IJNS
International Journal of Network Security

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

Vol. 21, No. 4 (July 2019)

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. 21, No. 4 (July 1, 2019)

AAVSS: Auxiliary Authorization Mechanism Based on Arbitrary-Angle VSS

Ying Chin Chen[1], Kuo Jui Wei[1], Jung San Lee[1], Ngoc Tu Huynh[2] and Jyun Hong Lin[1]
(Corresponding author: Jung San Lee)
Department of Information Engineering and Computer Science, Feng Chia University[1]
Taichung, 40724, Taiwan
Faculty of Information Technology, Ton Duc Thang University[2]
Ho Chi Minh City, Vietnam
(Email: leejs@fcu.edu.tw)
(Received Nov. 28, 2017; Revised and Accepted May 5, 2018; First Online Mar. 2, 2019)

Abstract

To provide seamless Internet services, most public buildings, including coffee shops, airports, and libraries, temporary personal computers are offered to users for network access. This has led to a potential risk that the secret information of the user may be leaked out once these temporary computers have been affected by Trojans. Generally, the service provider checks the authority of a user according to a series of authentication procedure. While a user enters the verification token into the public computers, an attacker may apply a key-logger to steal the password or personal information. In this article, we have first introduced the visual secret sharing technique with arbitrary-angle stacking to design an auxiliary authorization protocol. According to the stacked one-time password, users can have the access to network services without keying any secret information into the public computers. Moreover, the efficiency of AAVSS is favorable to resource-constrained mobile device.

Keywords: Auxiliary Authorization; Keylogger; Nearest-Neighbor Interpolation; VSS

1 Introduction

As the Internet brings convenience for the whole world, human beings now can purchase clothes online, chat online, organize conference online, watch the pay-TV, and enjoy other online products expediently. Corresponding to this trend and due to the heavy work burden on people in the modern society, it is in great need that people can access to various services or products just through the Internet without having to go out personally.

In recent decades, various kinds of network services are offered via the Internet. People expect that they are able to obtain the services everywhere and anytime. Thus, the verification mechanisms are necessary for confirming legibility of both service provider and user. Among those mechanisms, the password authentication [2–6, 9, 11–13] is an easy way to achieve this purpose. In such environment, a new user has to provide a pair of identity and password to the service provider in the registration. The server then keeps the secret information in the database once it has accepted the joining request. After that, the server maintains the service access according to the comparison between the received authentication token and the recorded one in the database.

Nevertheless, researchers have pointed out that this simple mechanism might suffer from the stolen verifier attack. Besides, it requires a large amount of memory to maintain the password table and corresponding information. Hereafter, the smart card has been introduced in the design of authentication mechanisms to solve this security problem and mitigate the storage consumption. Personal information and authentication token are kept in the card instead of the server database for validity proof. Hence, the risk of stolen verifier attack could be eliminated.

Subsequently, there have many attacks based on the information retrieved from the smart card. According to the extracted token, intruders can further mount malicious attacks, such as forgery attack and impersonation attack. In addition, two-factor authentication schemes have been designed for enhancing the entropy of verification token, in which the difficulty in compromising the system could be highly reinforced. Aside from a smart card, people adopt the fingerprint as the other factor for personal information protection. The sampling process of fingerprint, however, is a tough problem in the implementation. It is due to the high sensitivity of cryptographic function.

Actually, the validity is not the only essential requirement that these authentication mechanisms have to achieve. Many new challenges including anonymity, untraceability, efficiency, and resistance to recent attacks, have come out in designing a novel authentication mech-

Table 1: VSS stacking

s_1	s_2	$s_1 \vee s_2$
■	□	■
■	■	■
□	■	■
□	□	□

anism. Among them, a common situation is seldom considered in the field of authentication mechanism. That is, people often need to login a network system via a public computer or some other persons laptops. In this case, it is hard to guarantee that the login information could be well-protected. An attacker can install a key-logger program into a public computer; Thus, recording the identity and password of a login user [7, 11]. Once people cannot get rid of this temporary switch situation, how to prevent a personal secret from being intercepted or recorded must be firmly concerned in developing an authentication system.

In this article, we aim to propose an auxiliary authority mechanism based on the visual secret sharing (VSS) technique [8, 10], in which a user can use the mobile phone to switch a service to a temporary computer instead of entering any personal secret. For simplicity, the abbreviation of the new mechanism is defined as AAVSS. More precisely, a user can login the network systems with personal computer by a pre-defined authentication scheme. Whenever the personal computer is inaccessible, the user can employ the smartphone to obtain the one-time token (OTP) to have the access [3]. In AAVSS, the nearest-neighbor interpolation (NNI) is integrated into the VSS to achieve the arbitrary-angle stacking, while a smartphone is applied to record the base of VSS [1, 15]. According to International Telecommunication Union (ITU) statistics in 2015, there are more than 7.085 billion smartphone subscribers around the world [14]. The adoption of smart phone does make sense while being integrated into an authentication mechanism. Moreover, the computing ability of a smart phone is much higher than that of a smart card. This device can share parts of computation for verification. In particular, the property of arbitrary-angle stacking of VSS is difficult to complete; thus, it is the main challenge in developing AAVSS.

The rest of this article is organized as follows. Related works are introduced in Section 2, followed by the details of AAVSS in Section 3. The performance analysis is explained in Section 4. We make conclusions in Section 5.

2 Related Works

In the following, we introduce the concept of VSS and NNI. The VSS is used to embed OTP content into shares. Hereafter, a legal user can figure out OTP by stacking shares on the base kept in the smartphone. As to the NNI, it is adopted to achieve the arbitrary-angle stacking.

Table 2: The stacking principle of Lin *et al.*'s scheme

GS_1	GS_2^0	GS_2^{120}	GS_2^{240}	SP_1	SP_2	SP_3
□	□	□	□ ■	□	□	□ ■
		■	□ ■		■	□ ■
	■	□	□ ■	■	□	□ ■
		■	□ ■		■	□ ■
■	□	□	□ ■	■	■	■ □
		■	□ ■		□	■ □
	■	□	□ ■	□	■	■ □
		■	□ ■		□	■ □

2.1 Visual Secret Sharing

The VSS technique is first proposed by Naor and Shamir [10]. It distributes the pixel of a secret image SP to two transparent shares TS_1 and TS_2. S_1 and S_2 denote the corresponding position pixels of TS_1 and TS_2. The stacking operation is the OR boolean operation ($\vee$), as shown in Table 1. People cannot learn anything useful information of SP from only one share. By stacking those two transparencies, the content of SP can be revealed.

In 2014, Lin *et al.* have extended this idea to develop a multi-secret VSS mechanism based on random grid [8], in which the secrets can be shown by specific angle stacking. Given three secret images SP_1, SP_2, and SP_3, they randomly generate a grid base GS_1 and three temporary shares GS_2^0, GS_2^{120}, GS_2^{240} according to the principle in Table 2. Note that Lin *et al.* apply the exclusive-or operation to stack shares instead of OR function. How to construct those shares depends on two rules:

1) If the pixel of secret image is white, the pixel at the corresponding position of share must be the same as that of GS_1.

2) If the pixel of secret image is black, the pixel at the corresponding position of share must be the opposite one of GS_1.

Hereafter, a user who collects base GS_1 and share GS_2 is able to reveal SP1. Rotating GS2 by 120°, the user further extracts SP_2. Finally, the user can figure out SP_3 by rotating GS_2 by 240°.

2.2 Nearest-Neighbor Interpolation

The NNI is a technique used to adjust the scale of image or to rotate an image [1, 15]. In AAVSS, we apply it to re-define the pixel coordinates of share after rotating. Suppose that the new pixel coordinate is (x', y'), as illustrated in Figure 1. Due to the rotation operation, the

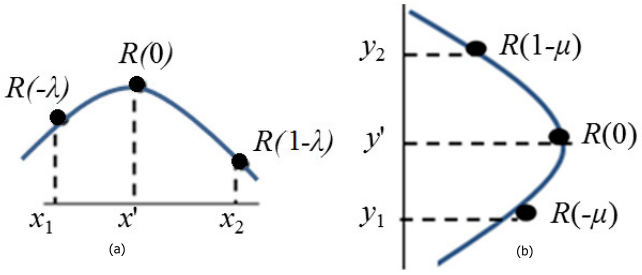


Figure 1: The new coordinate (x', y') of pixel after rotating

values of (x', y') might not be integers. Note that (x_1, y_1) and (x_2, y_2) are the neighbor coordinates with integer appearance. Thus, we can use NNI to transfer these values into integers for fulfilling the image format. We first calculate the difference between (x', y') and neighbors based on Equation (1). Subsequently, the new coordinate is modified as (x", y") according to Equation (2) and Equation (3). The function R(u) is used to check the nearest coordinate of (x', y'), and u is the parameter for distance judgment.

$$\left\{ \begin{array}{l} \lambda = x' - x_1, x_1 \le x' \le x_2 \\ \mu = y' - y_1, y_1 \le y' \le y_2 \end{array} \right\} \quad (1)$$

$$\left\{ \begin{array}{l} x'' = R(-\lambda)x_1 + R(1-\lambda)x_2 \\ y'' = R(-\mu)y_1 + R(1-\mu)y_2 \end{array} \right\} \quad (2)$$

$$R(u) \quad = \quad \left\{ \begin{array}{l} 1 \text{ if } -0.5 \le u \le 0.5 \\ 0 \text{ otherwise} \end{array} \right\} \quad (3)$$

Figure 2 is an example illustrating how to decide the new coordinate. Black dots of Figure 2(a) are the original coordinates, while white ones are the rotated positions. Suppose $(x', y') = (2.7, 1.4)$ in Figure 2(b). (x_1, y_1), (x_1, y_2), (x_2, y_1) and (x_2, y_2) are (2, 1), (2, 2), (3, 1), and (3, 2), respectively. Thus, we can have $(x'', y'') = (3, 1)$ according to the above equations.

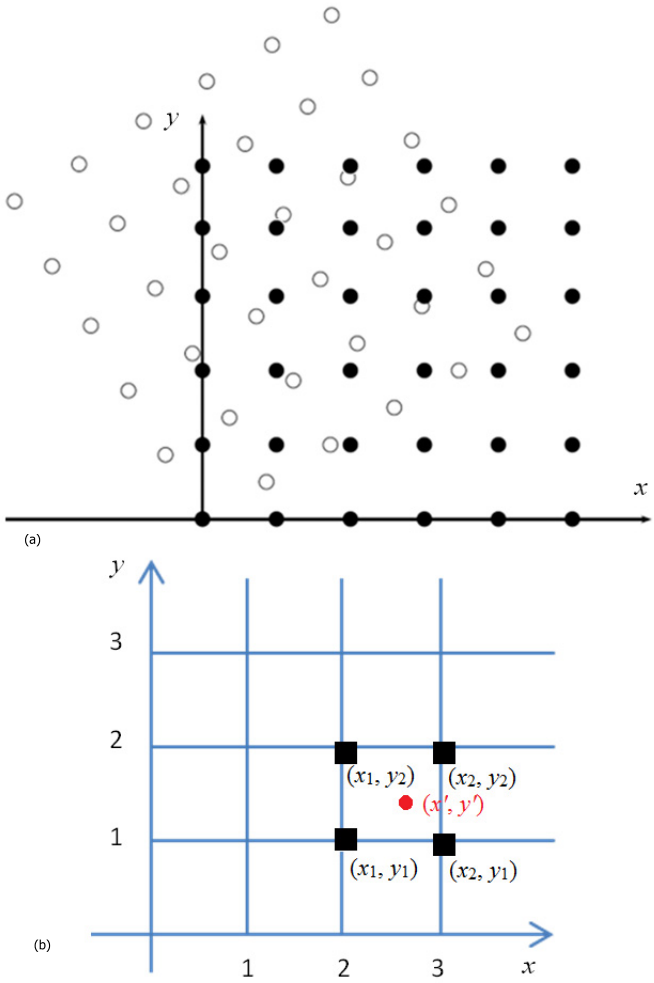


Figure 2: Example of coordinate re-definition

3 The Details of AAVSS

In this section, we describe the details of the auxiliary authority mechanism, including the setup phase, transference phase, and VSS share generation phase. In the first phase, we explain the environment setup and essential assumptions. The authority transference protocol is introduced in the second phase. How to achieve the arbitrary-angle stacking is presented in the final phase.

3.1 Setup Phase

There are four roles in AAVSS, MU (Mobile User), AS (App Server), PC (Public personal Computer), and S (Service provider). MU is a mobile subscriber with an intelligent cellular phone containing a secret base S_1. S_1 is shared between AS and MU. AS is an App server which is responsible for generating VSS base and share. MU has to register at AS and installs an App connecting to AS. S is a server offering a specific service over the Internet. AS and S have shared a secure channel, while MU is a registered user of S and possesses a pair of identity and password. PC is the so called non-personal computer. Notations used in the article are defined in Table 3.

3.2 Transference Phase

Once MU switches to a PC, MU can obtain an OTP token to access network service instead of entering a real password according to the following procedure. The flowchart of the transference phase is illustrated in Figure 3.

Step 1. MU starts the App and sends a request to AS, including ID_{MU}, IDs, and RS.

Step 2. AS generates a four-digit secret number FS and sends it to S along with ID_{MU} via a pre-shared secure channel. Note that FS is an OTP for this pro-

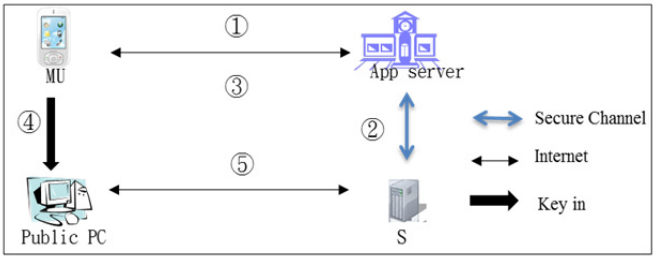


Figure 3: Transference flowchart

Table 3: Notations used in AAVSS

IDs	The identity of service provider S
ID_MU/PW_MU	The identity/password of MU for S
RS	A random seed for generating four rotation angles(R_1, R_2, R_3,R_4)
FS	A four-digit secret number, $FS = P_1P_2P_3P_4$
S_1	The base of MU

tocol run. AS then figures out four rotation angles (R_1, R_2, R_3, R_4) according to the received RS. With these four angles, AS employs the procedure of subsection 3.3.2 to construct four temporary shares S_{2a}, S_{2b}, S_{2c}, and S_{2d} containing P_1, P_2, P_3, and P_4, respectively.

Step 3. AS returns $S_2 = S_{2d}$ to MU.

Step 4. MU applies RS to obtain (R_1, R_2, R_3, R_4) and then stacks S_2 on S_1. Rotating S_2 with R_4, P_4 could be generated. Keeping on rotating R_3, R_2, and R_1 sequentially, MU could have P_3, P_2, and P_1. MU keys in ID_{MU} and FS= (P_1, P_2, P_3, P_4) into the web page on PC.

Step 5. S checks if (ID_{MU}, FS) is the same as the one sent from AS. If they are correct, S accepts the request; otherwise, the connection is terminated.

3.3 VSS Share Generation Phase

Here we explain how to achieve the arbitrary-angle VSS stacking. As we use the concept of strongbox to lay out the OTP content, all the shares and base are in the shape of circle. The secret base S_1 kept in MU and AS is constructed in subsection 3.3.1, while the share S_2 is generated in subsection 3.3.2. All pixels are divided into eight types of 3×3 block, as shown in Table 4. Note that P_1, P_2, P_3, and P_4 are displayed in four pictures, and the content of these four pictures are located separately. If the stacked result of a block contains four white pixels, it is regarded as the white pixel of an OTP digit. In case that the stacked results of a block possesses nine black pixels, it is considered as the black pixel of an OTP digit.

Table 4: Block types

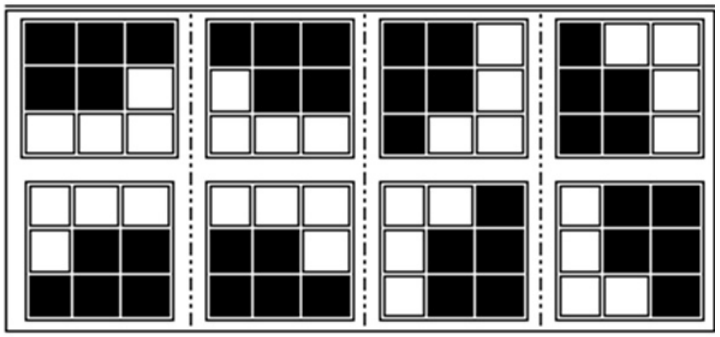
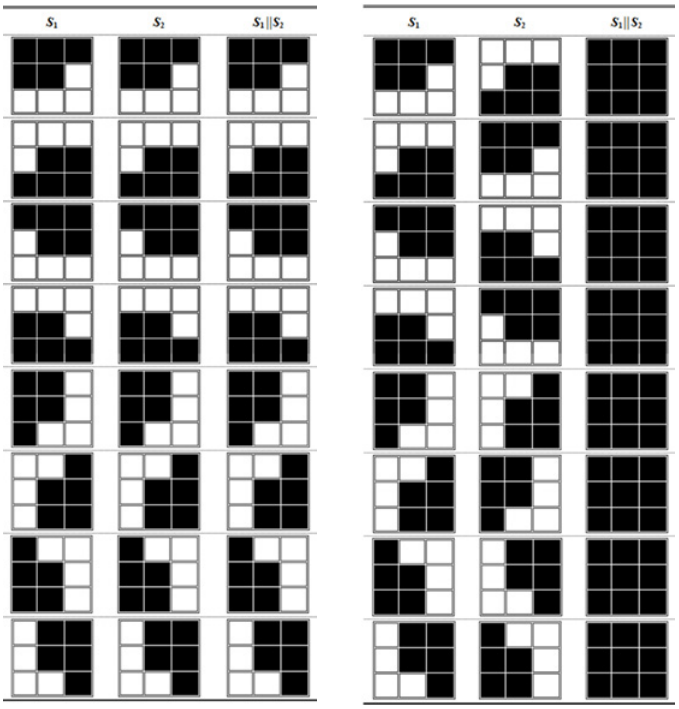


Table 5: Stacking rules

(a) Stacking rules of white blocks (b) Stacking rules of black blocks



3.3.1 The Generation of S_1

Once AS received and accepted the registration of MU, it randomly selects blocks from Table 4 to generate a secret base S_1, as displayed in Table 6(a). S_1 is sent to MU and is kept in its database related to ID_{MU}.

3.3.2 The Generation of S_2

The S_2 consists the blocks from Table 4 according to the rules of Table 5. For a white pixel of an OTP digit, AAVSS generates a block of share corresponding to the position of S_1 based on Table 5(a) such that the stacked result ($S_1 \vee S_2$) could be white. As to a black pixel of an OTP digit, AAVSS picks a block of share corresponding to the position of S_1 based on Table 5(b) such that the stacked result ($S_1 \vee S_2$) could be black.

