Gen_mCFS***, ***Sign_mCFS*** and ***Verify_mCFS***. Detailed description of the algorithm is shown in Algorithm 1.

Algorithm 1 mCFS Signature Algorithm

Gen_mCFS

Select a (n,k) Goppa code C randomly over F_2, of which the error correcting capacity is t, and the parity check matrix is $\boldsymbol{H}$, select a valid syndrome decoding algorithm γ;

Select a $(n-k)\times(n-k)$ invertible matrix $\boldsymbol{Q}$ over F_2, and a $n\times n$ permutation matrix $\boldsymbol{P}$ randomly;

Select a public secure hash function $h:\{0,1\}^* \rightarrow F_2^{n-k}$;

Define $< h, t, \boldsymbol{H}^{pub} = \boldsymbol{QHP} >$ as public parameters of the system, and $< \boldsymbol{Q}, \boldsymbol{H}, \boldsymbol{P}, \gamma >$ as the users private key.

Sign_mCFS$(msg, \boldsymbol{Q}, \boldsymbol{P}, \gamma)$

For the signer needs to sign a message msg, the signature process is as follows:

Calculate the hash value of the message msg, $s = h(msg)$;

Randomly select $i \in \{1, 2, \ldots, 2^{n-k}\}$, by using the secret decoding algorithm γ to try to decode $s_i = \boldsymbol{Q}^{-1}h(s||i)$, until i_0 has been found, which meets the existence of $\gamma(s_i)$;

If $v = \gamma(s_{i_0})$, the signature value is $(i_0||v\boldsymbol{P})$.

Verify_mCFS$(msg, i, u, \boldsymbol{H}^{pub})$

Set $< msg, i||u >$ as the message-signature pair of the receiver, the verify process is:

Calculate $a = h(h(msg)||i)$ as well as $b = \boldsymbol{H}^{pub}u^T$;

Signature is valid if and only if $a = b$.

3.2 Performance Analysis

Dallot et has conducted a rigorous formal proof of CFS and mCFS signature algorithm, that the security of the algorithm is reduced to the SD and GD problem under the Random Oracle model. Because of the high level of the security, most of the current code based signature scheme is designed on the basis of CFS.

Even though mCFS algorithm has very high level security, its realization efficiency, namely the speed of signature, is rather low, which is caused by too many syndrome decoding attempts. The analysis of mCFS signature algorithms success probability as below:

For pre-selected Goppa codes $(n = 2^m, k = n - mt)$, we assume that the number of decodable syndrome is N_d, the number of overall syndrome is N_t, obviously

$$N_t = 2^{n-k} = 2^{mt} = n^t \tag{1}$$

The weight of error vector which has decodable syndrome has to be less than the error correcting capacity t, hence

$$N_d = \sum_{i=0}^{t} \binom{n}{i} \approx \binom{n}{t} \approx \frac{n^t}{t!} \tag{2}$$

And the approximate success probability of mCFS signature algorithms is

$$P_s = \frac{N_d}{N_t} \approx \frac{\frac{n^t}{t!}}{n^t} = \frac{1}{t!} \tag{3}$$

That is to say, every $t!$ times attempts can only get one decodable syndrome. With t increasing, this number could grow relatively fast, such as set $t = 10$, a signature can be obtained after trying $10! = 3628800$ times averagely. In some earliest literatures [6] authors proposed $t = 9$. But under Bleichenbacher's attack [11], this parameter is no longer safe, the parameter $m = 15$, $t = 12$ or $m = 16, t = 10$ is recommended. In the long term, with new attack methods proposed, value of t will unavoidably growing larger and larger. In order to obtain a valid signature, the signing speed becomes lower and lower with numbers of syndrome decoding attempts growing exponentially, and at the same time, implement efficiency would become worse.

The main reason of the inefficiency of mCFS signature algorithms is generally the s_i calculated from hash value of the message is not a decodable syndrome of liner code C. In order to decode successfully, it has to find a decodable s_i through trying so many different s_i. A valid decoding and successful signature based on finding a proper s_i which is exactly within decoding capacity of C. In order to improve the efficiency of signature, the

original algorithm needs to be improved, so that the calculated s_i itself or at least in a great probability should be a decodable syndrome required.