In the following, we describe how to generate four tem-

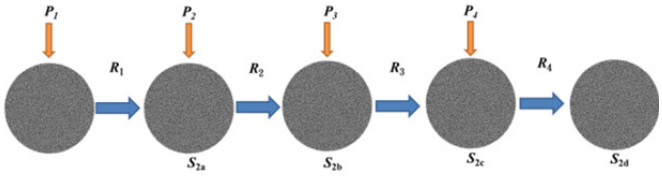


Figure 4: The procedure of generating four shares

porary shares to achieve arbitrary-angle stacking. The procedure is illustrated in Figure 4.

Step 1. $P_1P_2P_3P_4$. AAVSS random picks up four digits(P_1, P_2, P_3, P_4) as the OTP of this section.

Step 2. S_{2a}. AAVSS refers to the block positions of S_1 and P_1 to generate corresponding block of the first share. For a white pixel of P_1, AAVSS randomly selects a block of Table 4 for the corresponding position in the share. As to a black pixel of P_2, it picks the opposite block of S_1 for the corresponding position in the share. According to NNI, it rotates this share by the angle $(-R_1)$ to form S_{2a}.

Step 3. S_{2b}. AAVSS refers to the block positions of S_1 and P_2 to generate corresponding block of the second share. Note that we only consider the black pixel of P_2, P_3, and P_4 to avoid pixel interference problem. For a black pixel of P_2, AAVSS selects the opposite block of S_1 and replaces the corresponding block of S_{2a} with this one. According to NNI, it rotates this share by the angle $(-R2)$ to form S_{2b}.

Step 4. S_{2c}. AAVSS refers to the block positions of S_1 and P_3 to generate corresponding blocks of the third share. For a black pixel of P_3, AAVSS chooses the opposite block of S_1 and replaces the corresponding block of S_{2b} with this one. According to NNI, it rotates this share by the angle $(-R_3)$ to form S_{2c}.

Step 5. S_{2d}. AAVSS refers to the block positions of S_1 and P_4 to generate corresponding block of the last share. For a black pixel of P_4, AAVSS finds the opposite block of S_1 and replaces the corresponding block of S_{2c} with this one. According to NNI, it rotates this share by the angle $(-R_4)$ to form S_{2d}.

Step 6. $S_2 = S_{2d}$, as shown in Table 6(b).

3.4 Performance Analysis

To prove the practicability of AAVSS, we have simulated the system to examine the performance. In particular, the sizes of base and share are set to be 128×128 pixel, 192×192 pixel, 256×256 pixel, 384×384 pixel, and 512×512 pixel, which are suitable for the mobile device. We use desktop to simulate the AS. CPU used for the server is AMD FX-6300 Six-Core 3.5GHz with 8GB RAM.

Table 6: S_1 and S_2

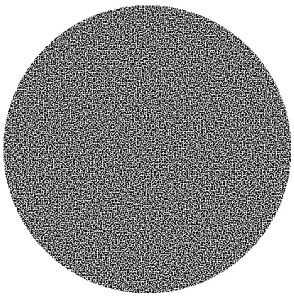
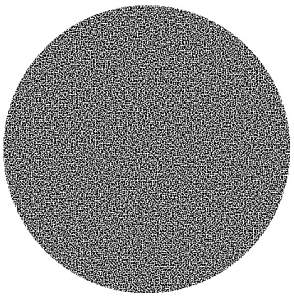
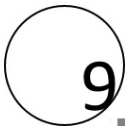
(a) S_1	(b) S_2
	

Table 7: Secret digits of an OTP

P_1	P_2	P_3	P_4
			

The operating system is Window 7 with 64bit. The used program language is Microsoft Visual Studio 2013 C++. The examination includes the recognition of stacked results and the generation efficiency of base and share. We first generated ten OTP's (i.e., FS's) which contain four digits (P_1, P_2, P_3, P_4). Each four-digit number is spread on four pictures, as shown in Table 7.

For each base size, we have generated a base S_1 according to Table 4. The results are displayed in Table 8. For each base, we have generated two sets of (S_{2a}, S_{2b}, S_{2c}, S_{2d}) according to two sets of (P_1, P_2, P_3, P_4) and two sets of rotation angle (R_1, R_2, R_3, R_4). The outcomes are illustrated in Table 9. Undoubtedly, nothing could be revealed. The details of rotation angles and stacked outcomes are listed in Table 10. Here is an example to illustrate how to reveal the OTP in the first row of Table 10. Given the set of rotation angle (139°, 192°, 85°, 301°), we first stacked S_2 on S_1 and rotated S_2 by 301°. Then we obtained a recognizable digit 4. We kept on rotating S2 by 85° to have digit 9, by 192° to reveal digit 2, and by 139° to extract digit 3. Note that 3294 is the OTP for this protocol run. It is clear that all the operations could lay out a recognizable digit from the human vision perception. Thus, the arbitrary-angle stacking is confirmed in AAVSS.

To highlight the contribution of AAVSS, we also inspected the efficiency for constructing base and shares.

Table 8: S_1 with different size (Pixel)

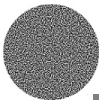
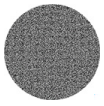
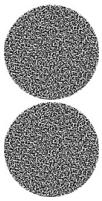
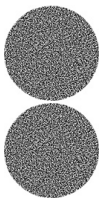
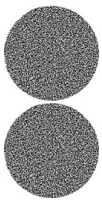
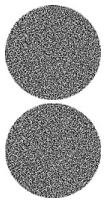
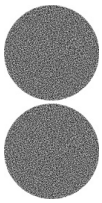
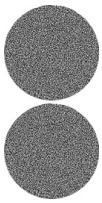
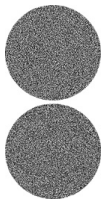
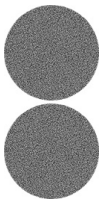
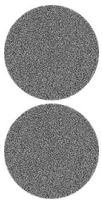
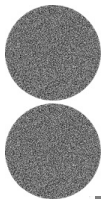
128×128	192×192	256×256	384×384	512×512
				

Table 9: S_2 with different size (Pixels)

Size	S_{2a}	S_{2b}	S_{2c}	S_{2d}
128 × 128				
192 × 192				
256 × 256				
384 × 384				
512 × 512				

For each size of test image, we performed the system two thousand times to gather the statistics. Two thousands random seeds are generated to produce two thousand sets of rotation angles. Similarly, two thousand sets of secret numbers (OTPs) are embedded into corresponding shares. The average time each step is shown in Table 11.

The generation of base S_1 is quick as it is constructed randomly according to the blocks of Table 4. The time for producing the first temporary share S_{2a} is longer since we need to consider the black and white pixels of the secret digits. By contrast, the time for creating S_{2b}, S_{2c}, and S_{2d} is shorter as only the black pixels of secret digit are referred to.

In Table 11, S_2 displays the time for completing a share that will be transferred over the Internet. Namely, it is the time summation of generating those four temporary shares. We have selected five sizes of test image, which are suitable for smartphone appearance. Actually, it is easy for people to recognize the content of stacked result,

Table 10: Stacked results

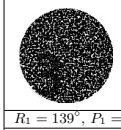
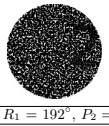
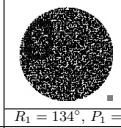
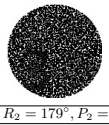
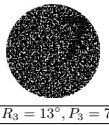
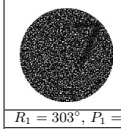
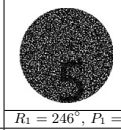
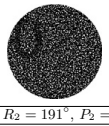
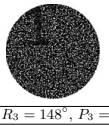
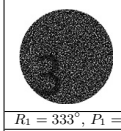
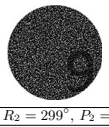
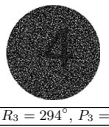
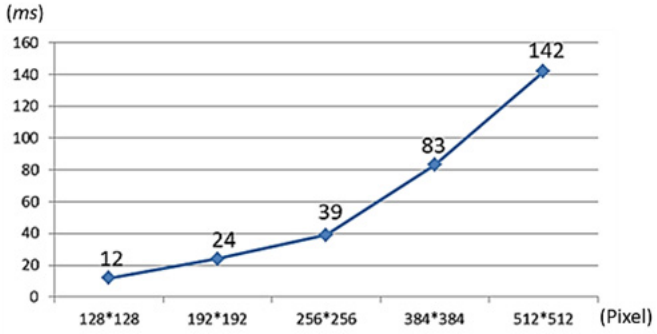
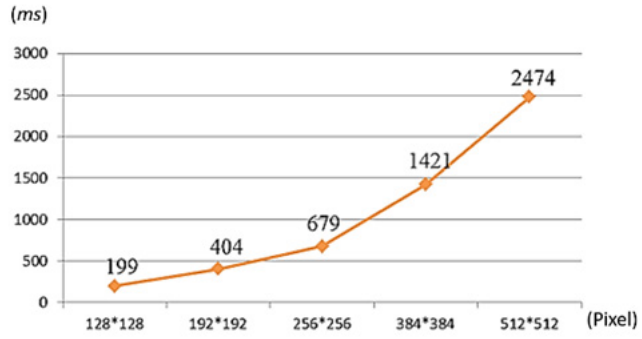
Size	$S_1 \vee S_{2a}$	$S_1 \vee S_{2b}$	$S_1 \vee S_{2c}$	$S_1 \vee S_{2d}$
128 × 128	 $R_1 = 139°, P_1 = 3$	 $R_1 = 192°, P_2 = 2$	 $R_3 = 85°, P_3 = 9$	 $R_4 = 301°, P_4 = 4$
	 $R_1 = 134°, P_1 = 0$	 $R_2 = 179°, P_2 = 3$	 $R_3 = 13°, P_3 = 7$	 $R_4 = 131°, P_4 = 9$
192 × 192	 $R_1 = 303°, P_1 = 7$	 $R_2 = 114°, P_2 = 6$	 $R_3 = 173°, P_3 = 4$	 $R_4 = 221°, P_4 = 9$
	 $R_1 = 246°, P_1 = 5$	 $R_2 = 191°, P_2 = 0$	 $R_3 = 148°, P_3 = 1$	 $R_4 = 65°, P_4 = 8$
256 × 256	 $R_1 = 333°, P_1 = 3$	 $R_2 = 299°, P_2 = 9$	 $R_3 = 294°, P_3 = 4$	 $R_4 = 346°, P_4 = 5$
	$R_1 = 282°, P_1 = 2$	$R_2 = 321°, P_2 = 1$	$R_3 = 187°, P_3 = 7$	$R_4 = 126°, P_4 = 0$
384 × 384	$R_1 = 156°, P_1 = 3$	$R_2 = 194°, P_2 = 5$	$R_3 = 90°, P_3 = 2$	$R_4 = 257°, P_4 = 8$
	$R_1 = 34°, P_1 = 1$	$R_2 = 96°, P_2 = 4$	$R_3 = 25°, P_3 = 9$	$R_4 = 207°, P_4 = 7$
512 × 512	$R_1 = 316°, P_1 = 5$	$R_2 = 279°, P_2 = 8$	$R_3 = 33°, P_3 = 7$	$R_4 = 114°, P_4 = 2$
	$R_1 = 40°, P_1 = 1$	$R_2 = 54°, P_2 = 2$	$R_3 = 254°, P_3 = 9$	$R_4 = 46°, P_4 = 6$

Table 11: Time for generating base and shares

Time *(ms)*						
Size (Pixel)	S_1	S_{2a}	S_{2b}	S_{2c}	S_{2d}	S_2
128×128	12	51	50	49	49	199
192×192	24	107	99	99	99	404
256×256	39	182	166	165	166	679
384×384	83	388	344	345	344	1421
512×512	142	673	600	601	600	2474



(a) S_1



(b) S_2

Figure 5: The fold-line graphs of base and shares

as shown in Table 10, no matter what kinds of size we selected. Taking the size of 256×256 for instance, the average time for creating a share for MU is only 679 ms, which is acceptable in integrating a mobile phone as the second factor of authentication. Again, we have provided the fold-line graphs of efficiency under different sizes in Figure 5. They can help to make sense of the practicability of AAVSS in playing an auxiliary authority procedure.

3.5 Conclusions

In this article, we first used VSS technique with arbitrary-angle to develop an auxiliary procedure for preventing real password from being recorded in a public computer. In particular, NNI is adopted to achieve the arbitrary-angle stacking, which can be considered as a strongbox. The security of the applied VSS can be referred to that of [8]. With limited time and trial constraints, AAVSS can effectively produce an OTP to protect secret information when the personal computer is inaccessible.

References

[1] H. C. Chao and T. Y. Fan, "Random-grid based progressive visual secret sharing scheme with adaptive priority," *Digital Signal Processing*, vol. 68, pp. 69–80, 2017.

[2] D. Chen, N. Zhang, Z. Qin, X. Mao, Z. Qin, X. Shen, and X. Y. Li, "S2m: A lightweight acoustic fingerprints-based wireless device authentication protocol," *IEEE Internet of Things Journal*, vol. 4, no. 1, pp. 88–100, 2017.

[3] N. Haller, *The S/Key One-Time Password System*, RFC 1760, 1995.

[4] A. E. Jr, W. Kantrowitz, and E. Weiss, "A user authentication scheme not requiring secrecy in the computer," *Communications of the ACM*, vol. 17, no. 8, pp. 437–442, 1974.

[5] M. A. Khan, "A survey of security issues for cloud computing," *Journal of Network and Computer Applications*, vol. 71, pp. 11–29, 2016.

[6] D. Kreutz, O. Malichevskyy, E. Feitosa, H. Cunha, R. da R. Righi, and D. D. J. de Macedo, "A cyber-resilient architecture for critical security services," *Journal of Network and Computer Applications*, vol. 63, pp. 173–189, 2016.

[7] T. Limbasiya, M. Soni, and S. K. Mishra, "Advanced formal authentication protocol using smart cards for network applicants," *Computers & Electrical Engineering*, vol. 66, pp. 50-63, 2018.

[8] K. S. Lin, C. H. Lin, and T. H. Chen, "Distortionless visual multi-secret sharing based on random grid," *Information Sciences*, vol. 288, pp. 330–346, 2014.

[9] Y. J. Liu, C. C. Chang, and S. C. Chang, "An efficient and secure smart card based password authentication scheme," *International Journal of Network Security*, vol. 19, no. 1, pp. 1-10, 2016.

[10] M. Naor and A. Shamir, "Visual cryptography," in *Workshop on the Theory and Application of of Cryptographic Techniques*, pp. 1–12, 1994.

[11] S. Sagiroglu and G. Canbek, "Keyloggers," *IEEE Technology and Society Magazine*, vol. 28, no. 3, 2009.

[12] R. Song, "Advanced smart card based password authentication protocol," *Computer Standards & Interfaces*, vol. 32, no. 5-6, pp. 321–325, 2010.

[13] H. M. Sun, Y. H. Chen, and Y. H. Lin, "oPass: A user authentication protocol resistant to password stealing and password reuse attacks," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 2, pp. 651–663, 2012.

[14] S. Teltscher, E. Magpantay, I. Vallejo, L. Kreuzenbeck, D. Korka, V. Gray, D. Olaya, M. Hilbert, M. Minges, N. Delmas, *et al.*, "Measuring the information society," *Geneva: International Telecoms Union*, 2013. (`https://www.itu.int/en/ITU-d/Statistics/Documents/publications/mis2013/MIS2013_without_Annex_4.pdf`)

[15] C. C. Wu, S. J. Kao, and M. S. Hwang, "A high quality image sharing with steganography and adaptive authentication scheme," *Journal of Systems and Software*, vol. 84, no. 12, pp. 2196–2207, 2011.

Biography

Ying-Chin Chen is pursuiting her MS degree in Information Engineering and Computer Science in Feng Chia University, Taichung, Taiwan. Her current research interests include information security and visual secret sharing.

Kuo-Jui Wei received the Ph.D. degree in Information Engineering and Computer Science in 2016 from Feng Chia University, Taichung, Taiwan. His current research interests include information security and mobile communications.

Jung-San Lee has worked as a professor in the Department of Information Engineering and Computer Science at Feng Chia University, Taiwan, since 2017. His current research interests include information security and mobile communications.

Ngoc-Tu Huynh is a Lecturer with Faculty of Information Technology, Ton Duc Thang University, Ho Chi Minh City, Vietnam. Her current research interests include information security and cryptography.

Jyun-Hong Lin received MS degree in information engineering and computer science in Feng Chia University, Taichung, Taiwan in 2015. His current research interests include image processing, and network security.

E-commerce Trade Consumption Payment Security and Privacy Based on Improved B2C Model

Linzhu Hu

(Corresponding author: Linzhu Hu)

Chongqing University of Science and Technology
Campus city, Shapingba district, Chongqing 401331, China
(Email: lzhu1981@yeah.net)

(Received Sept. 22, 2018; revised and accepted Apr. 11, 2019; First Online May 26, 2019)

Abstract

With the popularity of the Internet and smart phones, e-commerce based on the Internet has rapidly developed by relying on its particular merits. However, the openness of the Internet makes payment security and privacy protection become the key of e-commerce development. This study gave a brief introduction of both traditional and improved Business to Customer (B2C) e-commerce and preformed the analogue simulation on shift left long (SLL) security protocol based on double encryption algorithm and traditional encryption algorithm under different sizes of data. The result showed that the double encryption algorithm could has lower complexity for encrypting and decrypting data, enabling to shorten the time of the encryption and decryption of the data; in terms of security, the decryption integrity of the data that was encrypted by double encryption algorithm was lower, and was basically garbled without logic. Thus, the security is guaranteed. In conclusion, the third-party privacy server in the improved B2C model can effectively guarantee the payment and privacy security of consumers.

Keywords: Business To Customer; Double Encryption Algorithm; E-Commerce; Payment Security

1 Introduction

With the popularity of the Internet, e-commerce, which is different from traditional commerce, has gradually developed. With the help of the Internet, e-commerce can initiate business transactions anytime and anywhere, and no physical cash is needed in this process [1]. However, for business operation, whether traditional or electronic, the most important thing is the protection of information [2], including transaction fund and personal information of both buyers and sellers [5, 15].

Traditional business [18, 21] is based on the real world, and in a state of "face to face", the buyer and the seller can completely rely on the only biological characteristic to confirm the information's reality and safety, but buyers and sellers of the electronic commerce with the virtual Internet cannot meet directly, so security protocol is used to ensure information security certification [6]. Studies on the e-commerce security are as follows. Yi *et al.* [14] brought up a formal analysis method to verify quantum cryptography electronic payment protocol security. The results showed that the agreement was not satisfactory because of the logical flaws. After improving and using formal analysis to verify again, it could be found out that defects were made up for. Mandal [16] put forward a kind of electronic payment system based on authentication key exchange protocol.

This case introduced an effective owner tracking mechanism to identify the malicious customers. At the same time, the automatic validation of the Internet security protocols and applications simulated the security of the scheme to prove that its replay and man-in-the-middle attack were safe. Mlke *et al.* [8] suggested to use a kind of privacy protection e-commerce protocol (PPEP) which would decouple or unlock online trade and consumer identity to provide anonymity for online shoppers in the e-commerce websites. What's more, they also brought up a PPEP plan which enabled merchants to perform customer management without disclosing the identity of customers to merchants. This study briefly introduced the traditional Business to Customer (B2C) and the improved B2C e-commerce model and simulated shift left long (SLL) security protocol based on double encryption algorithm and traditional encryption algorithm under different sizes of data.

2 Traditional B2C Model E-Commerce

As shown in Figure 1, the fundamental frame structure of traditional B2C mode [4, 20] consisted of the third-

party payment platform, buyer browser, seller website and logistics platform [12]. The execution flow of traditional B2C mode transaction protocol [7, 11] is shown by the ordinal arrow in Figure 1.

1) The buyer looks through goods in the seller website through the browser;
2) The seller provide goods information for buyers in the websites;
3) After logging in the website, buyer places an order for the goods and chooses the third-party payment platform to pay;
4) After receiving the payment order, the seller submits it to the third party platform;
5) The buyer confirms the payment transfer operation of the order in the third-party platform;
6) The third-party platform feed back the payment processing results to both the seller and the buyer;
7) The seller issues and processes orders and submits processing information to the third-party platform;
8) The buyer confirms the receipt of goods on the third party platform after he has received the goods satisfactorily;
9) The third-party platform transfers the buyer's payment to the seller's account.

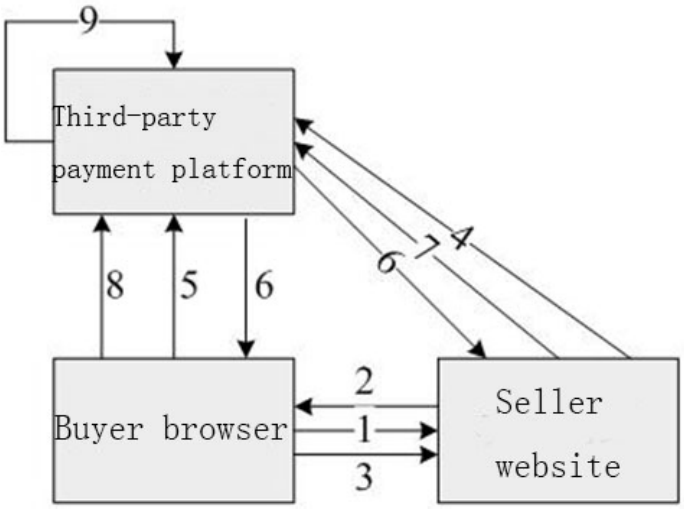


Figure 1: The traditional B2C mode model

Traditional B2C model adopts the third-party payment platform to ensuring the payment security, however, it still has some disadvantages in the physical application [3, 17]:

1) The seller is eager to deliver the goods after receiving the order without confirming the payment information of the buyer;
2) After the seller issues the goods, the buyer cancels the order due to malice or unexpected factors, resulting in the seller's property and goods being empty;
3) Because of the logistics platform, the seller delivers the goods, but the buyer who is "received" has not actually received the goods;
4) The order information of the buyer can be found on all three platforms in the circulation process, increasing the risk of privacy disclosure.

3 Improved B2C Mode E-Commerce

As shown in Figure 2, to solve the four shortcomings of the traditional mode mentioned above, the traditional B2C trade mode was expanded by third-party privacy server [9] and logistics platform, and the original functions of modules remain unchanged.

The execution flow of improved B2C mode transaction protocol is shown by the ordinal arrow in Figure 2 [13]:

1) The buyer browses the goods on the seller's website through the browser;
2) The seller provides the buyer with the commodity information on the website;
3) The buyer registers the address and other privacy information in the third-party privacy server, and obtains the corresponding ID serial number;
4) The third-party privacy server transfers the order to the seller;
5) The seller transfers the received order information to the third-party payment platform, where the privacy information in the order is replaced by the ID serial number obtained before;
6) Payment platform transfers the results of feedback to buyer and seller;
7) The seller delivers goods according to the order;
8) The logistics platform informs the buyer that the goods are received;
9) The logistics platform notifies the buyer's received information to the third-party privacy server;
10) Attending logistics platform can't learn the buyer's private information.

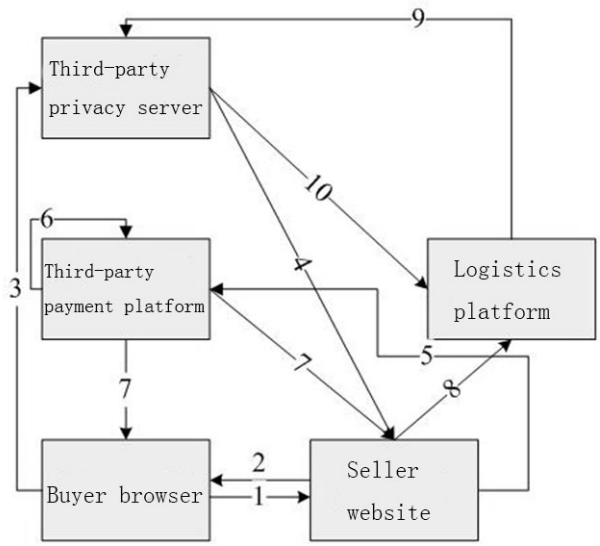


Figure 2: The improved B2C mode model

4 Double Encryption Algorithm

In the traditional B2C e-commerce model, the data interaction between modules is through Hyper Text Transfer Protocol (HTTP), but the data transferred by HTTP protocol transmission in general is the plaintext which has no encryption process. When conducting e-commerce transactions, data transferred is extremely easy to be intercepted or faked by a third party, meanwhile, the both sides of transmitting and receiving information can't confirm identity of each other. In the improved B2C e-commerce model, third-party privacy server and logistics platform are added. The third-party privacy server provides the whole model with Secure Socket Layer (SSL) security agent protocol based on double encryption algorithm [22]. SSL protocol can provide secure communication privacy protection for both sides of data transmission.

As shown in Figure 3, the third-party privacy server will judge the data type after receiving it from the buyer's browser, and if it is PI, the flag bit of SSL will be flag1=0, flag2=0; then the public key of the payment gateway in the SSL protocol is used to encrypt the PI, and obtain the payment encryption package CPI, which is then filled into the actual data recorded in the SSL protocol. If it is an order information OI, it is populated directly into the actual data in the SSL record; after obtaining the actual data recorded by SSL, the Hash function algorithm is applied to perform summary calculation on the actual data, the sequence generated by the sequence generator and the encryption key of PI. The obtained summary data was MAC data. The SSL record generated in the first few steps shall be encrypted by applying the symmetric key [10,19] in the SSL protocol negotiated by both parties to generate the transmission ciphertext Crecord-SSL.

The double encryption algorithm mentioned above is

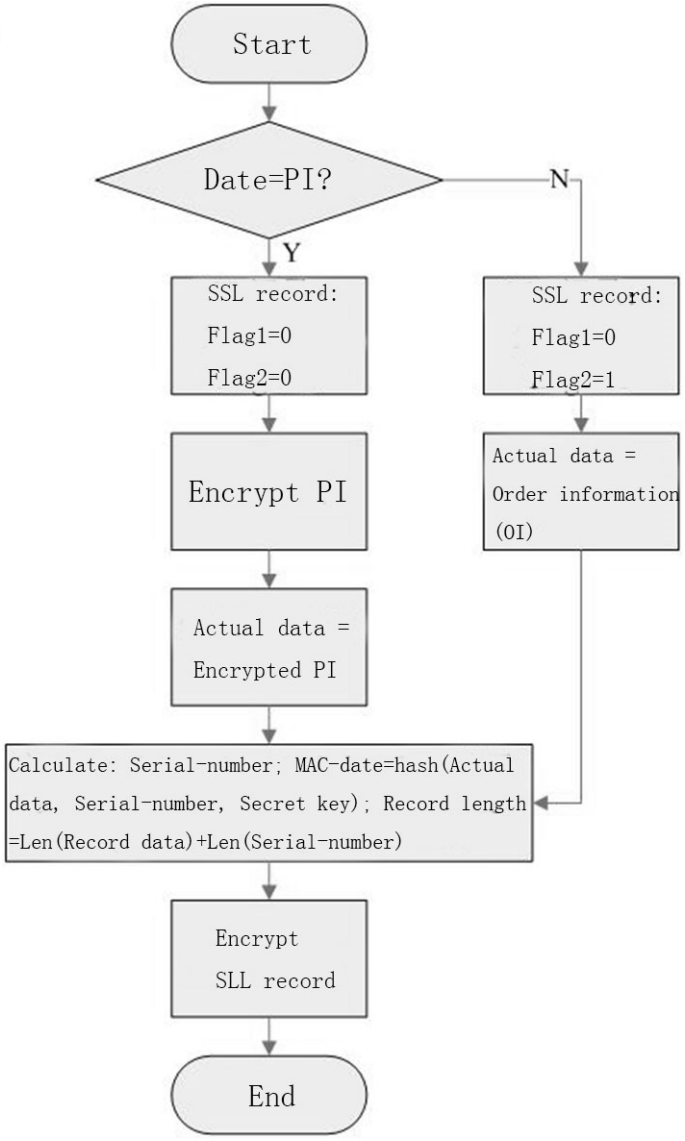


Figure 3: Double encryption algorithm flow

applied in the third-party privacy server to obtain the ciphertext Crecord-SSL of payment information and order information, and then transmitted to the server of the seller's website. After receiving ciphertext Crecord-SSL, the server performs symmetric decryption according to the negotiated symmetric key, then checks the integrity of SSL record data, and judges whether the record data is order information or payment information according to the flag bit. If it is the order information, the server will extract the information and put it into storage for processing. In the case of payment information, the payment encryption package CPI will be transmitted to the third-party payment platform, and the public key of the payment gateway is used to decrypt it and wait for the payment result. If successful, the logistics platform is informed to deliver the goods.

5 Simulation Experiment

5.1 Experiment Environment

The experiments in this study were performed on a lab server with server configuration of Windows 7 system, I7 processor, and 16 Gbytes of memory. The coding of SSL security protocol based on double encryption algorithm and SSL security protocol based on double encryption algorithm was implemented using C++.

5.2 Experiment Methods

Data packets with different sizes of order information and payment information were set, and data packets were encrypted and decrypted through SSL security protocol based on double encryption algorithm and SSL security protocol based on double encryption algorithm. The experiment was repeated 100 times and the average of the total time required to encrypt and decrypt the data packets under both algorithms was counted.

Similarly, data packets with different sizes of order information and payment information were set, and the data packets were encrypted respectively through SSL security protocol based on double encryption algorithm and SSL security protocol based on double encryption algorithm. Then the encrypted data packets were informally decrypted to simulate the situation where the orders and payment information were stolen, meanwhile, the maximum decryption time was set as 60 min to prevent the decryption time from being too long. The cracked ciphertext was compared with the original text to obtain the decryption integrity.

5.3 Experiment Results

5.4 Time Complexity

As shown in Figure 4, for a data packet of 1 M, the total time required for encryption and decryption by the traditional encryption algorithm was 78.7 ms, and the total time of the double encryption algorithm was 31.2 ms; for a data packet of 10 M, the traditional encryption algorithm required 600.3 ms, and double encryption algorithm required 245.1 ms; for a data packet of 20 M, the traditional encryption algorithm needed 1181.5 ms, the double encryption algorithm needed 487.2 ms; for a data packet of 30 M, the traditional encryption algorithm needed 1765.9 ms, the double encryption algorithm needed 695.3 ms; for a data packet of 40 M, the traditional encryption algorithm required 2377.5 ms, and the double encryption algorithm required 1103.2 ms. It could be seen that no matter which algorithm was, as the data to be encrypted increased, the total time required for encryption and decryption increased, and the difference of the required time between the two algorithms became increasingly obvious starting from 10 M, and the time required by the double encryption algorithm was significantly smaller than that of the traditional algorithm. It showed that the double encryption algorithm had lower time complexity and better performance.

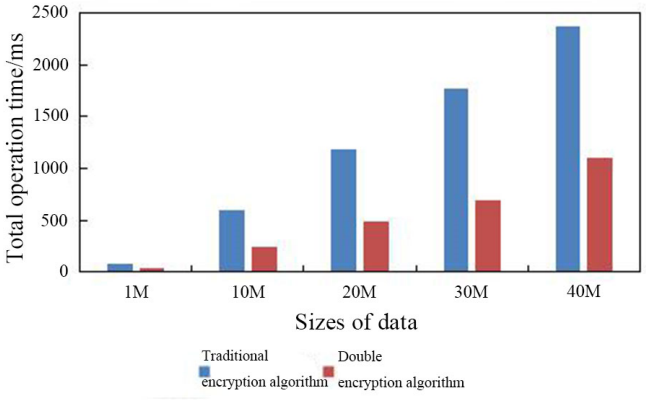


Figure 4: Total time of encryption and decryption of different sizes of data by different algorithms

5.5 Security Analysis

As shown in Figure 5, after 60 minutes of decryption, the integrity of the traditional encrypted data packet of 1 M was 10.2%, and the integrity of the double encrypted data packet was 8.1%; the integrity of the traditional encrypted data packet of 10 M was 8.2%, and the integrity of the double-encrypted data packet was 5.3%; the integrity of the traditional encrypted data packet of 20 M was 5.1%, and the integrity of the double-encrypted data packet was 3.2%; the integrity of the traditional encrypted data packet of 30 M was 2.2%, and the integrity of the double-encrypted data packet was 0.8%; the integrity of the traditional encrypted data packet of 40 M was 0.9%, and the integrity of the double encrypted data packet was 0.2%. It could be seen that with the increasing of the encrypted data packet, the integrity of the decrypted data was significantly reduced. After the data packet of 1 M

was decrypted for 60 minutes, it could be seen that there were several logical characters. In the situation of the data packet of 20 M, only a few logic characters were available. In the situation of the data packet of 40 M, the decrypted data was basically garbled. Both algorithms could prevent decrypting to a certain extent, and the decrypted data which was encrypted by double encryption algorithm had lower complexity and higher safety.

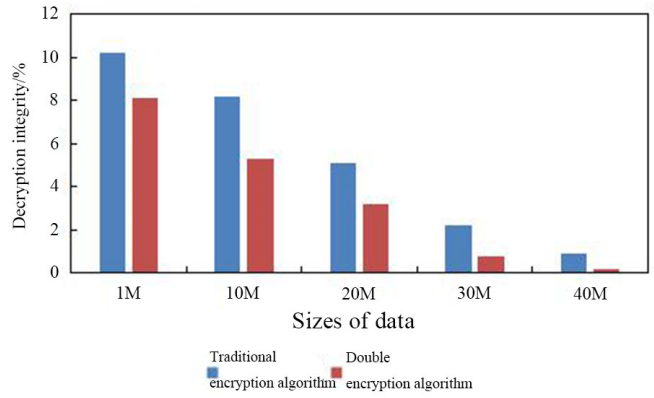


Figure 5: Security of encryption of different sizes of data by two algorithms

6 Conclusion

This article simply introduced the traditional B2C and improved B2C of e-commerce model. The improved electronic business model of B2C compared with the traditional one increased two modules as a third party privacy and logistics platform. Third party privacy server used SSL security protocol based on double encryption algorithm to improve the payment security and privacy protection of e-commerce. Then, the performance of SLL security protocol based on double encryption algorithm and traditional encryption algorithm in encrypting data of different sizes was simulated. The result was that the total time required for encryption and decryption of both algorithms increased with the increase of encrypted data.

The double encryption algorithm had less time complexity and less total time for decryption and encryption, and was more suitable for private information exchange of e-commerce. It was found that the data packet with larger size had significantly reduced decryption complexity after 60 min decryption of the data packet encrypted by the two encryption algorithms. Both algorithms could prevent decrypting to some extent, and the data that was encrypted by the double encryption algorithm had lower decryption integrity and higher security. To sum up, the added third-party privacy server based on double encryption algorithm in the improved B2C e-commerce mode can ensure the security of payment and privacy.

References

[1] S. A. Chaudhry, M. S. Farash, H. Naqvi, *et al.*, "A secure and efficient authenticated encryption for electronic payment systems using elliptic curve cryptography," *Electronic Commerce Research*, vol. 16, no. 1, pp. 113–139, 2016.

[2] Z. Djuric, D. Gasevic, "FEIPS: A secure fair-exchange payment system for internet transactions," *The Computer Journal*, vol. 58, no. 10, pp. 2537–2556, 2015.

[3] T. H. Feng, M. S. Hwang, and L. W. Syu, "An authentication protocol for lightweight NFC mobile sensors payment," *Informatica*, vol. 27, no. 4, pp. 723–732, 2016.

[4] E. Y. Huang, C. J. Tsui, "Assessing customer retention in B2C electronic commerce: an empirical study," *Journal of Marketing Analytics*, vol. 4, no. 4, pp. 172–185, 2016.

[5] M. S. Hwang, C. C. Lee, Yan-Chi Lai, "Traceability on low-computation partially blind signatures for electronic cash", *IEICE Fundamentals on Electronics, Communications and Computer Sciences*, vol. E85-A, no. 5, pp. 1181–1182, May 2002.

[6] M. S. Hwang, Li-Hua Li, "A New Remote User Authentication Scheme Using Smart Cards", *IEEE Transactions on Consumer Electronics*, vol. 46, no. 1, pp. 28–30, Feb. 2000.

[7] M. S. Hwang, I. C. Lin, L. H. Li, "A simple micropayment scheme", *Journal of Systems and Software*, vol. 55, no. 3, pp. 221–229, Jan. 2001.

[8] M. Ike, K. Sarac, "PPEP: A deployable privacy preserving e-commerce protocol for electronic goods," in *International Conference on Communication & Network Security*, 2016.

[9] S. E. Kaplan, R. J. Nieschwietz, "A web assurance services model of trust for B2C e-commerce," *International Journal of Accounting Information Systems*, vol. 4, no. 2, pp. 95–114, 2015.

[10] M. M. Kiani, A. Raza, K. D. Gill, "Centralized collaborative reputation model for B2C E-Commerce," *17th IEEE International Multi Topic Conference*, pp. 450–455, 2014.

[11] N. Knego, "Importance of assortment for B2c electronic commerce in some EU countries," *Economy & Business Journal*, vol. 10, no. 1, pp. 94–102, 2016.

[12] I. C. Lin, M. S. Hwang, C. C. Chang, "The general pay-word: A micro-payment scheme based on n-dimension one-way hash chain", *Designs, Codes and Cryptography*, vol. 36, no. 1, pp. 53–67, July 2005.

[13] J. Ling, M. Jun, Z. Yang, "Customer-perceived value and loyalty: how do key service quality dimensions matter in the context of B2C e-commerce?," *Service Business*, vol. 10, no. 2, pp. 301–317, 2016.