4 Efficient Code Based Digital Signature Algorithm

In this section, we first construct a code based hash function and then on the basis of it, we improve the mCFS signature algorithm to obtain an efficient code based digital signature algorithm.

4.1 Code Based Hash Functions

The method for constructing a code based hash function is first proposed by Augot et, which is based on Merkle-Damgard design principle [16, 8], namely, a compression function f permits to loop calculate the given message several rounds for obtaining an iteration value as the hash value of the message. It can be proved that the security of the hash function constructed in accordance with this method has no less security than the compression function [1]. Bernstein and Meziani et improved the implement efficiency of the original method respectively [4, 17]. Such constructing methods can be concluded as:

Set compression function as f, and the input is s bits, the output is r bits ($r < s$). To derive hash value of the given message msg, it needs to do a number of loop iterations by using function f:

- The first round: Select the initial vector IV of length r; Select $s - r$ bits from a given message msg, sign as m_0, and concatenate it with the IV as the initial input vector of f with length s, then get r bits initial output;
- Starting from the second round, feed r bits pre-round output back to the input, similar as the first round, select $s - r$ bits from the message msg as m_i in order, concatenate r with m_i as the input vector of f. Calculate the new r bits output.
- Loop this process until the message is taken out. During the final round, if the remaining bits of the message msg are insufficient to $s - r$ bits, randomly select some bits to meet the requirement. The final output of the function f is the hash value of the message msg.

Figure 1 shows this iterative process. In the construction of the hash function mentioned above, the compression function f is the most important part, and even the security of hash function also depends on the security of f. A kind of constructing method of compression function f based on coding difficult problem is presented below.

First select a (n, k) Goppa codes, where $n = 2^m, k = n - mt$, and select a positive integer $w|n$. It is clear that $w = 2^{m'}, m' < m$. Set $l = n/w = 2^{m-m'}$.

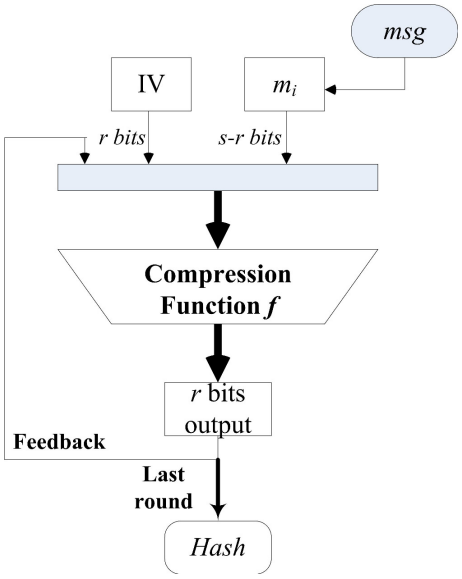


Figure 1: The diagram of hash iterations

For any word c of length n, it can be divided into w blocks of equal length, each block contains l bits. If a word c of weight w within each block $((i-1)l, il]$ happens to have only one 1, c is called as *regular word*.

Set $\boldsymbol{H}$ as the parity check matrix of Goppa codes, which is a $(n - k) \times n$ matrix. Divide $\boldsymbol{H}$ into w submatrix $\boldsymbol{H}_i, i = 1, 2, \ldots, w$ in accordance with the following method

$$\boldsymbol{H} = (\boldsymbol{H}_1, \boldsymbol{H}_2, \ldots, \boldsymbol{H}_w) \tag{4}$$

of which$\boldsymbol{H}_i = (h_{(i-1)l+1}, h_{(i-1)l+2}, \ldots, h_{il})$, and$h_j$ is the jth column of the matrix $\boldsymbol{H}$.

Next we define compression function $f : F_2^s \to F_2^r$, where $s = wlog_2 l$, and $r = n - k = mt$ is the number of matrix $\boldsymbol{H}$'s rows.

For any $x \in F_2^s$, x is divided into the w blocks of equal length in accordance with the same way, that is $x = (x_1, x_2, \ldots, x_w)$, and $x_i \in F_2^{log_2 l}$. Convert x_i into numbers between 0 to $l - 1$. Select the $(x_i + 1)$th column of the matrix $\boldsymbol{H}_i$, that is $h_{(i-1)l+x_i+1}$. Calculate $z = \sum_{i=1}^{w} h_{(i-1)l+x_i+1}$, then the output of the compression function is $f(x) = z$.