[14] Y. Liu, X. Liu, J. Wang, *et al.*, "Security analysis of electronic payment protocols based on quantum cryptography," in *International Conference on Information Science & Control Engineering*, IEEE, 2017.

[15] J. W. Lo, H. M. Lu, T. H. Sun, and M. S.Hwang, "Improved on date attachable electronic cash," *Applied Mechanics and Materials*, vol. 284, pp. 3444–3448, 2013.

[16] S. Mandal, S. Mohanty, B. Majhi, "Design of electronic payment system based on authenticated key exchange," *Electronic Commerce Research*, vol. 18, no. 2, pp. 359–388, 2018.

[17] D. L. Paris, M. Bahari, N. A. Iahad, "Business-to-customer (B2C) electronic commerce: An implementation process view," in *3rd International Conference on Computer & Information Sciences*, pp. 19–24, 2016.

[18] M. Pasquet, S. Gerbaix, "Instant payment versus smartphone payment: The big fight?," in *IEEE Third International Conference On Mobile And Secure Services (MobiSecServ'17)*, pp. 1–3, 2017.

[19] L. G. Pee, "Customer co-creation in B2C e-commerce: does it lead to better new products?," *Electronic Commerce Research*, vol. 16, no. 2, pp. 1–27, 2016.

[20] M. S. Hwang and P. C. Sung, "A study of micropayment based on one-way hash chain", *International Journal of Network Security*, vol. 2, no. 2, pp. 81–90, Mar. 2006.

[21] S. Walczak, G. L. Borkan, "Personality type effects on perceptions of online credit card payment," *Journal of Theoretical and Applied Electronic Commerce Research*, vol. 11, no. 1, pp. 5, 2016.

[22] X. Wang, Y. Jia, L. Guo, "Study on the function of computer technology in the electronic commerce environment security and risk assessment," in *International Conference on Intelligent Transportation*, pp. 784–786, 2016.

Biography

Linzhu Hu, born in 1981, gained the master's degree from China University of Petroleum. She is working as a lecturer in Chongqing University of Science & Technology. Her research direction is international business and trade. She had hosted and participated in the project: In 2013, she participated in the school-level education reform project "Research on the Construction of Applied Talents Training Platform for International Economics and Trade Majors Based on Discipline Competition"; In 2015, she participated in the school-level teaching reform project "Application Research of CLIL Bilingual Teaching Model Based on Applied Undergraduate Brand Characteristics"; In 2016, she guides students to participate in the first prize of the Cross-Strait College Students International Trade Competition; In 2017, she hosted the school-level "Business Negotiation Practice" "School Plan" project.

Efficient Hierarchical Key Management Scheme for VR/AR Database Systems

Tsung-Chih Hsiao[1], Yu-Min Huang[2], Yu-Fang Chung[3], Tzer-Long Chen[4], and Tzer-Shyong Chen[5]
(Corresponding author: Tzer-Shyong Chen)
School of Arts, Southeast University, Nanjing, China[1]
Department of Statistics, Tunghai University, Taiwan[2]
Department of Electrical Engineering, Tunghai University, Taiwan[3]
Department of Information Technology, Ling Tung University,Taichung, Taiwan[4]
Department of Information Management, Tunghai University, Taiwan[5]
(Email: arden@thu.edu.tw)
(Received Feb. 26, 2018; Revised and Accepted Aug. 13, 2018; First Online Mar. 9, 2019)

Abstract

With the growth of Virtual Reality (VR) and Augmented Reality (AR) in technologies such as artificial intelligence, wireless, 5G, big data, massive compute, industrial 4.0 and virtual stores. This paper improved the secure mechanism which existed some shortcomings. In order to accomplish the decentralized environment access control, it also proposed another new mechanism to achieve the requirements on the nonspecific internet. Besides, considering the security on storing and controlling and the use of administrative privileges of the VR and ARclouds is necessary. With the new mechanism, the problems such as mobile security or acting calculation which derived from VR and AR could be solved. This new research achieves a better circumstance. Developer staff's responsibility can be allocated; the systems can be compatibly integrated; on the other hand, the users' privacy of personal information can be strictly protected.

Keywords: Augmented Reality; Database System; Mobile Security; Privacy; Virtual Reality

1 Introduction

For the VR/AR data in the Internet standardization needs and standards system, research data security standards. Development of general requirements such as general requirements, architecture, testing and evaluation; development of common standards such as Internet and digital interconnection interface, logo resolution, data internet platform and security.The updating of VR/AR system and the insurance system, specifications of the Developer codes and information have several major problems. The consistency of the coding system and the data exchange format are not uniform and the expression ability in the VR/AR information system is quite lacking, etc. Under development of applicable VR/AR information, standards and practical application in the market, VR/AR information system needs to face the problem.

According to the number of VR/AR users, network size and other indicators, VR/AR user data security has become on of the the world's largest issue [6, 15].VR/AR information system equips a data system that is called VR/AR database system. Database is a set of related data collection, and the operation of the database must rely on the Database Management System, DBMS, to operate [5]. The database system is a program that controls the classification of the database and the access to the data [12]. According to the VR/AR information system, the transmitting of VR/AR information or related information, will inevitably use the network. Based on the problems of user's privacy, the developer enterprise's internal and external networks must be comprehensively planned. User's demand is divided into the following points: response, availability, quality, adaptability, security, affordability, expected growth. However, the safety of electronic VR/AR information has also become an important issue, especially for the user's access rights and in different time-range norms. We attempt to utilize mathematical methods to go through the data encryption and decryption which can strictly protect user's VR/AR information [1, 7].

DBMS can be divided into three types – hierarchical, network, relational [2]. The application of this information management system is widely used for relational purposes; however, due to the system used in developer enterprises rely on each other, the system will reject those people who attempt to get access to the user's information.

The feature of blockchain technologies may bring us more reliable and convenient services [9]. In a traditional public-key encryption, the sender has to authenticate that

the invoked public key is the legitimate public key for the intended receiver [10].Therefore, in this paper, we propose an integrated hierarchical access mechanism and the characteristics of the database system. By storing the decryption key in the Lagrange interpolation polynomial method, the VR/AR confidential information and users' privacy can be effectively protected [8]. A lot of related works have been proposed to solve access control problems [3, 4, 11, 13, 14].

2 Proposed Work

2.1 VR/AR Database Integration of VR/AR Systems

VR/AR systems contain a lot of information in the database, such as VR/AR records. As a result, VR/AR hardware and software communication are regulated. Due to different systems and equipments in different companies, leads to various incompatibilities between VR/AR devices and platforms. The Application Programming Interface standard describes how the VR/AR application or game engine renders its content and receives the data. If both of these core elements are standardized across all VR/AR hardware and software products, there will be an explosion in industry adoption and innovation.DBMS is mainly responsible for processing all data storage and retrieval operations. It can also modify data integration, data consistency check rules, controlling single or multi user's authorization, and data protection, etc. These are one part of the operation of the VR/AR system. For those people who intent to obtain information or even reveal other user's information, the system will cause compatibility obstacles to make the hackers unsuccessfully retrieve the user's information.

In order to make the system manager more convenient access to user information, we seek for access keys to secure confidential files while considering the safety issues in the transmission process. Therefore, this research method through the public encryption system and Lagrange interpolation of VR/AR data encryption protection, through the key authentication management center issued legal authority user decryption key, allowing users to access to the decryption key secret documents, strict management of the user data.

2.2 An Improved Access Scheme

Our goal is to construct the key allowing a server to access a particular document. We generalize the decryption polynomial $F_{DK_j}(x)$ subject to the following criterion (Table 1).

$$F_{DK_j}(x) = \begin{cases} DK_j, & \text{if server } S_i \text{ has permission} \\ & \text{to access } j \text{ document} \\ C, & \text{Otherwise} \end{cases} \quad (1)$$

for $C \neq DK_j$.

We aim to enhance the security over the decryption key to avoid potential exploration of information revealed by a third party.

2.3 Key Production

The decryption key can be generated through the following steps.

Step 1: Select large prime numbers p and q in random as the roots of finite field $GF(p)$. Number g and p remain public.

Step 2: Each confidential document will use non-repetitive decryption key DK_j, $j = 1, 2, \cdots, n$ with n denoting the number of documents.

Step 3: Choose non-repetitive secret key K_i, $i = 1, 2, \cdots, m$, where m is the number of servers which are about to visit confidential documents.

Step 4: The mobile agent owner uses a set of interpolation polynomial at with ID_j represents the number of DK_j. If $DK_i \leq S_i$, S_i has permission to get the decryption key DK_j. We construct $F_{DK_j}(x)$ as below.

$$F_{DK_j}(x) = x + DK_j - [\sum_{DK_j \leq S_i} x_{ij} l_{ij}(x) + \prod_{DK_j \leq S_i} a(l_{ij}(x))R], \quad (2)$$

where $l_{ij}(x)$ is the Lagrange interpolation polynomial formulated as:

$$l_{ij} = \prod_{t=1, t\neq i}^{m} (\frac{x - x_{1j}}{x_{ij} - x_{1j}}) \cdots (\frac{x - x_{i-1,j}}{x_{ij} - x_{i-1,j}})(\frac{x - x_{i+1,j}}{x_{ij} - x_{i+1,j}}) \cdots (\frac{x - x_{mj}}{x_{ij} - x_{mj}}) \quad (3)$$

We have Hash Function noted as:

$$a(l_{ij}(x)) = \begin{cases} l_{ij}(x) - 1, & \text{if } l_{ij}(x) = 1 \\ 1, & \text{Otherwise} \end{cases} \quad (4)$$

and R stands as a random real number.

2.4 Key Derivation

The decryption polynomial $F_{DK_j}(x)$ will derivate the decryption key through access permission from section above.

1) Server S_i providing a decription key DK_j for which the far-end server will be able to access the j document.

2) Server S_i substitutes its secret key K_i and decryption key ID_j for the public decryption polynomial equation $F_{DK_j}(x)$ to get DK_j. This access can be carried through the following derivation.

Table 1: Parameters for generating the decryption process

Symbols	Definition
CA	The key authentication management center
S_i	Server (System User)
ID_t	Number of confidential documents
K_i	The private key corresponding to each legitimate user
DK_t	Corresponds to the IDt's decryption key
$F_{DK_t}(x_i, t)$	Public decryption polynomial for retrieving decryption key

If $DK_i \leq S_i$, the secret key K_i provides x_{ij} and the Lagrange interpolation polynomial turns to be

$$l_{ij}(x_{ij}) = \prod_{t=1, t\neq i}^{m} \left(\frac{x - x_{tj}}{x_{ij} - x_{tj}}\right) = 1 \tag{5}$$

while the same x_{ij} we have $l_{ij}(x_{ij}) = 0$, for $i' \neq i$ or $j' \neq j$. This gives us the Hash Function as shown below:

$$a(l_{ij}(x_{ij})) = \begin{cases} l_{ij}(x_{ij}) - 1, & \text{for } i, j \\ 1, & \text{for } i' \neq i \text{ or } j' \neq j. \end{cases} \tag{6}$$

Thus, we have Equation (7):

$$\begin{aligned} \prod_{DK_j \leq S_i} a(l_{ij}(x_{ij}) &= 1 \cdots 1[l_{ij}(x_{ij}) - 1]1 \cdots 1 \\ &= 0. \end{aligned} \tag{7}$$

$$\sum_{DK_j \leq S_i} x_{ij} l_{ij}(x_{ij}) = x_{ij}. \tag{8}$$

Finally we put Equations (7), (8) into decryption polynomial in order to get decryption key DK_j:

$$F_{DK_j}(x_{ij}) = x_{ij} + DK_{ij} - x_{ij} = DK_j. \tag{9}$$

If $l_{ij} \neq 0, 1$, then we have

$$F_{DK_j}(x) = x + DK_j - [\sum_{DK_j \leq S_i} x_{ij} l_{ij}(x) + R]$$

Which means it would not be the desired decryption key DK_j either.

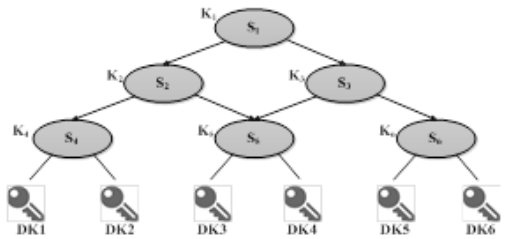


Figure 1: Access control architecture diagram for mobile agents

3 Example

According to the hypothetical hierarchical scheme (Figure 1), we assume that server S_6 has permission to access document 5 and 6. We demonstrate how the decryption process works to give the access key to the server. Suppose that we want to provide the decryption key to document 1, 3, 6 for server S_6, we have the following calculation shown below if S_6 uses input key x_{66}.

$$\begin{aligned} l_{16}(x_{66}) &= \left(\frac{x_{66} - x_{26}}{x_{16} - x_{26}}\right)\left(\frac{x_{66} - x_{36}}{x_{16} - x_{36}}\right)\left(\frac{x_{66} - x_{46}}{x_{16} - x_{46}}\right) \\ &\quad \left(\frac{x_{66} - x_{56}}{x_{16} - x_{56}}\right)\left(\frac{x_{66} - x_{66}}{x_{16} - x_{66}}\right) = 0 \end{aligned} \tag{10}$$

$$\begin{aligned} l_{36}(x_{66}) &= \left(\frac{x_{66} - x_{16}}{x_{36} - x_{16}}\right)\left(\frac{x_{66} - x_{26}}{x_{36} - x_{36}}\right)\left(\frac{x_{66} - x_{46}}{x_{36} - x_{46}}\right) \\ &\quad \left(\frac{x_{66} - x_{56}}{x_{36} - x_{56}}\right)\left(\frac{x_{66} - x_{66}}{x_{36} - x_{66}}\right) = 0 \end{aligned} \tag{11}$$

$$\begin{aligned} l_{66}(x_{66}) &= \left(\frac{x_{66} - x_{16}}{x_{66} - x_{16}}\right)\left(\frac{x_{66} - x_{26}}{x_{66} - x_{26}}\right)\left(\frac{x_{66} - x_{36}}{x_{66} - x_{36}}\right) \\ &\quad \left(\frac{x_{66} - x_{46}}{x_{66} - x_{46}}\right)\left(\frac{x_{66} - x_{56}}{x_{66} - x_{56}}\right) = 1. \end{aligned} \tag{12}$$

We also obtain values of Hash Function with $(l_{16}(x_{66})) = 1$, $a(l_{36}(x_{66})) = 1$, $a(l_{66}(x_{66})) = 0$ since Lagrange interpolation polynomials have been evaluated from Equations (10), (11), (12). Therefore, we can retrieve decryption key through decryption polynomial shown below:

$$F_{DK_6}(x_{66}) = x_6 6 + DK_6 - [x_{66} + 0] = DK_6. \tag{13}$$

For the case that the server S6 uses a input key such as $x_* \neq x_{66}$ or $x_* \neq$ any linear combination of key x_{11}, x_{22}, x_{33}, $\cdots$ and so on, then we have:

$$\begin{aligned} l_{16}(x_*) &= \left(\frac{x_* - x_{26}}{x_{16} - x_{26}}\right)\left(\frac{x_* - x_{36}}{x_{16} - x_{36}}\right)\left(\frac{x_* - x_{46}}{x_{16} - x_{46}}\right) \\ &\quad \left(\frac{x_* - x_{56}}{x_{16} - x_{56}}\right)\left(\frac{x_* - x_{66}}{x_{16} - x_{66}}\right) = c_1, c_1 \neq 0, 1 \end{aligned} \tag{14}$$

$$\begin{aligned} l_{36}(x_*) &= \left(\frac{x_* - x_{16}}{x_{36} - x_{16}}\right)\left(\frac{x_* - x_{26}}{x_{36} - x_{36}}\right)\left(\frac{x_* - x_{46}}{x_{36} - x_{46}}\right) \\ &\quad \left(\frac{x_* - x_{56}}{x_{36} - x_{56}}\right)\left(\frac{x_* - x_{66}}{x_{36} - x_{66}}\right) = c_2, c_2 \neq 0, 1 \end{aligned} \tag{15}$$

$$\begin{aligned} l_{66}(x_*) &= \left(\frac{x_* - x_{16}}{x_{66} - x_{16}}\right)\left(\frac{x_* - x_{26}}{x_{66} - x_{26}}\right)\left(\frac{x_* - x_{36}}{x_{66} - x_{36}}\right) \\ &\quad \left(\frac{x_* - x_{46}}{x_{66} - x_{46}}\right)\left(\frac{x_* - x_{56}}{x_{66} - x_{56}}\right) = c_3, c_3 \neq 0, 1. \end{aligned} \tag{16}$$

Then we get the decryption equation shown below:

$$\begin{aligned} F_{DK_6}(x_*) &= x_* + DK_6 - [x_*(c_1 + c_2 + c_3 + R)] \\ &\neq DK_6. \end{aligned} \quad (17)$$

Which means the server fails to get the decryption key, that is, the confidential documents have been protected from illegal attempts.

Other failure case could be the one that we suppose the server S_6 uses input key x_{56}, then we calculate Lagrange interpolation first , the results shown below:

$$\begin{aligned} l_{16}(x_{56}) &= (\frac{x_{56}-x_{26}}{x_{16}-x_{26}})(\frac{x_{56}-x_{36}}{x_{16}-x_{36}})(\frac{x_{56}-x_{46}}{x_{16}-x_{46}}) \\ &\quad (\frac{x_{56}-x_{56}}{x_{16}-x_{56}})(\frac{x_{56}-x_{66}}{x_{16}-x_{66}}) = 0 \end{aligned} \quad (18)$$

$$\begin{aligned} l_{36}(x_{56}) &= (\frac{x_{56}-x_{16}}{x_{36}-x_{16}})(\frac{x_{56}-x_{26}}{x_{36}-x_{36}})(\frac{x_{56}-x_{46}}{x_{36}-x_{46}}) \\ &\quad (\frac{x_{56}-x_{56}}{x_{36}-x_{56}})(\frac{x_{56}-x_{66}}{x_{36}-x_{66}}) = 0 \end{aligned} \quad (19)$$

$$\begin{aligned} l_{66}(x_{56}) &= (\frac{x_{56}-x_{16}}{x_{66}-x_{16}})(\frac{x_{56}-x_{26}}{x_{66}-x_{26}})(\frac{x_{56}-x_{36}}{x_{66}-x_{36}}) \\ &\quad (\frac{x_{56}-x_{46}}{x_{66}-x_{46}})(\frac{x_{56}-x_{56}}{x_{66}-x_{56}}) = 0. \end{aligned} \quad (20)$$

Hence, we have decryption polynomial equation calculated below:

$$F_{DK_6}(x_{56}) = x_{56} + DK_6 - (x_{56} \times 0 + R) \neq DK_6. \quad (21)$$

Eventually, we know that the server uses the wrong input key x_{56}, which causes the decryption polynomial $F_{DK_j}(x)$ to generate the false decryption key that can't open the confidential documents at all.