Theorem 1. *The output of the compression function f above is equivalent to calculating a syndrome of a regular word of length n and weight w, that is, for any $x \in F_2^s$, a regular word c could be found which meets $\mathbf{H}c^T = f(x)$.*

Proof. First of all, according to the definitions above,

$$f(x) = \sum_{i=1}^{w} h_{(i-1)l+x_i+1}. \tag{5}$$

Define a word $c = (c_1, c_2, \ldots, c_n)$ of length n is as follows:$c_j = 1 \Leftrightarrow \exists x_i, (i-1)l + x_i + 1 = j$. That means existing a x_i, after converting it to a decimal number, the

selected column number is corresponding to the location label j of c_j. Due to calculating a syndrome of a word is equivalent to adding matrix $\boldsymbol{H}$'s columns which are corresponding to non-zero bits of the word, by definition, $f(x)$ is exactly the syndrome of word c, namely $\boldsymbol{H}c^T = f(x)$.

According to the definition of c, c have and only have one 1 within each block $((i-1)l, il], i = 1, 2, \ldots, w$. So that c is a regular word of weight w. □

Based on the compression function above, we define a code based hash function $h_c : \{0,1\}^* \to F_2^r$ as below:

For a given message msg, choose (n, k) Goppa codes and get a compression function f in accordance with the definition above. Through using Augot loop iteration method several times to compress message msg by f, we can obtain a bit string of length r as hash values $h_c(msg)$. The function h_c can apply on arbitrary length message msg, and the output is a bit string of length $r = n - k$.

Theorem 2. *As the above definition, the output of code based hash function h_c is a syndrome of a regular word of length n and weight w.*

Proof. According to the loop iteration constructing methods of the hash function, the output hash value of the final round is also the output of the function f. According to Theorem 1, for any message msg, $h_c(msg)$ is a syndrome of a regular word of length n and weight w. □

We analyze the security of hash functions having this structure: obviously the one-way character of hash function h_c relies on a special SD problem:

Input: A $(n-k) \times n$ matrix $\boldsymbol{H}$ over finite field F_2, vector $s \in F_2^{n-k}$, integer $k > 0$;

Output: A regular word $x \in F_2^n$, its weight $w(x) \leq k$, and satisfies the condition $\boldsymbol{H}x^T = s$.

Augot called this problem as Regular Syndrome Decoding (**RSD**) Problem. It can be proved that this is a NP complete problem [1].

4.2 An Efficient Digital Code Based Signature Algorithm

The mCFS algorithm,which is proposed by Dallot et, can improve the original CFS signature algorithm [7] with stronger security. In this section, by applying the code based hash function h_c given in 4.1, we aim to improve the implement efficiency of mCFS to obtain an efficient signature algorithm mCFS$_c$. This signature algorithm can greatly improve the signature efficiency of mCFS without any decrease of security. mCFS$_c$ also includes three phases: ***Gen_mCFS***$_c$, ***Sign_mCFS***$_c$ and ***Verify_mCFS***$_c$. Algorithm 2 gives the details of the algorithm.

The correctness of verify process in Algorithm 2 can be proved as below: If $< msg, R'||u >$ is a legitimate pair of message-signature through signature process above, it can get the equation as follows:

$$\begin{aligned} b &= \boldsymbol{H}^{pub}u^T = \boldsymbol{HP}(v\boldsymbol{P})^T = \boldsymbol{HPP}^T v^T \\ &= \boldsymbol{H}v^T = s = h_c(h_c(msg)||R') = a \end{aligned}$$

Algorithm 2 mCFS$_c$ Signature Algorithm

1: ***Gen_mCFS***$_c$
2: Selects a (n, k) Goppa codes C randomly over F_2, with the correcting error capacity t and the parity check matrix $\boldsymbol{H}$, a valid decoding syndrome algorithm γ;
3: Randomly select a $n \times n$ permutation matrix $\boldsymbol{P}$ over F_2;
4: Choose a positive integer $w \leq t$ and $w|n$, and construct code based hash function $h_c : \{0,1\}^* \to F_2^{n-k}$;
5: Define $< h_c, t, \boldsymbol{H}^{pub} = \boldsymbol{HP} >$ as system public parameters, and $< \boldsymbol{H}, \boldsymbol{P}, \gamma >$ as the users private key.
6: ***Sign_mCFS***$_c(msg, \boldsymbol{P}, \gamma)$
7: Set the message of the signer is msg, and the signature process is:
8: Choose a one-time random number $R \in \{1, 2, \ldots, 2^{n-k}\}$, and calculate $s = h_c(h_c(msg)||R)$;
9: Set $v = \gamma(s)$, so that the signature value is $(R||v\boldsymbol{P})$.
10: ***Verify_mCFS***$_c(msg, R', u, \boldsymbol{H}^{pub})$
11: Set the received message signature pair is $< msg, R'||u >$, the verify process is:
12: Calculate $a = h_c(h_c(msg)||R')$ and $b = \boldsymbol{H}^{pub}u^T$;
13: Signature is valid if and only if $a = b$.