4 Analysis of Security

We would discuss from the viewpoint of attackers to compromise the proposed scheme to confirm the method is secure. The attackerssteals from outside.They hacks valuable information in order to accumulate money. This situation could result in the divulgence of confidential information and damages. Accordingly, this becomes a serious issue in the process of security analyses. Regarding external attack, attackers with the knowledge of public parameters are not able to obtain any decryption key DK_t and, consequently, they are not able to obtain any confidential file. If the external attackers wish to extract the secret key SK_i from the interpolation function parameters $x_{ij} = ID_j||g^{SK_i}(\mathrm{mod}p)$, then they have to solve the Discrete logarithm problem; which is known to be computationally infeasible since p is a large prime.

The reversed attack is defined as the user with lower authority intends to access the higher level. If a user successfully carries out a reversed attack on a user who has higher authority, then the attacker could illegally obtain the secret key to access to the confidential documents. After hacking the information successfully, the attackers may tend to sell it which would result in loss to the organization. It is thus important to prevent reversed attack.

5 Conclusions

The biggest challenge for VR/AR development is the security of privacy user data. This paper introduces Virtual Integrated VR/AR-information Systems. This concept is used to achieve dependence and provide a safe environment for enterprise institutes to exchange information online based on user's right management mechanism to access the confidential documents, that manager can efficiently achieve the user's complete VR/AR information. However, there is a risk of data transfer. Theft or tampering of data on the Internet so that the user or server permission to add a hierarchical access control, to ensure that patients in the premise of data confidentiality and safety, effective and safe for authorized manager to use, not has been authorized users, to ensure user privacy will not be violated.

Acknowledgments

This study was supported by the introduction of talents Southeast University Scientific Research Projects (Nos. 3213048201).

References

[1] E. Bierman, T. Pretoria and E. Cloete, "Classification of malicious host threatsin mobile agent computing," in *Proceedings of the Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists on Enablement Through Technology*, no. 5, pp. 34–49, 2002.

[2] A. Castiglione, *et al.*, "Hierarchical and shared access control," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 4, pp. 850–865, 2016.

[3] A. Castiglione, A. D. Santis, B. Masucci, F. Palmieri, A. Castiglione and X. Y. Huang, "Cryptographic hierarchical access control for dynamic structures," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 10, pp. 2349–2364, Oct. 2016.

[4] W. Y. Chao, C. Y. Tsai and M. S. Hwang, "An improved key-management scheme for hierarchical access control," *International Journal of Network Security*, vol. 19, no. 4, pp. 639–643, July 2017.

[5] V. El-khoury, N. Bennani and A. M. Ouksel, "Distributed key management in dynamic outsourced databases: A trie-based approach," in *First International Conference on Advances in Databases, Knowledge, and Data Applications*, Gosier, pp. 56–61, 2009.

[6] J. Kasurinen, "Usability issues of virtual reality learning simulator in healthcare and cybersecurity," *Procedia Computer Science*, vol. 119, pp. 341–349, 2017.

[7] K. C. Laudon and J. P. Laudon, *Management Information Systems*, Pearson, Chapter 6: Information systems Organizations and Strategy, pp.143, 2011.

[8] H. Y. Lin, D. J. Pan, X. X. Zhao and Z. R. Qiu, "A rapid and efficient pre-deployment key scheme for secure data transmissions in sensor networks using lagrange interpolation polynomial," in *International Conference on Information Security and Assurance (ISA'08)*, pp. 261–265, 2008.

[9] I. C. Lin and T. C. Liao, "A survey of blockchain security issues and challenges," *International Journal of Network Security*, vol. 19, no. 5, pp. 653–659, Sept. 2017.

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

[11] J. W. Lo, M. S. Hwang, C. H. Liu, "A simple key assignment for access control based on polynomial," *The Arabian Journal for Science and Engineering*, vol. 38, no. 6, pp. 1397–1403, June 2013.

[12] Z. M. Ozsoyoglu and J. Wang, "A keying method for a nested relational database management system," in *Eighth International Conference on Data Engineering*, Tempe, AZ, pp. 438–446, 1992.

[13] T. H. Sun and M. S. Hwang, "A hierarchical data access and key management in cloud computing," *ICIC Express Letters*, vol. 6, no. 2, pp. 569–574, 2012.

[14] S. H Tang, X. Y. Li, X. Y. Huang, Y. Xiang and L. L. Xu, "Achieving simple, secure and efficient hierarchical access control in cloud computing," *IEEE Transactions on Computers*, vol. 65, no. 7, pp. 2325–2331, July 2016.

[15] X. C. Zhang and X. L. Zhao, "Based on virtual reality technology security intelligent wireless network resources dynamic allocation method research," *Science Technology and Engineering*, vol. 15, pp. 283–288, 2017.

Biography

Tsung-Chih Hsiao received the Ph.D. in the Department of Computer Science and Engineering, National Chung Hsing University, Taiwan. He is currently an associate professor in the School of Arts at Southeast University, China. Research fields include Information Security, Cryptography, and Network Security.

Yu-Min Huang received the Ph.D. in the Department of Statistics at the University of Minnesota Twin Cities, United States. She is currently an assistant professor in the Department of Statistics at the Tunghai University, Taiwan. Research fields include Statistical Inference, Multivariate Statistics, and Statistical Computing.

Yu-Fang Chung received a B.A. degree in English Language, Literature and Linguistics from Providence University in 1994, an M.S. degree from Dayeh University in 2003, and a Ph.D. degree from National Taiwan University in 2007, both in Computer Science, Taiwan. She is currently a professor in the Departments of Electronic Engineering and Information Management at Tunghai University, doing research, i.e., Information Security and Cryptography.

Tzer-Long Chen received the Ph.D. in the Department of Information Management, National Taiwan University, Taiwan. He is currently an assistant professor in the Department of Information Technology at Lingtung University, Taiwan. Research fields include Information Security, Cryptography, and Network Security.

Tzer-Shyong Chen received the Ph.D. in the Department of Electrical Engineering (Computer Science) at National Taiwan University, Taiwan. He is currently a professor in the Department of Information Management at Tunghai University, Taiwan. Research fields include Information Security, Cryptography, and Network Security.

An Algorithm of the Privacy Security Protection Based on Location Service in the Internet of Vehicles

Peng-Shou Xie[1,2], Tian-Xia Fu[2], and Hong-Jin Fan[2]
(Corresponding author: Tian-Xia Fu)
Research Center of Engineering and Technology for Manufacturing Informatization of Gansu Province[1]
School of Computer and Communication, Lanzhou University of Technology[2]
No. 287, Lan-Gong-Ping Road, Lanzhou 730050, China
(Email: fukaix_wang@163.com)
(Received Nov. 14, 2017; Revised and Accepted June 15, 2018; First Online Feb. 24, 2019)

Abstract

The safe and comfortable location based services are required for users in the internet of vehicles, at the same time, the privacy and confidential requirements are indispensable. Enciphering merely user′s information cannot guarantee the safety of user′s privacy, and the query itself may leak the user′s location information and identity one. Aimed at the problem of privacy security in location based service under the internet of vehicles environment, in this paper, through analysis of LBS privacy security technology, in the V2I system of internet of vehicles, combining K area with fake names anonymous technology, a kind of improved PPA-IOV privacy protection algorithm is formed. Experimental verification indicated that the algorithm performs a higher anonymous success rate for users in the car network environment and reduced the average anonymous space, thus the service quality of user′s query is improved.

Keywords: Anonymous Success Rate; Internet of Vehicles; Location Based Services; Privacy Protection

1 Introduction

With the popularization of cars and ownership increased significantly nowadays, the road conditions become more and more complex and the safety of road situation is not optimistic as well. When users in the internet of vehicles enjoy the great convenience provided by location based services (LBS), their own identity information is involved inevitably. Location coordinates and the contents queried of user privacy information are exposed to the network. Location based service is the main basis for the server to process the service request. The more accurate the location information provided by the user, the more accurate the service information returned to their after the server query processing. However, the query service request that contains the exact location information is recorded in the location server, which will undoubtedly open the door to malicious attackers to steal the user′s location privacy and query privacy. If the location server is not trusted or the communication process that between the user and the location server is unsafe, the user′s location information or query content may be stolen or disguised by the malicious attacker. The attacker deduces the user′s trajectories and status according to obtain the user′s location information, or from the user query content to infer the user to travel to the destination and so on, which will bring serious privacy threats to user [6].

In recent years, a variety of user's location privacy protection scheme has been proposed. Zhang *et al.* [16] presented a personalized LBS (P, L, K) anonymous model based on sensitivity, and on this basis, formed the privacy protection query anonymous algorithm under the use of mesh division and pseudonym users. By searching for the user′s neighborhood space iteratively, the purpose of protecting the privacy of the query was achieved. Li *et al.* [13] proposed a mobile-cloud framework, which is an active approach to eradicate the data over-collection. By putting all users′ data into a cloud, the security of users′ data can be greatly improved. Che *et al.* [4] put forwarded a location anonymity algorithm based on P2P and dynamic grids. The grid is used to divide the space provided by the LBS service provider, and dynamically adjust its size of the space according to the user′s required anonymity and the number of users.

Arain *et al.* [1] proposed Dynamic Pseudonym based multiple mix zone (DPMM) technique by analyzing limitations of existing methods related to location privacy with mix zones, such as RPCLP, EPCS and MODP, which ensures the highest level of accuracy and privacy, and addresses the issues related with existing location privacy protection techniques. Han *et al.* [7] analyzed the location K-anonymity technique, which requires that when

an user sends a LBS request, its location information is undistinguishable from other location information of at least $k-1$ users, which can effectively resist query tracking. Jin *et al.* [8] improved the positional K-anonymity algorithm based on a quad-tree-like scan using a bottom-up approach. Based on the Casper model, this method after scanning the information of the lowest-level grid, and choose to iterate upwards cells to improve spatial resolution if it satisfies the minimum anonymity requirements. Some progresses have been made in the solution to privacy protection, but there is still a problem that the contradiction between the effect of privacy protection and the accuracy of the processing results is difficult to balance. Importantly, there are relatively few researches on privacy protection methods specific to the environment of internet of vehicles.

Considering the shortcomings of the above researches, we propose a Privacy Preservation Algorithm-Internet of Vehicles (PPA-IOV) correspondingly. By the PPA-IOV method, privacy information can be well protected. Through the analysis of service structure of LBS and location privacy protection technology in internet of vehicles [12], the PPA-IOV is improved immensely by mean of combining K anonymous area and pseudonym technique in P2P network structure. The simulation experiments indicate that the PPA-IOV can get a better success rate and improves the accuracy of the service by reducing the anonymous area without exposing the exact location of the user comparing with the SCAPGID [5] algorithm and P2P-IS-CA-HL [4] algorithm.

The rest of the paper is organized as follows. The related location privacy protection technology of the algorithm and design of PPA-IOV in internet of vehicles are introduced in Section 2. The key Steps of the algorithm and the realization process in detail are introduced in Section 3. The detection performance of the proposed algorithm are analyzed and compared through two sets of simulation experiments in Section 4. Section 5 concludes the solutions.

2 Privacy Protection Algorithm - Internet of Vehicles (PPA-IOV)

2.1 Related Location Privacy Protection Technology

In LBS, the most widely used privacy protection model is the location K-anonymous model [15]. The basic idea is that the location of the mobile user is satisfied the location k-anonymity when a mobile user′s location cannot be distinguished from the location of other k-1 mobile users. There are three basic techniques for implementing location k-anonymity: Dummy location, spatial cloaking, and spatial-temporal cloaking [9].

1) Dummy location.

   Generally speaking, publishing location pseudonym information is putting the false location of service request instead of the real location. As shown in the Figure 1, the small rectangle represents the user object in the LBS. The black dot indicates the location information of the user. The hollow dot indicates the user's false location submitted to the location server. One of the important advantages of pseudonym technology is that users can generate dummies on their own without the need for any other communication protocol components. When using the pseudonym technology, the attacker does not know where the users real position is. Therefore, the farther the distance between true and false positions is, the higher the security factor is, but the worse the service quality is. It is possible to obtain the correct query result if the user object provides a false position for L1 instead of L2 in Figure 1, but it is also easier to expose the actual position of L1 than L2.

2) Spatial cloaking.

   The main idea of this approach is to replace the user's precise location information with a spatial hiding area that requires the location of the user and at least other k-1 users. For example, the real position of a user is u, and the idea of space clocking is to expand this point into a hidden range A, as shown in the dotted ellipse area in the Figure 2, use the hidden area A instead of the user′s location u, and to ensure that the probability of user occurrence on each location in the area is the same. So that the attacker can only know that the user is in the hidden area, but cannot tell the user's exact location. The size of the hidden area is proportional to the degree of privacy protection and inversely proportional to the quality of the service.

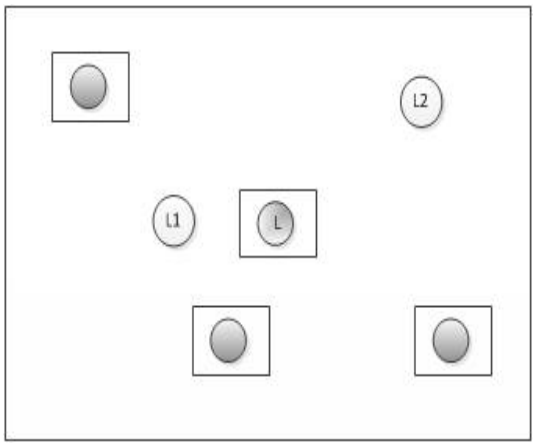


Figure 1: Diagram of the dummy location

3) Spatial-temporal cloaking.

   The basic idea of spatial-temporal cloaking is to delay the response time when constructing a hidden area of space, with a temporal and spatial hidden area

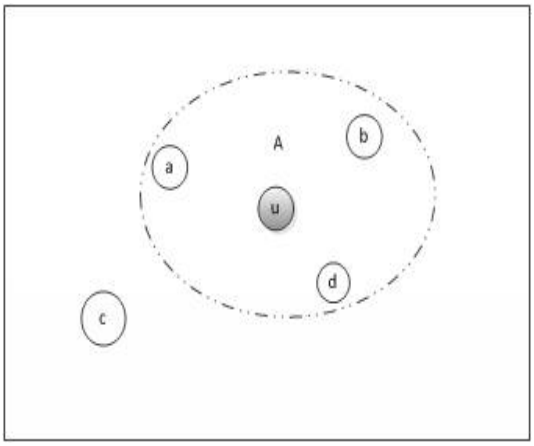


Figure 2: Diagram of the spatial cloaking

instead of the user′s precise location, which also contains the location of the user and at least other k-1 users. As a result of the delay in response time, so during this time, there will be more users and more queries. The greater the density of users, the smaller the hidden area is, and thereby improved the degree of privacy protection and the quality of service. However, apart from the shortcomings of spatial clocking, spatial-temporal cloaking technology also increases service response time.

2.2 PPA-IOV Design

In the V2I communication system of the internet of vehicles, vehicle nodes can communicate with the roadside unit (RSU) directly by using the P2P structure. The P2P structure [10] can not only use the knowledge of other nodes in the network and enrich the anonymous technology diversity when compared with the independent structure, but also eliminate the need for third-party anonymous server, to avoid the risk of information disclosure by attacked server when compared with the central structure. The RSU plays a certification and supervisory role in nodes within its coverage. In the environment of the internet of vehicles, the vehicle nodes are constantly changing and the formation of anonymous groups is transformed and updated in real time. The P2P structure relies on the collaboration of terminal users to achieve privacy protection. The nodes of the structure can communicate with each other through peer-to-peer network and establish a real-time assistance relationship. The security of a service model based on P2P structure depends on the selection of the anonymous areas. The privacy protection algorithm PPA-IOV proposed in this paper is described as follows:

1) The vehicle requests the identity authentication for RSU to obtain the node identity, and records it as V_i. The corresponding pseudonym node is generated by the pseudonym generation algorithm, denoted it as V_{pi}. The value of i is in the range of 1 to n, and n represents the number of vehicles passing through the RSU. Each vehicle node corresponds to only one corresponding vehicle pseudonym node.

2) The terminal nodes carry out peer-to-peer communication with the neighboring vehicle nodes (V_i or V_{pi}) through the ad-hoc network.

3) A certain node V_i within the signal coverage of an RSU actively forms an anonymous group with other nodes. A node can appear within multiple anonymous groups.

4) Set a fixed value K, when the number of nodes within the anonymous group reaches the K value, no new nodes will be added. Otherwise, then return to 3).

5) The K-anonymous region of the RSU is denoted as R_j, where j is in the range of 1 to m, and m represents the number of anonymous areas within an RSU, R_j including the K nodes in the anonymous region, R_j=$\{V_1, V_{p1}, V_2, V_{p3}, \ldots, V_K\}$.

6) Anonymous area R_j randomly selects a user as an agent to send the query to the RSU. The RSU receives a query from the entire anonymous area, and returns the reply to the entire anonymous area when the RSU responds the query result.

The model of the privacy protection algorithm is shown in Figure 3 below:

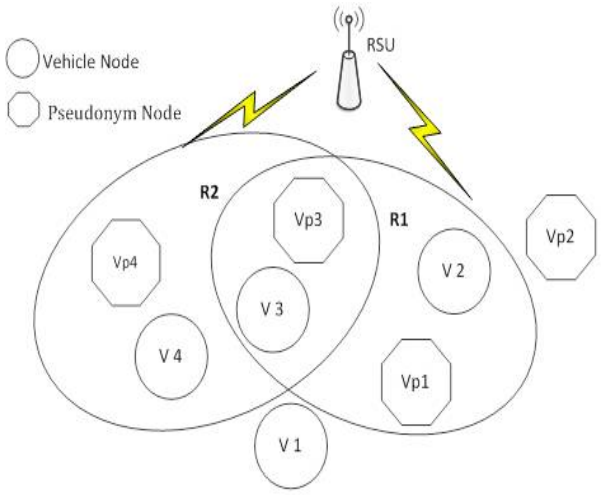


Figure 3: The model of the privacy protection algorithm

Take the Figure 3 for example, V_i (i=1, 2, 3, 4) represents the vehicle node, V_{pi} represents the corresponding pseudonym node, and R_i represents the K anonymous area within the RSU. From the figure, both node V_3 and V_{p3} are in the R_1 area, but also in the R_2 area. At the same time, V_2 in the R_1 area, but there can be no V_{p2}. In general, while the anonymous region R_j is formed, the elements in the R_j (j=1,2) region have both real vehicle nodes and fake position nodes, but the real nodes V_i

and the corresponding pseudonym nodes V_{pi} not must appear simultaneously in the R_j region. The algorithm ensures that the generated pseudonym nodes are at the same level as the real node. When the RSU returns the service response information, the node issuing the request information takes the initiative to obtain the response information, and the other nodes ignore it. Since the K-anonymous area sends out the request information as a whole, it is not easy for eavesdroppers to discern sends the service request accurate identification of the vehicle. At the same time, the use of the method can also effectively solve the difficult problem of the formation of anonymous group of vehicle nodes .