5 Performance Analyses of Algorithms

This section focuses on the security analysis and efficiency analysis of the code based signature algorithm mCFS$_c$ mentioned in Section 4.2 and comparing it with other existing code based signature algorithms.

Between the three kinds of construction method of building code based digital signature algorithm, the second one, based on zero-knowledge identification algorithm and the Fiat-Shamir paradigm, always have very long signature length [20], roughly 120 Kbits. The third method, constructing a special subset of the syndrome space as the foundation of digital signature algorithm, have been proved only be used as one-time signature [20]. So, the first method, represented by mCFS, is the mainstream of code based signature and we only compare our algorithm with the mCFS algorithm.

5.1 Security Analysis

First of all, we analyze the security. Compared with the mCFS signature algorithm, the primary difference is replacing the random hash function h with the code based hush function h_c. The point is the essence of this change is that it substitutes random hash function for a trapdoor hash function, and the trapdoor information is decoding

Table 1: The security comparison of two algorithms

Signature algorithm	Dependent problems	Hardness of problems
mCFS	SD, GD	NP complete
mCFS_c	RSD, GD	NP complete

Table 2: The efficiency comparison of two algorithms

Signature algorithm	Hash times	Decoding times	Hash times($t=9$)	Decoding times($t=9$)
mCFS	$t!+1$	$t!$	362881	362880
mCFS_c	2	1	2	1

Table 3: The signature time consumption (in seconds) of two algorithms

(m,t)	(15,7)	(15,8)	(15,9)	(16,7)	(16,8)	(16,9)
mCFS	189.58	2570.48	35562.24	442.51	7862.41	57697.92
mCFS_c	0.052	0.073	0.109	0.096	0.203	0.327

algorithm γ of selected Goppa codes. For this type of hash functions, anyone who knows the trapdoor information can effectively calculate the inverse of the hash value, or else, any useful values cannot be provided without the trapdoor information.

In mCFS_c, decoding algorithm γ is the signer's private key which couldn't be obtained but the signer. The security of this hash function can be guaranteed so long as the absolute confidentiality of private key. So the security of mCFS_c is equivalent to mCFS. Hence, this change does not result in any reduction of security. Table 1 shows the security comparison of these two algorithms.

5.2 Efficiency Analysis

According to Algorithm 2, during the process of signing message msg, it has to perform twice hash computation and once syndrome decoding algorithm. According to the Theorem 2, the output of hash functions h_c is a syndrome of a regular word of weight w which does not exceed decoding capacity t of the selected Goppa codes. Therefore anyone who has the secret syndrome decoding algorithm γ can always effectively obtain one regular word of length n and weight w. Compared with mCFS algorithm average $t!$ times attempts to get a decodable syndrome, the biggest advantage of mCFSc is greatly improving signature speed by relieving plenty of decoding attempts. In the long term, this algorithm provides a fundamental method to liberate algorithm from the restriction of code parameter t, so that we can obtain high security by choosing very large t without any reduction of signature speed. Table 2 shows the efficiency comparison of these two algorithms.

In Table 2, parameter t takes the classical value 9. In order to obtain higher security, this value should be increased, and $t = 10$ or $t = 12$ is recommended [11]. It is easy to see with t increases, the consumption of mCFS will increase rapidly, while the consumption of our algorithms mCFS_c remains very low. In order to resist the new attacks in the future, the value of t will unavoidably growing larger and larger, and the implement efficiency of mCFS will become worse and worse, while mCFS_c always has good performance.

5.3 Experimental Results

In this section, we give some experimental results to reveal the efficiency difference between mCFS and mCFS_c. Because of the similarity of ***Gen*** and ***Verify*** phases of these two algorithms, we only count the time consumption of ***Sign*** phase, the most time-consuming phase in Algorithm 1 and Algorithm 2.

The software we used is Magma V2.12, running on 64 bit Windows7 operating system, and the hardware parameters are: Intel Core i7-4710, 2.50GHz, 4GB RAM. The decoding algorithm for Goppa codes is the *Patterson* algorithm [21].

We first selected six different Goppa codes with different parameters m and t. For each code we selected 20 text files with size of 10kB and counted the average time consumption of ***Sign*** phase in these two different algorithms. The experimental results are show in Table 3.

6 Conclusions

As the most important code based digital signature algorithm, the security and implement efficiency of CFS has been extensively studied since it was first proposed. However, with parameter increasing very quickly, its still hard to fundamentally solve the sharp reduction of the signing speed. The further application of the algorithm is therefore seriously limited.

This article proposed and analyzed an improved code based signature algorithm $mCFS_c$ by introducing code based hash function into mCFS algorithm. $mCFS_c$ algorithm can be expected to avoid repeated decoding syndrome attempts to find a decodable syndrome, which increases the signature speed. In addition, compared with mCFS, the signature time can be greatly reduced without any reduction of error correcting capacity t. Meanwhile, the new method has the same security as mCFS algorithm. Therefore it is a more practical code based signature algorithm.

Acknowledgments

This paper was supported by the National Nature Science Foundation of China (Program No. 61272037, 61472472, 41504115); Natural Science Basic Research Plan in Shaanxi Province of China (Program No. 2015JQ6262, 2016JM6033) and Scientific Research Program Funded by Shaanxi Provincial Education Department (Program No.15JK1669, 15JK1661). The authors gratefully acknowledge the anonymous reviewers for their valuable comments.

References

[1] D. Augot, M. Finiasz, and N. Sendrier., "A family of fast syndrome based cryptographic hash functions," in *Progress in Cryptology (Crypto'05)*, pp. 64–83, 2005.

[2] E. Berlekamp, "Goppa codes," *IEEE Transactions on Information Theory*, vol. 19, no. 5, pp. 590–592, 1973.

[3] E. R. Berlekamp, R. J. McEliece, and H. C. A. Van Tilborg, "On the inherent intractability of certain coding problems," *IEEE Transactions on Information Theory*, vol. 24, no. 3, pp. 384–386, 1978.

[4] D. J. Bernstein, T. Lange, and et al. C. Peters, "Really fast syndrome-based hashing," in *Progress in Cryptology (AFRICACRYPT'11)*, pp. 134–152, 2011.

[5] P. L. Cayrel and P. Véron, "Improved code-based identification scheme," *Computer Science*, arXiv:1001.3017, 2010. (`https://arxiv.org/abs/1001.3017`)

[6] N. T. Courtois, M. Finiasz, and N. Sendrier, "How to achieve a mceliece-based digital signature scheme," in *Advances in Cryptology (ASIACRYPT'01)*, pp. 157–174, 2001.

[7] L. Dallot, "Towards a concrete security proof of courtois, finiasz and sendrier signature scheme," in *Research in Cryptology*, pp. 65–77, Berlin Heidelberg: Springer, 2008.

[8] I. B. Damgard, "A design principle for hash functions," in *Advances in Cryptology (CRYPTO'89)*, pp. 416–427, Springer New York, 1990.

[9] D. Engelbert, R. Overbeck, and A. Schmidt, "A summary of mceliece-type cryptosystems and their security," *Journal of Mathematical Cryptology*, vol. 1, no. 2, pp. 1–51, 2007.

[10] M. Finiasz, "Parallel-cfs," in *Selected areas in cryptography*, pp. 159–170, 2011.

[11] M. Finiasz and N. Sendrier, "Security bounds for the design of code-based cryptosystems," in *Advances in Cryptology (ASIACRYPT'09)*, pp. 88–105, Berlin Heidelberg: Springer, 2009.

[12] P. Gaborit, C. Lauradoux, and N. Sendrier, "Synd: a fast code-based stream cipher with a security reduction," in *IEEE International Symposium on Information Theory*, pp. 186–190, Nice, France: IEEE Information Theory Society, 2007.