3 The Key Steps of the PPA-IOV Algorithm

3.1 Pseudonymous Generation Algorithm

In the privacy protection algorithm proposed in this paper, after obtaining the RSU authentication, the user generates a fake location node (dummy). And the two message nodes can send and receive the information normally and form an anonymous area together with the neighbor nodes, then send the message to the RSU. After receiving the service reply message sent back from the cloud server, the RSU broadcasts the response information to the anonymous area. The requesting user only needs to take the initiative to extract the necessary information. In this way, even if the RSU roadside unit stores a set of position data, it cannot distinguish the real position data from them.

As the road navigation service, the user must send the location data continuously in the LBS service of the Internet of Vehicles. Generally speaking, each object distance that can be moved is limited within a fixed time. If the dummy is randomly generated, the difference between the real location data and the fake position of the dummy node can be easily detected by the observer. In this case, the location anonymity is reduced. In order to avoid this situation, the dummy must not be completely different from the real location data. For this purpose, we first propose the following pseudonymous generation algorithm to double the pretender for real location data. First add the following assumptions:

a) All users generate the same number of virtual objects. Namely, each vehicle user generates only one virtual node.

b) In addition to location data, the user does not send other personal information.

c) The location data of the user in the process of generating the pseudonym node remains the same.

d) The user location information remains the same within the unit time. Over a period of time, it automatically re-authenticate and generates a pseudonym node.

The pseudo-code of the pseudonymous generation algorithm is shown in Algorithm 1. In this algorithm, the position of the dummy depends on the location coordinates of the real vehicle nodes at the last moment.

Algorithm 1 Pseudo-code of pseudonymous generation algorithm

Input: Positions of entities at $t-1$, m, n
Output: Positions of pseudonymous at t

```
Define a dummy structure {
    double x;  // x coordinate
    double y;  // y coordinate
    double t;  // time
}
Assignment entity[] to the Input
for i = 1; i < n; i++ do
  // rand(x,y): generate a random number between x and y
  dummie[i].x ← rand(entity[i].x-m, entity[i].x+m);
  dummie[i].y ← rand(entity[i].y-m, entity[i].y+m);
  dummie[i].t ← (entity[i].t)+1;
end for
Output the contents of the dummies[]
```

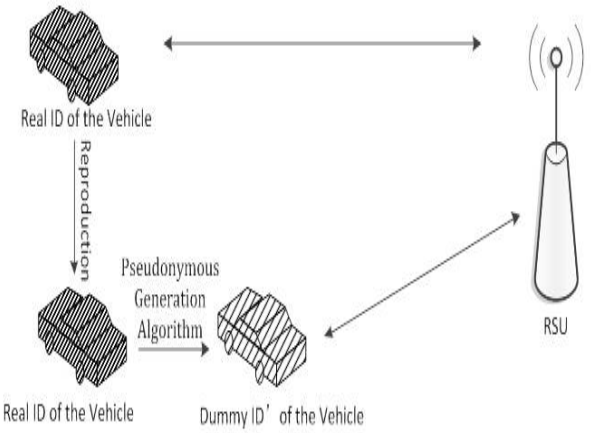


Figure 4: The model of the pseudonym generation algorithm

Set the basic information of the vehicle node: I=$\{ID$,$(x,\ y)$,t-1$\}$, where ID represents the node identification of the real vehicle, $(x,\ y)$ represents the position coordinates of the vehicle, and t represents the current time. After the pseudonym generation algorithm, the target user generates a user whose basic information is I'= $\{ID,$ random $(x,\ y),\ t\}$, where I and I' exist in the car network system at the same time, to realize the protection of double nodes of user's identities. The pseudonym generation algorithm model is shown in Figure 4 .

3.2 K-anonymous Area Generation Algorithm

As mentioned in Section 2.1 above, the most widely used privacy protection model in LBS is the location K-anonymity model. The K-anonymous location privacy protection technology that based on the generalized performances better in terms of accuracy and practicability, is a commonly used location privacy protection technology. In the process of K anonymity, Nearest Neighbor Clock [3] is the most classic anonymous area formation algorithm. As shown in Figure 5 (a), where user A performs an LBS query with an anonymous degree of 3, and the two nearest neighbors of user A are user B and user C. The anonymous area of user A is a rectangular area indicated by a dotted line in the figure. And the two nearest neighbors of user B are user C and user D, and the anonymous area of user B is the rectangular area indicated by the dotted line in the Figure 5 (b).

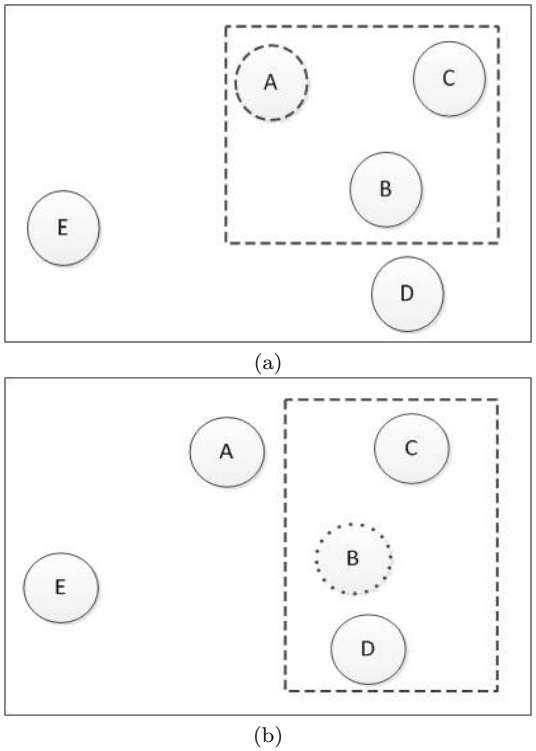


Figure 5: The nearest neighbors of anonymous area; (a) The nearest neighbors of anonymous area A; (b) The nearest neighbors of anonymous area B

The construction of anonymous areas is directly related to the accuracy of service search. If the anonymous area is too large, the service quality will be reduced. If the anonymous area is too small, the location of the user will be exposed. Associating the special trajectory of vehicle operation in the internet of vehicles, under normal circumstances, it only travels on a fixed road. The path of the vehicle is limited, and the area of the vehicle node is also limited.

The regular circular or rectangle anonymous area is not applicable under the environment of the internet of vehicles. Due to the fact that only one-way or two-way road systems are commonly used to drive vehicles on specific road sections, if the anonymous area is simply constructed in a circular shape, it will cause nearly half the invalid anonymous space in the anonymous area when the requesting node is located on the side of the road, thus causing the inaccuracy of the return of the query result. In addition, the vehicle road construction is not a regular pattern owing to the complex and changeable road terrain, but as the road condition changes constantly. So it is necessary to adopt an irregular anonymous area selection method to meet the needs of the anonymous area construction demand. For this reason, a K-anonymous area study based on boundary irregular polygons [14] is developed.

Algorithm 2 Pseudo-code of K-anonymous area generation algorithm

1: Input A_{min}, A_{max}, K
2: // **Step 1:** Peer search step
3: $List \leftarrow \{\varnothing\}$
4: U broadcast a request to the peers V_i
5: **for** $i = 1$; $i < k$; $i++$ **do**
6: check t_s of the v_i node
7: **if** $t_s \geq t - \Delta t$ **then**
8: $list \leftarrow list \sqcup$ {the received location information of node }
9: **else**
10: abandon the node
11: **end if**
12: **end for**
13: // **Step 2:** Cloaked Area step
14: $T \leftarrow$ the point that x-coordinates is the largest and smallest or the y-coordinates is the largest and smallest in List
15: $A \leftarrow$ a minimum bounding irregular polygons of all users in T
16: $S \leftarrow$ the acreage of the area A
17: **if** node is in the anonymous area A **then**
18: $List2 \leftarrow$ {the location information of node}
19: $List \leftarrow List$ - $List2$
20: **end if**
21: **while** $List = \{\varnothing\}$ **do**
22: Calculate the acreage S of A
23: **end while**
24: **if** $S < A_{min}$ **then**
25: recruit new nodes, and execute lines 6 to 23 to ensure that A satisfies the minimum area privacy requirement
26: **else**
27: **if** $S > A_{max}$ **then**
28: re-execute the Algorithm 2
29: **end if**
30: **end if**
31: **Return** A

When a user needs to query a specific service nearby,

the required anonymous requirements are first identified: the minimum area of anonymous area (A_{min}), the maximum anonymous area (A_{max}), and the anonymity degree (K). As the vehicle node is constantly moving, the longer the peer position information is the user cached, the lower the information timeliness of the nodes, the greater the offset of the position of the vehicle node, and the lower the accurate of the hidden area, so parameters Δt be used to control the obsolescence of the node caching information. And it also represents the life of the vehicle identity beacon. When $t_s \geq t - \Delta t$, the user node information is valid, where t_s is the point in time when the node caches location information and t represents the occurrence time of the query. Otherwise the requesting vehicles will automatically access to the RSU to obtain new identity authentication.

The generation of K anonymous areas in the PPA-IOV algorithm is described as follows.

Input: Set the service query to Q, Q ={ID, (x, y), v, M, t), where ID represents the node identify of the requesting vehicle, (x, y) represents the current position coordinates of the vehicle, and v represents the current velocity of the vehicle, M represents the content of query, and t represents the occurrence time of the event.

Output: Generate an irregular polygonal area that is not less than K nodes.

1) The user U first builds an anonymous area with the help of nearest neighbor anonymous algorithm to recruit neighbors. U checks the freshness of the nearest neighbor node cache information in turn. If the node within the last peer search time, that is $t_s \geq t - \Delta t$, the user U requests its neighbor to return its list and the size of the t_s. If the neighbor peer cache time information is too old, that is $t_s \leq t - \Delta t$, U will discard the node information. Through multiple operations, a result set $List$={(x_1, y_1), (x_2, y_2), $\cdots$, (x_k, y_k)}, with K node location coordinates that satisfies the K anonymity degrees is obtained.

2) Join all the points with the x-coordinates is the largest and smallest or the y-coordinates is the largest and smallest in the $List$ set into the set T. Such as T ={(x_{min}, y_a), (x_{max}, y_b) $(x_c, y_{min})(x_d, y_{max})$, $\cdots$}. The anonymous area made up of the points in the T set is recorded as A.

3) The x-coordinates is the largest and smallest or the y-coordinates is the largest and smallest of all points outside the A region are added to set T, and the anonymous region of the points in the T is recorded as a new anonymous region A. Repeatedly executed 3) until all nodes in $List$ are included in the A.

4) After completing the above steps, the regions formed in T are irregular polygonal regions with K nodes.

5) Calculate the acreage S of the polygon anonymous area A by dividing polygons into multiple triangles. Set the vertices in T to be arranged in counter clockwise order of $a_1, a_2, \cdots, a_m$. The coordinates of the vertices are in turn(x_1, y_1), (x_2, y_2), $\cdots$, (x_m, y_m), and the acreage of the anonymous area is:

$$S = \frac{1}{2}\Big(\begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} + \begin{vmatrix} x_2 & y_2 \\ x_3 & y_3 \end{vmatrix} + \cdots + \begin{vmatrix} x_{m-1} & y_{m-1} \\ x_m & y_m \end{vmatrix} + \begin{vmatrix} x_m & y_m \\ x_1 & y_1 \end{vmatrix} \Big). \quad (1)$$

6) If S<A_{min}, then add new neighbor nodes into the set $List$ and re-execute Step 2) to Step 5) until $S \geq A_{min}$. If $S > A_{max}$, then anonymous failure.

The pseudo-code of the K-anonymous area generation algorithm is shown in Algorithm 2.

After the K anonymous area is obtained, a user is randomly selected from the anonymous area as an agent for sending the query to the RSU. The RSU receives a query from the entire anonymous area and returns the reply to the entire anonymous area. Assume that the K anonymous area does not change during the time of the server returning the query result.

The K anonymous area generation algorithm makes the requesting user distribute evenly in the anonymous area, and realizes the probability that the requesting node is attacked at $1/K$. Compared with the circular anonymous area, the problem of the requesting user is located in the anonymous center, and the probability of attack is far greater than $1/K$ is well solved. In fact, in the case of the same number of nodes, the area of the irregular polygon based on the boundary is also smaller than the area of the anonymous area based on the circle or rectangle.

Figure 6 shows the idea of K anonymous area adjustment. The boundary point in $List$ is found to join the set T firstly (the boundary point is the maximum or minimum point of the x or y in the List set), in Figure (a), which constitutes a ring anonymous area. The anonymous area A contains A={1, 2, 3, 4, 5, 6, 13, 14, 15, 17, 20, 21}, $A_{outside}$={7, 8, 9, 11, 12, 16, 18, 19, 22}. The boundary point {8, 11, 7, 22} is found again outside the anonymous area A to join the set T in Figure (b), so the boundary set T and the anonymous region A are respectively: T={1, 11, 8, 6, 7, 17, 21, 22}, A={1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 20, 21, 22}, at this time only two nodes {18,19}outside the anonymous area A . Then node 19 is incorporated into the T set, and the resulting anonymous area A contains all the nodes in the $List$. So far, the irregular K- anonymous region is formed, as shown in Figure (c).

4 Algorithm Implementation

In order to verify the effect of the algorithm, the experiment is carried out on the platform of processor Intel (R)

Table 1: The experimental configuration parameters of this paper

module category	parameters	value
traffic scene	The number of lanes	6
	The length of Road /m	1500
	The width of Lane /m	3.5
	The speed of the node /(m/s)	10
	The number of user nodes	100,200,300,400,500
network communication	The acreage of minimum anonymous area (A_{min}/m^2)	12500
	The acreage of maximum anonymous area (A_{max}/m^2)	25000
	The time for caching record lose efficacy(Δt /s)	10
	The value of K	50,60,70,80,90,100

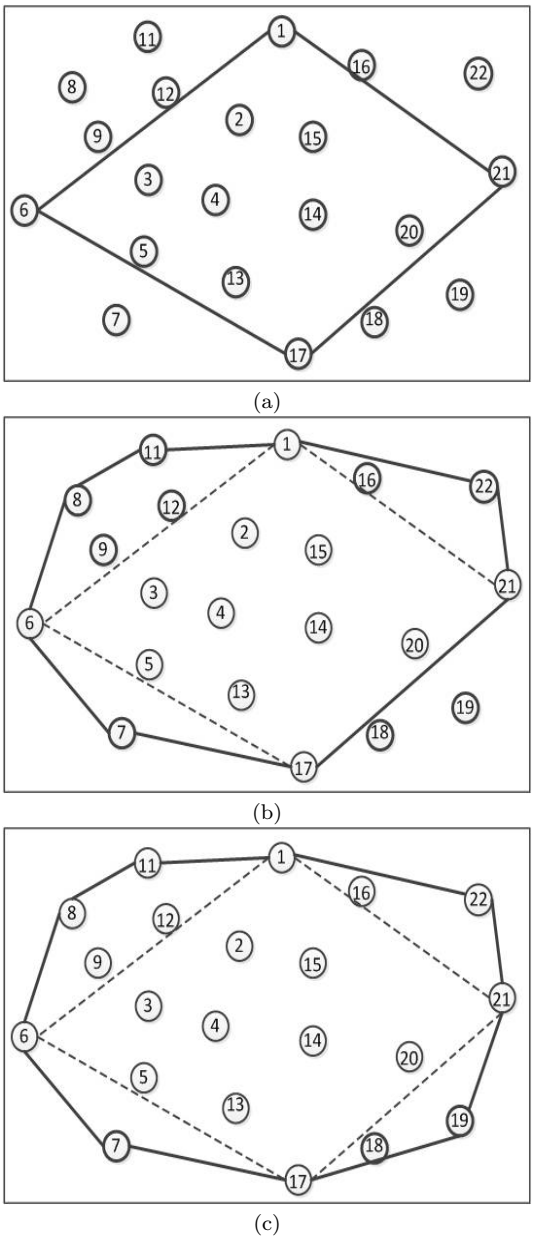


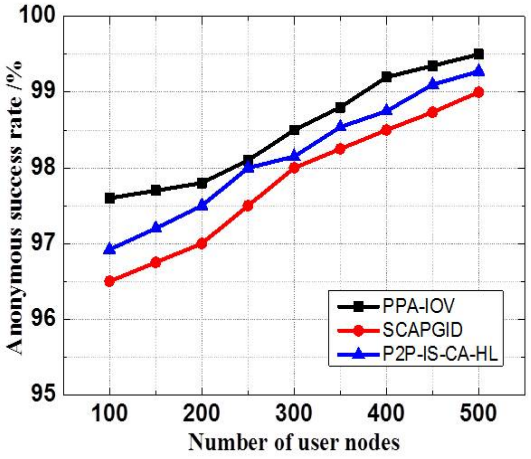
Figure 6: K anonymous regional adjustment

Core (TM) 2 Duo CPU E7500 3.0GHz with 4.0GB memory. First, the vehicle movement model is generated by simulation of urban mobility (SUMO) [2], because only the privacy security protection of nodes in the running process is discussed, so the setting of this paper experiment scenario is simple and only simulates a crossroad to a total of 6 straight lanes in two-way road. Supposing that a RSU signal is covered within 1500m∗1500m, and the number of nodes with different density is set up in the road to generate experimental scenes. Then the trace file generated in SUMO is connected to the network simulator NS-2 [11]. NS-2 reads the trace file to generate the vehicle node data. For better the result observation, the observation time was set at 200ms.

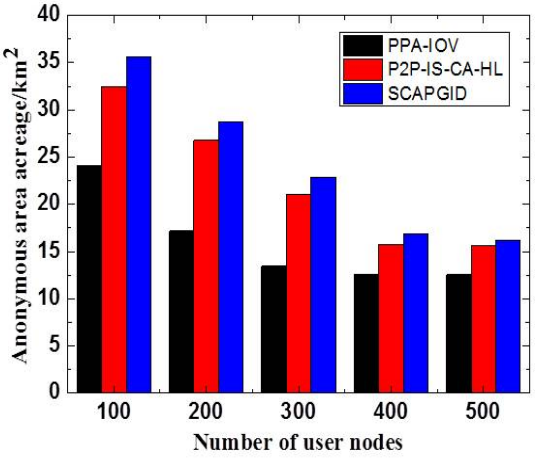
Algorithm SCAPGID divides the plane by the gird, and through the dynamic expansion of the grid and the user in the same grid to complete the location anonymous. In the P2P-IS-CA-HL Algorithm, the nodes in the anonymous area can share information and the anonymous area center can be adjusted. On each group of experiments, we were compared PPA-IOV with the SCAPGID and P2P-IS-CA-HL Algorithm varies in algorithm anonymous success rate and anonymous area size with the number of users changed within [100, 500] and the degree of anonymity changed within [50,100] respectively. The anonymous success rate indicates that the anonymity capacity of the privacy protection algorithm for the user's query request. While the smaller the anonymous area, the more accurate the quality of the query result obtained from the server.

The experimental process is divided into two parts: traffic scene and network communication. The table 1 shows the experimental configuration parameters.

Figure 7 shows the anonymous success rate and anonymous regions with the number of nodes changes respectively. It can be seen from the Figure (a) that the algorithm anonymous success rate of the three algorithms increases with the increasing of the number of nodes. Because the more nodes, the more assisted neighbor nodes in the anonymous area with the same large area, resulting in an increase in the anonymous success rate. The reason for the anonymous success rate in the PPA-IOV algorithm is relatively higher than that the SCAPGID and



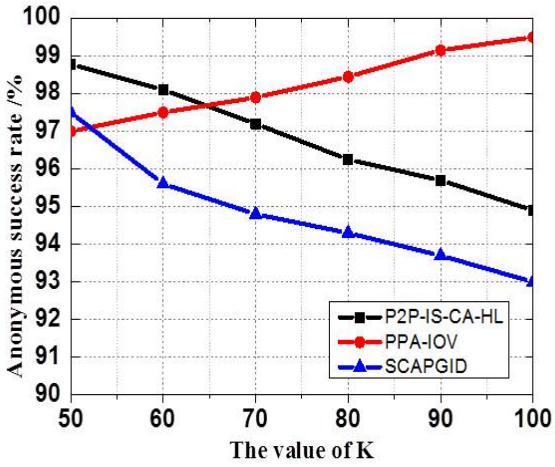
(a)



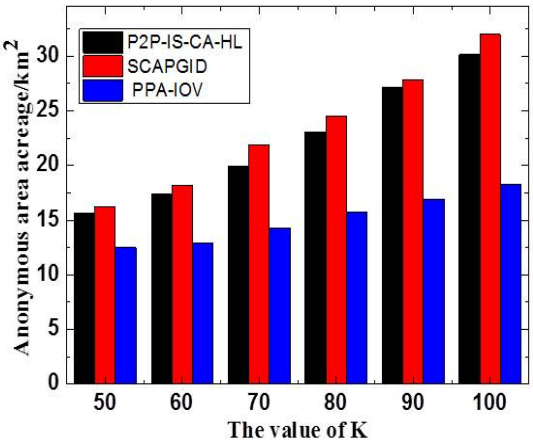
(b)

Figure 7: The anonymous success rate and anonymous area size varies with the number of user nodes. (a)The anonymous success rate varies with the number of user nodes. (b)The anonymous area size varies with the number of user nodes.

P2P-IS-CA-HL algorithm is that the pseudonym nodes in the PPA-IOV algorithm play a role and the number of nodes per unit area is larger. The Figure (b) shows that the anonymous area acreage of the three algorithms decreases as the number of users increases. In the case of fewer users, the algorithm needs to meet the degree of anonymity by enlarging the anonymous area. The SCAPGID and P2P-IS-CA-HL algorithm restrict the expansion way of the anonymous area by the idea of grid amplification and information sharing respectively, then causes the anonymous area to be far larger than the way that the PPA-IOV algorithm only expands the anonymous area through its own nodes. When the number of users reaches a certain degree and the nodes in the smallest anonymous area have already satisfied the degree of anonymity, then the anonymity area gradually stabilizes in the smallest anonymous area. Similarly, the pseudonymous node in the PPA-IOV algorithm makes the number of nodes in the entire environment higher than the other two algorithms, so it is easier to meet the anonymity in the case of a small number of users.



(a)



(b)

Figure 8: The anonymous success rate and the anonymous acreage under different k anonymity changes; (a)The anonymous success rate under different k anonymity changes; (b)The anonymous acreage under different k anonymity changes

Figure 8 depicts the variation of the anonymous success rate and the area of the anonymous area under different K anonymity levels respectively. The number of nodes at this time is set to 200. The graph displayed that the PPA-IOV algorithm increases the anonymous success rate as the increase of K anonymous, while the SCAPGID and P2P-IS-CA-HL algorithm showed an upward trend. This is because when the degree of anonymity increases, users need to recruit more neighbors help to collect enough peer

position information to fuzz its location, and these two algorithms are more likely to encounter network partitioning problem due to their own region amplification manner. Network partitioning problem is the number of users residing in the network partition is less than the required anonymous level K. Comparing to the SCAPGID and P2P-IS-CA-HL algorithm, the advantage of the PPA-IOV algorithm is the generation of pseudonym nodes and becomes the effective nodes in the anonymity area, which making the number of nodes in the whole experimental environment higher than them. Therefore, the area of anonymous acreage increases is very small with the K anonymity increasingly. While the SCAPGID and P2P-IS-CA-HL algorithm needs to accumulate the gird and share area of the anonymous regions to expand the anonymous area so that the number of nodes within it satisfies with K anonymity, which requires a larger area of anonymous area.

Furthermore, the above experiments indicated that the size of anonymous areas is not only related to the degree of anonymity K but also to the size of the smallest anonymous area. When the value of K is small, the decisive factor is the smallest anonymous region for the anonymous area. At this point, the higher node density in the region, the more nodes contained, and the greater the anonymous success rate. When the K value is larger, the nodes in the smallest anonymous area do not satisfy k anonymity, then anonymous areas spread out until the K nodes to meet, so the anonymous area formed in the environment of small node density is larger, and the success rate of anonymity increases.

5 Conclusions

This article around the theme of the terminal vehicle node privacy protection under the V2I system in the internet of vehicles, briefly describes three major privacy protection technologies at present. Then the privacy protection algorithm PPA-IOV by combining the characteristics of P2P structure is put forward. The content of the formation of pseudonyms and K anonymous areas in the algorithm are mainly introduced. Two groups of experiments via changing the number of user nodes and K anonymous respectively are performed, and the PPA-IOV algorithm is compared with the original algorithm SCAPGID and P2P-IS-CA-HL algorithm. The results showed that the algorithm proposed in this paper increases node density in the environment due to the application of pseudonym nodes, which improved the algorithm anonymous success rate. Furthermore the algorithm for getting the smaller anonymous space area and improve the quality of the query service.

Acknowledgments

This study was supported by the National Science and Technology Support Program of China (2012BAF12B19). The authors gratefully acknowledge the anonymous reviewers for their helpful comments and suggestions.

References

[1] Q. A. Arain, Z. L. Deng, I. Memon, A. Zubedi, J. C. Jiao, A. Ashraf, and M. S. Khan, "Privacy protection with dynamic pseudonym-based multiple mix-zones over road networks," *China Communications*, vol. 14, no. 4, pp. 89–100, 2017.

[2] M. Behrisch, D. Krajzewicz, and M. Weber, *Simulation of Urban Mobility*, 2014. (`http://sumo.dlr.de/2014/SUMO2014.pdf`)

[3] S. Bhattacharyya and G. Sanyal, "Feature based audio steganalysis (FAS)," *International Journal of Computer Network & Information Security*, vol. 4, no. 11, pp. 62–73, 2012.

[4] H. R. Che, Y. Z. He, and J. Q. Liu, "Spatial cloaking algorithm based on peer-to-peer and grid id," *Netinfo Security*, vol. 2015, no. 3, pp. 28–32, 2015.

[5] C. Y. Chow, M. F. Mokbel, and X. Liu, "Spatial cloaking for anonymous location-based services in mobile peer-to-peer environments," *GeoInformatica*, vol. 15, no. 2, pp. 351–380, 2011.

[6] P. Gormley, M. Mcbride, and J. Harkin, "Internal location based services using wireless sensor networks and rfid technology," *International Journal of Computer Science & Network Security*, vol. 6, no. 4, pp. 108–113, 2006.

[7] J. M. Han, Y. Lin, J. Yu, J. Jia, and L. Q. Zheng, "Lbs privacy preservation method based on location k-anonymity," *Journal of Chinese Computer Systems*, vol. 35, no. 9, pp. 2088–2093, 2014.

[8] F. S. Jin, Z. S. Ye, and H. Song, "A similar quadtree based on location k-anonymity algorithm," *Transactions of Beijing Institute of Technology*, vol. 34, no. 1, pp. 68–71, 2014.

[9] J. Karjee and H. S. Jamadagni, "Data accuracy models under spatio-temporal correlation with adaptive strategies in wireless sensor networks," *International Journal of Network Security*, vol. 4, no. 1, pp. 2152–5064, 2013.

[10] E. Kim, J. Kim, and C.Lee, "Efficient neighbor selection through connection switching for p2p live streaming," *Journal of Ambient Intelligence & Humanized Computing*, vol. 2018, no. 1, pp. 1–11, 2018.

[11] N. Kumar, M. Kumar, and R. B. Patel, "A secure and energy efficient data dissemination protocol for wireless sensor networks," *International Journal of Network Security*, vol. 15, no. 6, pp. 490–500, 2013.

[12] C. T. Li, M. S. Hwang, and Y. P. Chu, "A secure and efficient communication scheme with authenticated key establishment and privacy preserving for vehicular ad hoc networks," *Computer Communications*, vol. 31, no. 12, pp. 2803–2814, 2008.

[13] Y. Li, W. Dai, Z. Ming, and M. Qiu, "Privacy protection for preventing data over-collection in smart

city," *IEEE Transactions on Computers*, vol. 65, no. 5, pp. 1339–1350, 2016.

[14] G. A. D. Luna, P. Flocchini, N. Santoro, G. Viglietta, and M. Yamashita, "Meeting in a polygon by anonymous oblivious robots," *31st International Symposium on Distributed Computing (DISC'17)*, vol. 91, no. 14, pp. 1–15, 2017.

[15] Y. J. Wu, *Privacy Preserving Data Publishing : Models and Algorithms*. China: Tsinghua University press, 2015.

[16] F. X. Zhang and C. H. Jiang, "Research on lbs (p,l,k) model and its anonymous algorithms," *Netinfo Security*, vol. 2015, no. 11, pp. 66–70, 2015.

Biography

Peng-shou Xie He was born in Jan. 1972. He is a professor and a supervisor of master student at Lanzhou University of Technology. His major research fields include theory and technology of Internet of Things, technology and application of Internet of Manufacturing Things, theory and method of information system engineering, software theory and methodology.

Tian-xia Fu She was born in Aug. 1992. She is a master student at Lanzhou University of Technology. Her major research fields include security for privacy preservation in internet of vehicles.

Hong-jin Fan He was born in Mar. 1993. He is a master student at Lanzhou University of Technology. His major research fields include big data security in intelligent transportation.

Managing Computer Security, Risk Analysis and Threat Using ISO 31000:2009: Case Study at Seiyun Community College, Yemen

Abdullah A. Al-khatib[1] and Mohammed A. Hassan[2]
(Corresponding author: Abdullah A. Al-khatib)
Faculty of Computer Science, Landshut University of Applied Science[1]
Am Lurzenhof 1, 84036, Landshut, Germany
Department of Information Systems, Seiyun Community College[2]
Seiyun, Yemen
(Email: S-aalkha@haw-Landshut.de)
(Received Jan. 13, 2018; Revised and Accepted June 18, 2018; First Online Mar. 17, 2019)

Abstract

In modern digital era, all organizations are completely dependent on Computers and related devices. For that reason, it is mandatory for the organization to manage computer security in order to run smoothly. This paper discusses a case study of Seiyun Community College (SYNCC) for managing computer security, analyzing the risk and threat to the organization's computer system. This case study is performed using the data collection method from the archives. The findings show that SYNCC is lacking in computer system security behind ISO 31000:2009 standard which leads to failure in running the organisation. Finally, from the analysis, it is recommended that SYNCC develop a security plan to cover all the aspects of the information and communication technology. Educating the users is also needed in order to implement the security policy.

Keywords: Analyzing Risk; Analyzing Threat; Enterprise Risk; ISO 31000:2009; Risk management

1 Introduction

Higher educational Institutes are currently faced with a great demand for human, technological, and environmental resources, among others, that must be harmoniously related so that they can provide and optimise the different services. These services include communication over long distances, thereby reducing geographical limits, the storage of vital and large amounts of user information, the provision of support to users and a number of functions that can be provided with the interrelation of these resources. Also the growth of educational Institutes has increased the risks that are implicitly due to changes or updates that are generated by this growth. For this reason, those Institutes must be prepared to face these risks and to generate a support guide for Institutes to use.

An increasing number of higher educational Institutes are beginning to worry about this issue and they are asking how this can be implemented it and what costs can be incurred. Companies that are now entering the world of business are the ones that are most exposed since their ignorance of the subject leads to greater vulnerability. Before starting to function in an orderly manner, companies may have absorbed losses generated by some risk not taken into account or not controlled, so the guide will be based on the standard ISO 31000 that is fully defined for Risk management [12].

Seiyun Community College (SYNCC) is a governmental College in Yemen that is established in 2003. Presently, the College has five departments that are located in geographically separated buildings. Communication between these buildings is based entirely on Internet Protocol (IP) network. National and International communications, on the other hand, are heavily dependent on Internet services. The Information and Communication Technology (ICT) center provides services to the College. Prominent among the services it provides access to the Internet, Email services, staff and student information systems, and ICT Capacity Building Training (CBT). ICT center also operates wireless and fibre networks for intra campus connectivity. ICT vision of the College is to have Information Technology equipment for the entire community of the SYNCC and electronic classrooms and other equipment for educational delivery. The ICT mission of the College is to provide access to information and knowledge to the SYNCC community and beyond as a way to improve Teaching, Learning, Research and Community Services for its stakeholders.

There are a number of risk management standards and frameworks have been developed by countries and na-

tional standards bodies. Some of such standards and frameworks are: CoCo (Criteria of Control) standard which was developed by the Canadian Institute of Chartered Accountants in 1995, IRM (Institute of Risk Management) is one of the best-established and most widely used risk management standards was produced by the IRM in 2002 in co-operation with AIRMIC and Alarm, BS 31100 is published by British Standards Institution in 2008, COSO ERM is framework produced by the Committee of Sponsoring Organizations of the Treadway Committee in 2004. COSO ERM framework (2004) contains all of the elements of the earlier Internal Control version COSO (1992), Turnbull is framework produced by the UK's Financial Reporting Council in 2005. The standard that had the widest recognition was the Australian Standard AS 4360 (2004). AS 4360 was withdrawn in 2009 in favour of ISO 31000. The latest addition to the available standards is the international standard ISO 31000:2009 which was published in 2009. Although some standards are better recognized than others, organizations should select the approach that is most relevant to their particular circumstances. [7].

ISO 31000:2009 [10] is an international standard published in 2009 that provides a generic approach to risk management, establishes minimum principles required for effective management, proposing an integrated framework from the definition of strategies, planning, administration, information and communication, policies, values and culture. In this way, risk management is systematic, transparent and credible in any scope and context. Although it is a standard, it is not the objective of the standard to establish a uniform implementation scheme in organizations, their design and implementation will vary according to the particularities (objective, context, structure, operations, functions, processes, products, services, assets, *etc.*) of each entity.

ISO 31000:2009 is used as a methodology to harmonize the risk management process, and in the current or future models that the organization has, it is necessary to consider that this standard is not used as a basis for certification. Its implementation is established with a systematic application of policies, procedures and practices to the activities of communicating, consulting, establishing the context, identifying, analyzing, evaluating, treating, monitoring and reviewing the risks [5]. The success of the model depends on the adequate structuring of the framework providing the foundations and criteria that will be embedded in the different levels of the organization. The framework allows management of risks to be aligned with a process adjusted to the context of the organization, and allows the assurance of timely communication that is used for decision-making and the allocation of responsibilities. Figure 1 shows the interrelation of the different components of the framework in an iterative way. It should be noted that this framework does not prescribe the management system of an organization but facilitates the integration of risk management with the entire administrative system, emphasizing that the process must be initiated

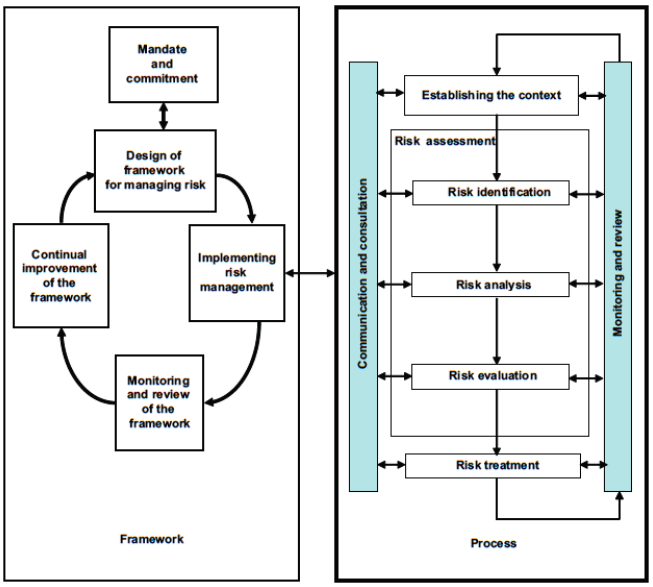


Figure 1: ISO 31000:2009 risk management framework and process [10]

in the top management of the company, showing its commitment and issuing Guidelines for risk management [8]. Once the principles are known, its framework to be followed by the organization in this risk management model shows the third pillar of the "Risk Management Process", which is shown in greater detail in ISO 31010: 2009 [9].

This study relies heavily on the literature and research in the context of risk management. Since ISO 31000 was initially published in November 2009 only a few academic research articles about standards have been published. Purdy [21], Shortreed [23], Leitch [15] and Aven [1] were among the first to examine the new research framework. While Purdy and Shortreed examined different aspects of the new standard in a positive tone, Leitch and Aven were very critical of the terminological and functional defects of the new standard. Health emergency management in mass gatherings and the applicability of ISO 31000 was examined by Hills [17] and examples of real-life mass gatherings with the Asia Pacific region were examined in the context of ISO 31000 which is of little value in the present study.

With the exception of study publications by professional organizations, such as PwC or Aon, risk management has been the focus of very little academic research thus far. Review of risk management related articles in academic journals and working papers has indicated that academic research on risk management is largely descriptive. In addition, findings of previous studies have been inconsistent. In [11] risk management consultancies and professional organizations have developed risk maturity models to investigate the performance of risk management. Models generally include a series of performance criteria which intend to measure how well the audited organization is performing in its risk management.

The contribution of this case study is that no scien-

tific studies on ISO 31000:2009 have been yet published in Yemen. Therefore, this present study is the first one to venture into that area. In addition, educational Institutes save time and find a clear and precise document and students will be able to obtain information more easily on the norms. This does not suggest that further research be discontinued, since as mentioned above, technology is an area that remains constantly changing and not being updated is considered as risk.

The remainder of this paper is organized as follows. Section 2 presents the Analysis and Discussion. The Recommendations and Critical Reviews are given in Section 3. Finally, Section 4 draws the Conclusions.