[13] L. K. Grover, "A fast quantum mechanical algorithm for database search," *Annual Acm Symposium on Theory of Computing*, pp. 212–219, 1996.

[14] R. J. McEliece, "A public-key cryptosystem based on algebraic coding theory," *DSN progress report*, vol. 42, no. 44, pp. 114–116, 1978.

[15] C. A. Melchor, P. Cayrel, and et al. P. Gaborit, "A new efficient threshold ring signature scheme based on coding theory," *IEEE Transactions on Information Theory*, vol. 57, no. 7, pp. 4833–4842, 2011.

[16] R. C. Merkle, "One way hash functions and des," in *Advances in Cryptology (CRYPTO'89)*, pp. 428–446, Springer New York, 1990.

[17] M.Meziani, O.Dagdelen, and et al. P. L.Cayrel, "S-fsb: An improved variant of the fsb hash family," *International Journal of Advanced Science and Technology*, vol. 35, pp. 73–82, 2011.

[18] H. Niederreiter, "Knapsack-type cryptosystems and algebraic coding theory," *Problems of Control and Information Theory*, vol. 15, no. 2, pp. 159–166, 1986.

[19] R. Overbeck, "A step towards qc blind signatures," *Iacr Cryptology Eprint Archive*, 2009.

[20] R. Overbeck and N. Sendrier, "Code-based cryptography," in *Post-quantum cryptography*, pp. 95–145, 2009.

[21] N. Patterson, "The algebraic decoding of goppa codes," *IEEE Transactions on Information Theory*, vol. 21, no. 2, pp. 203–207, 1975.

[22] M. K. Preetha, V. Sachin, and R. C. Pandu, "On provably secure code-based signature and signcryption scheme," *Iacr Cryptology Eprint Archive*, 2012.

[23] F. Ren, D. Zheng, and J. Fan, "Survey of digital signature technology based on error correcting codes(in chinese)," *Chinese Journal of Network and Information Security*, vol. 2, no. 11, pp. 1–10, 2016.

[24] P. W. Shor, "Algorithms for quantum computation: discrete logarithms and factoring," *Symposium on Foundations of Computer Science*, pp. 124–134, 1994.

[25] D. Zheng, X. Li, and K. Chen, "Code-based ring signature scheme," *International Journal of Network Security*, vol. 5, no. 2, pp. 154–157, 2007.

Biography

Fang Ren received his PhD degree in cryptography from Xidian University in 2012. Now he is an associate professor of Xi'an University of Posts and Telecommunications. His research interests include information security and code based cryptography.

Dong Zheng received his PhD degree from Xidian University in 1999. Now he is a professor of National Engineering Laboratory for Wireless Security,Xi'an University of Posts and Telecommunications. His research interests include cloud security, code based systems and other new cryptographic technology.

Weijing Wang received her Master degree in information security from Xi'an University of Posts and Telecommunications in 2017. Her research interests include information security and biometrics protection.

Reviewers (Volume 19, 2017)

Vijaya Lakshmi A
R Ch A Naidu
Abdeldime Mohamed Salih Abdelgader
Slim Abdelhedi
Mohd Faizal Abdollah
Raihana Syahirah Abdullah
Ashwini B Abhale
Syed Hasan Adil
Muhammad Najmi Ahmad Zabidi
Asimi Ahmed
Abdul-Gabbar Tarish Al-Tamimi
Shahid Alam
Sara Ali
Ashraf Hamdan Aljammal
Ali Mohamed Allam
Rengarajan Amirtharajan
Dawar Asfandyar
Saeed Bahmanabadi
Pijush Barthakur
Eihab B Bashier
Sunny Behal
Krishna Bhowal
Monowar H. Bhuyan
Tianjie Cao
Zhengjun Cao
Zhenfu Cao
Chi-Shiang Chan
Ayantika Chatterjee
Jan Min Chen
Liang Chen
Zhen Hua Chen
Chin-Ling Chen
Hu Chengyu
Abbas Cheraghi
Rachid Cherkaoui
Kaouthar Chetioui
Shu-Fen Chiou
Mohamedsinnaiya Chithikraja
Hassan Chizari
Tae-Young Choe
Kim-Kwang Raymond Choo
Yung-Chen Chou
Christopher P. Collins
Dipen Contractor
Ranjan Kumar Dash
Subhasish Dhal
Ahmed Drissi
Jiao Du
Qi Duan
Azz El Arab El Hossaini
Ahmed A. Elngar
Saad En Niari
Yousef Farhaoui
Xingbing Fu
Tiegang Gao
Amzari Ghazali
Krishan Kumar Goyal
Ke Gu
Rui Guo
C P Gupta
Sandeep Raj Gurung
Arash Habibi Lashkari
Tanmoy Halder
Yasir Hamid
Mody Ali Hamza
Disha Handa
Charifa Hanin
Seyed Hashemi
Abubaker Wahaballa Hassan
Ali Hassan
Amir Hassani Karbasi
Thaier Hayajneh
Pipat Hiranvanichakorn
Gwoboa Horng
Osama Hosam Eldeen
Peng Hu
Xiong Hu
Yu-Chen Hu
Xuexian Hu
Sk Hafizul Islam
Remani Naga Venkata Jagan Mohan
Amit Jain
N Jeyanthi
Shaoquan Jiang
Lin Zhi Jiang
Zeng Jianping
Wang Jie
Zhengping Jin
Adri Jovin
Omprakash Kaiwartya
Kannan Karthik
Asia Samreen Khan
Md. Al-Amin Khandaker
Basappa Bharamappa Kodada
Pramote Kuacharoen
Sajja Ratan Kumar
K S Anil Kumar
Sunil Kumar
Beesetti Kiran Kumar
Saru Kumari
Cheng-Chi Lee
Jung-San Lee
Jiping Li
Chun-Ta Li
Yang-Bin Lin
Liang Liu
Guanfeng Liu
Guangjun Liu
Rongxing Lu
Zhe-Ming Lu
K. Shantha Kumari Luke Jayakumar
Ming Luo
Lintao Lv
Sagar Bhaskar Mahajan
Tanmoy Maitra
Yassine Maleh
Arun Malik
Bo Meng
Suhail Qadir Mir
Anuranjan Misra
Hamdy M. Mousa
Arif - Muntasa
Zulkiflee Muslim
Amitava Nag
Loris Nanni
Syed Naqvi
Amin Nezarat
Siaw-Lynn Ng
Nasrollah Pakniat
Leon Pan
Mrutyunjaya Panda
Mohamed M.Y. Parvees
Kanubhai K Patel

Kailas Ravsaheb Patil
Basavaraj Patil
Gerardo Pelosi
Jiaohua Qin
Chuan Qin
Kashif Naseer Qureshi
Balakrishnan R R
Hashum Mohamed Rafiq
Benjamin W. Ramsey
Anurag Rana
Rama Chandra Rao
Vimalathithan
Rathinasabapathy
Dhivya Ravi
Jianguo Ren
Ali Safa Sadiq
Magdy M. Saeb
Arun Kumar Sangaiah
Arindam Sarkar
Solomon Sarpong
Sathish Ku Sathish Ku
Neetesh Saxena
Michael Scott
Resmi Sekhar
Thamizh D Selvam
Irwan Sembiring
Vrushank Shah
Kareemulla Shaik
Tarun Narayan Shankar
Udhayakumar Shanmugam
Sandeep Singh
Rajeev Sobti
Vuda Sreenivasarao
Deris Stiawan
Siva Shankar Subramanian
Manesh T
Nedal Mohammad Tahat
Maryam Tanha
Abebe Tesfahun
Miaomiao Tian
Geetam Singh Tomar
Zouheir Mustapha Trabelsi
S.C. Tsaur
Venkanna U
Karthikeyan Udaichi
Jyothsna V
Janani V S
Raghav V. Sampangi
Vandani Verma
Osman Wahballa
Ding Wang
Feng Wang
Yiming Wang
Hongbin Wang
Jianghong Wei
Fushan Wei
Chenhuang Wu
Keke Wu
Lei Xu
Zhao Xu
Chunxiang Xu
Prasant Singh Yadav
Xu Yan
Zheng Yang
Jun Ye
Venkatramana Reddy Yeddula
Milad Yousefi
Huifang Yu
Jianping Zeng
Wen Zhang
Xiaojun Zhang
Fangguo Zhang
Jie Xiu Zhang
Mingwu Zhang
Yinghui Zhang
Xingwen Zhao
Ming Zhao
Luo Zhiyong
Zhiping Zhou
Frank Zhu

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.