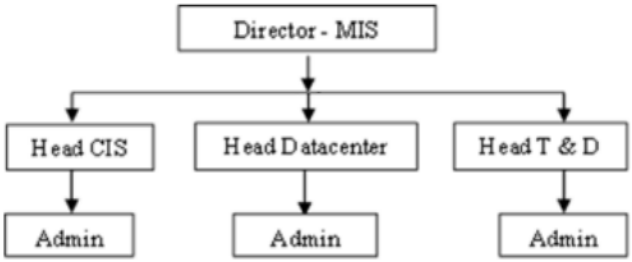


Figure 2: Security structure in SYNCC

2 Analysis and Discussion

For the development of the risk analysis within SYNCC, the standardization (ISO) 31000:2009 methodology was selected because it requires the participation of people who are directly involved in the operation of the critical information assets. Standardization is done in order to minimize risks and provide a higher quality service to its students by fully protecting its largest asset, which is the information on the aspects of reliability, integrity and availability, as each client makes strategic decisions [25].

The organizational, technological, and analysis aspects of an information security risk evaluation are complemented by ISO31000: 2009 risk management process. ISO 31000: 2009 is organized around these steps (illustrated in Figure 1), enabling organizational personnel to assemble a comprehensive picture of the information security needs.

SYNCC College's risk management process is consistent with the activities identified within the framework of IS0 31000:2009 method. To establish risk management process a reference was made to ISO 31000:2009. The process described below is based on ISO 31000:2009.

2.1 Communication and Consultation

The objectives at this step are to develop an effective communication process that serves as a basis for decision-making and for implementing the action plans required [10] in addition to identifying the College's Information and Communication Technology (ICT) assets such as information, hardware, human resources and the provision of policies and procedures to protect them.

2.1.1 Security Organization

This structured management framework directly monitors and controls the implementation of information security in SYNCC as shown in Figure 2. In this organization, the Director of Management Information Systems (Director-MIS) is the Chief information officer, followed by head of various units such as head of Datacenter, head of Corporate Information System (Head CIS) and head of Training and Development (Head T&D). Each individual unit has a number of staff called 'admin' that handles information system responsibility.

As can be seen from the diagram of Security Structure in SYNCC (Figure 2) each unit and admin handles some information system equipment and data that possesses security challenges. The College management is looking into the Director-MIS for the College information systems while head of units and admin are responsible for their individual units and services respectively. Reporting and implementation of security decisions starts from the admin to Director-MIS.

2.1.2 Human Resources

The SYNCC College has no centralized and well-documented security policy regarding employees, students, contractors, and part-time staff, but individual units have their own (Library and students affairs) registration/clearance policy where one will be asked to tender College belongings in his position before he is finally given clearance. In addition, the College has asked its new staff to sign or write an acceptance letter based on what has been stated in the appointment letter. The same thing applies to contractors, but here the College asked them to sign an agreement form that describes how the project will be executed, as well as penalties for failures. SYNCC does not conduct any security awareness training for its staff during and after recruitment. In addition, the College has no security policy for staff or contractors exiting SYNCC College to tender any access rights or equipment besides the one mentioned above.

2.1.3 Assets

Asset in information systems refers to anything from data, device, to other components of the environment that support information-related activities. In general, assets refer to hardware (*e.g.* servers, Desktop, laptops and switches), software (*e.g.* mission critical applications and support systems) and confidential information [19]. SYNCC have many assets (Network or systems) requiring protection as the following:

- *Desktops*: 208 PC (170 PC in student access rooms, 10 in the library, and 28 shared among staff in various

Table 1: Internal and external factors in SYNCC

Internal Factors	Description	External Factors	Description
Incident platform	Software designed for incident management.	Suppliers	Third parties that provide support services in the process and that allow the fulfilment of the objectives of the same.
Staff employees	They work in the help desk area in their different positions.	Customers	Are the people requesting the service.
Resources	Computer and communication equipment.	Economy	Influences the process due to the demand of potential customers and capital for investment in research and acquisition of new technologies.
Procedures	Refers to the procedures established in SYNCC area and which are endorsed by the quality area.	New technology	Refers to the hardware and software update.

offices).

- *Laptops*: 23 (Dean, Deputy Deans, Head of Departments and Principal Officers).
- *Printers*: 34 (Dean, Head of Departments and Principal Officers).
- *Servers*: 5 (4 running Ubuntu Linux and 1 running Centos hosting services such as web applications, Drupal, Internet connection, e-mail, and student record databases).
- *Switches*: 36 (48 port Cisco & Juniper switches to provide connection on the campus).
- *Router*: 6 (50% from 3800 Cisco Router and 50% from 7200 Cisco router).
- *Firewall*: 2 (2 SRX650 Juniper firewall).
- *Power systems*: 128 KVA diesel generator and 250KVA inverter.

The servers within the Datacenter are linked using Cat6 cable, various buildings of the College are linked using fiber optic cable and the computers in access rooms are linked using Cat6 cables. The remainder are linked by an 802.11g wireless network with an access port.

Although SYNCC has firewall configured at border of the Internet to protect the publicly accessible servers, the security control mechanisms implemented at SYNCC are not sufficient. Below are some of the security vulnerabilities of SYNCC:

- *Virus protection*: Not present on any computer except the one that runs either free version or pirated, not up-to-date; generally, most users are aware of viruses but are a bit unsure about what they could do to prevent them.
- *Spam-filtering software*: Many users are complaining about spam, but no protection is in place.
- *Updates*: All Microsoft Windows systems including productivity tools are not up-to-date because they are pirated copies.
- *Laptop computers*: None of the laptop computers have security locks, stickers and engraving.
- *Wireless networking*: The wireless network is open to people who have wireless access capability to browse on the network.
- *Web browsing*: SYNCC does not have a policy on acceptable use, and no one is taking any security measures.
- *Backups*: SYNCC back up data on the server to an external USB hard disk drive; the backed-up data and the machines are located in same room which is unsatisfactory.

Besides the above-mentioned assets, SYNCC also have the following:

- Records of contracts with suppliers and vendors.
- E-mail database and archive of past e-mail messages.
- Request forms of various items.

- Documents of legal records stored in various filing cabinets.
- Financial records.
- Operations.

Operations refer to functions, procedures and methods of doing things within the College. In the College staff only follow the responsibilities and functions assigned to them by their officer. There is no authentic documented way of handling duties and responsibilities except that staff operate within the ambit of their functions and what they can do to assist the system. Contractors use the bill of quantities given to them when bidding for contract and later sign agreement document prepared by the College.

2.2 Establish Context

In this section, the organization articulates its objectives and defines internal and external criteria and evaluation factors to consider in risk management [6]. In other words, it is the set of circumstances that surround and condition risk management both internally and externally [24]. The internal context is the internal environment in which the College seeks to achieve its objectives and it establishes the mission, vision and objectives of SYNCC College, the policies that are implemented, the culture of the college, its structure, strategy, and everything that affects the internal operation of the same. The external context establishes the current rules that apply to the college according to its economic activity, public policies, demography, trade, technology, and everything that has a relationship with the college from an external environment and that affects its operation [3]. For this case study, the internal and external factors of SYNCC are shown in Table 1.

2.2.1 System Acquisition

This is basically IT driven and refers to the way the system is developed, its life circle, security assessment, and vulnerability tests [13]. Due to the lack of technical competency, not all of these were followed when deploying a system. SYNCC regularly updates Linux O/S and services such as nginx, mysql, and drupal, but does not have any documented plan for handling lost passwords and access rights.

2.2.2 Internal Standards and Regulation

These are sets of code of practice and rules for security in computing [20]. At SYNCC, although they follow some standards, they are not documented and binding. The following are some of the standards they use.

- *Password standard*: must be a minimum of six characters, which includes a combination of upper, lower and special characters.
- *E-mail user standard*: The user must be a member of the College community, recommended by his HOD and the user ID must include his/her name.
- *Operating systems standard*: They use Ubuntu Linux on the server side.
- *Application standard*: Mostly application that runs on Linux platform.
- *Non-access of porn materials*: As a standard, SYNCC have content filtering engine at the gateway.
- *Vendors*: a vendor must register in Yemen, pay taxes, and be certified by the College tender board.

2.2.3 Incident Management

This capacity is needed to respond to security incidents, including forensic investigations, remedial actions to mitigate, and escalate the incidents to the next level [20]. In the event of a security breach, a staff on duty will be contacted. The staff attending to user's requests such as hardware maintenance, software installation and connectivity issues. During serious incidents such as virus infections, equipment or Internet failure, staff on duty and another staff will help in solving the problem. Staff monitor the servers and firewall regularly to make sure that no breaches have occurred.

2.2.4 Business Continuity Plan

The main objective of business continuity management is to neutralize interruptions to business activities, to safeguard critical business processes from the consequences of major failures of information systems, and to ensure their quick recovery [26]. In SYNCC the following measures have been implemented to allow ICT business activities to cope with failures:

- They keep a backup of important information on different media types.
- Installed and configured multiple servers as backup.
- Multiple environmental control devices.
- Power backup such as uninterruptible power supplies (UPS) and Diesel generator.
- Maintain and upgrade staff skill regularly as well as recruiting more staff to avoid losing key staff.

2.3 Risk Assessment

Risk assessment involves the identification of risks followed by their evaluation or ranking. It is important to have a template for recording appropriate information about each risk. The evaluation is subject to the results obtained in the execution of the management since all the companies have characteristics that make them different.

2.3.1 Identify Risks

The risk is the possibility that an event occurs that will have an impact on the achievement of the objectives. Risk is measured in terms of impact and likelihood [4]. Identification is the most important step for risk management since the risk that is not identified at this point will not be taken into account in the subsequent analysis and therefore will not be evaluated. According to Renn [22], to identify the risk it is necessary to have an adequate tool or technique defined by each entity or person according to its criteria. In addition, the people who execute this activity must have the appropriate knowledge about the process to be evaluated. All risks must be taken into account which means that both those with applied controls and those who do not have to be listed should be listed.

In SYCC, the potential risks are identified using two techniques. The first one is what-if analysis technique in which simple questions and scenarios are assumed to see if they can help to identify new risks and see what plans they need or already have in place to manage with these events. The second technique that is used to identify the potential risks is the automated scanning tools.

2.3.2 Risk Analysis and Evaluation

The objectives at this step are to separate the minor risks from major ones. This process includes identifying threats to information or information systems, determining the likelihood of occurrence of the threat, and identifying system vulnerabilities that could be exploited by the threat [16].

The potential risks encountered at SYNCC can break down into the following nine main categories:

- Epileptic nature of electricity supply.
- Harsh weather conditions. Very hot during the summer and cold winter.
- Possibility of prolonged downtime due to foreseen or unforeseen circumstances such as political issues and equipment failure.
- Continuous rise in software and support costs due to increase in both demand and complexities.
- Misuse of login details by staff and users.
- Lack of adequate skills to manage and implement systems.
- Lack of computer security policies.
- Experienced staff leaving for higher paying jobs.
- Threats against physical security for Crimes, Robberies, and Terrorism.

In SYNCC, they believe that these risk categories, and accordingly the risks included, can be rated as high, medium and low. Table 2 shows risk categories and their ranks.

Table 2: Risk assessment

Threat	Rate
Epileptic Power supply	2
Harsh weather	1
Equipment downtime	1
Rise of software and support	1
Cyber-criminal activities	2
Misuse of login details	2
Lack of adequate skills	2
Lack computer security policies	2
Experience staff leaving for high payment	2
Threats against physical security	2

2.3.3 Assets Management and Valuations

The definition of the risk profile allows comparing the results of its rating with the criteria defined to establish the degree of exposure of the process to risks [6]. In this way, it is possible to distinguish between acceptable, moderate or unacceptable risks and to set the priorities of the actions required for their treatment. Here, they manage the most important assets in the following ways:

Identification: They identify their hardware assets based on location and usage.

Location: Here they locate equipment based on usage *e.g.*, core network equipment, servers, and server software are in the datacenter; access devices located in the departments and units.

Responsible person: This depends on the location and usage *e.g.* Datacenter staff are responsible for datacenter holdings while personal/official access devices are to be the responsibility of the person using them. Network section staff are responsible for network access devices such as wired and wireless.

Protection: They deploy many measures to protect the assets such as power backup, physical security, access control, permissions.

In SYNCC, based on the value of the assets their corresponding importance: 5-critical, 4-very important, 3-important, and 2-good are given as shown in Table 3.

2.3.4 Treat the Risks

Once the risk has been fully identified, it is necessary to carry out its treatment [18]. Therefore, it is important to establish all the possible actions to mitigate the risk. These actions must be realizable and be effective in terms of mitigating the risk, and include:

- Defining new strategies.
- Plans for improvement.

Table 3: Asset management and valuation

Asset	Category	Location	Owner	Rate
Student information system	Information	Datacenter	Datacenter	5
Web sites	Information	Datacenter	Datacenter	3
e-mail messages	Information	Datacenter	Datacenter	3
Backup data	Information	Datacenter	Datacenter	4
Bind DNS	Application	Datacenter	Datacenter	4
Nginx	Application	Datacenter	Datacenter	4
Apache	Application	Datacenter	Datacenter	4
Drupal	Application	Datacenter	Datacenter	4
Servers	Hardware	Datacenter	Datacenter	4
Border routers	Hardware	Datacenter	Datacenter	3
Firewall	Hardware	Datacenter	Datacenter	3

Table 4: Risk management

Asset	Value	Risk	Recovery cost	Priority	Mitigating risk
Student information system	5	High	High	High	Avoid or limit
Web sites	3	Low	Low	Low	Accept
e-mail messages	3	Medium	Low	Medium	Limit
Backup data	4	Medium	Low	Medium	Limit
Bind DNS	4	High	Low	High	Avoid or limit
Nginx	4	High	Low	High	Avoid or limit
Apache	4	High	Low	High	Avoid or limit
Zimbra	3	Medium	Medium	Medium	Limit
Drupal	4	High	Low	High	Limit
Server hardware	4	Low	Low	Low	Accept
Border routers	3	Medium	Low	Medium	Limit
Firewall	3	Low	Low	Low	Accept
Access switches	2	Low	Low	Low	Accept
Environmental conditions	3	Low	Low	Low	Accept
Power backup	3	Low	Low	Low	Accept
Developers	3	Medium	Medium	Medium	Limit
Sys/Network admin	3	Medium	Medium	Medium	Limit

- Physical readjustments.
- Process optimization.
- Others.

In addition, the objectives at this step are to develop and implement a management plan to mitigate identified risks. In addition, the process of taking actions to eliminate or reduce the probability of confidentiality, integrity, and availability of valued information assets being compromised to an acceptable limit [14].

There are three steps to risk mitigation: identify options, choose options, and implement options [2]. In SYNCC College, the risk is mitigated as follows:

- *Accept the risk*: When both the value of an asset and the risk are low, the risk is considered acceptable.
- *Limit the risk*: When the risk of an asset is high and can not be transferred, consider limiting it. This is done by updating systems, alternative power and restarting the system.
- *Avoid the risk*: Risk is avoided by either building an alternative system or by shutting down the system. Table 4 shows the summary of the risk mitigation process in SYNCC College.

2.4 Monitoring and Review

A regular process of review is performed to: identify new risks when they just arise, monitor existing risks and identify any changes that may influence the implemented risk controls, ensure that the existing risk controls are working effectively, and transfer the information on risks fully to appropriate parties. This allows the College to anticipate

and respond in advance to events that would otherwise cause damage to the College.

3 Recommendation and Critical Review

The overall discussion above analyzes how they manage computer security in SYNCC, Yemen. SYNCC started with assessing the ICT infrastructures and threats and finally identify risks and mitigate them. It is not clear whether the processes used by the College to categorize assets, threats and mitigation are correct since there are no formal documents that guides the staff in managing computer security. In addition, the staffs have no formal training in managing computer security as they depend only on residual knowledge. The methodology used in managing computer security does not conform to the minimum acceptable standard. The threats and risks identified were just very few and do not conform to each other. Many mitigations measures were not properly taken to contain risk posed by the threats. In general, SYNCC has to take into account the following recommendation in order to improve this security:

1) Develop computer security plan for SYNCC to cover all aspect of ICT and enforce the plan. The ICT staff can request the management to buy-in by making them believe in the benefits of securing the College information system. In addition, ICT staff can demonstrate the ROI (Return on Investment) when such policy is implemented.

2) Develop and implement acceptable policy such as password, email, internet access, backup and logins to secure server.

3) Risk assessment must not be based on assets values only, but on the likelihood of vulnerabilities and their impact to the system. Since all the assets have their inherent vulnerabilities, after the identification of vulnerabilities and their likelihood and impact, it is possible to identify the risk and suggest the adequate security controls to be implemented.

4) Educate users (i.e., employee and students) on computer security such as password, importance of security, safe browsing, virus prevention and updating operating systems.

5) SYNCC should use to secure data on IEEE 802.11 networks. For example, MAC address filter, WEP (Wireless Equivalent Privacy), WPA (Wi-Fi Protected Access), IEEE 802.11i (encapsulation of extensible authentication protocol) and IEEE 802.1X. In addition, they must apply function for protected from risks and threats as shown in Table 5.

6) There are new features SYNCC should use to save college assets. Like Gateway with VPN, Cosign integration in gateway, Bluesocket gateway, 802.1x, Checkpoint gateway and Radius Server with Uniqname and Passwords.

Table 5: Security function

Function	Property
Firewall	Blocks or permits traffic from each user based on their role.
Redirection	Monitors web traffic from unregistered users and redirects them to the gateway's server.
Web Server	Presents user with login web page.
Authentication	Authenticates using servers such as Active Directory, LDAP or RADIUS.
Role Based Access Control	Registered users are assigned a role. Roles can control access based on IP address, network, protocol, time and location.
QOS Server	Bandwidth per user can be limited by role.
VPN Server	A Virtual Private Network protects the privacy of all traffic from a user with encryption.

7) Encrypted Access for Students, Faculty and Staff of SYNCC: All students, faculty, and staff will be required to encrypt their traffic on the wireless network by use of a VPN client on their computers.

8) Information is an asset of great value to the College and as such must be protected. That is why in the implementation of information security should be sought. Safeguarding the accuracy and completeness of stored or transmitted information, the content of which should remain unchanged unless modified by authorized personnel.

9) SYNCC has some security risks, Table 6 describes these risks and how to overcome them.

4 Conclusion

Information system in SYNCC College has over time been expanded and updated, its components changed, and its software applications replaced or updated with newer versions. In addition, many personnel were hired and no security policies were implemented. These changes meant that new risks will surface and risks previously mitigated by some means may again become a concern. Thus, the

Table 6: Security risks

Assets	How to Protect
Student information system, Web sites and e-mail messages system	Implementation of information security should be sought, Safeguarding the accuracy and completeness of stored or transmitted information, the content of which should remain unchanged unless modified by authorized personnel.
Backup data	Having saved in a remote location keeps it safe in case anything goes badly wrong with your computer.
Bind DNS	By ASA, PIX, and FWSM firewalls, Cisco Intrusion Prevention System (IPS) and Cisco IOS NetFlow feature.
Nginx	Having configuring SSL, restricting access by IP and performing a security audit
Apache	Having keep up to date, protect from Denial of Service (DoS) attacks , permissions on ServerRoot directories and Watching Your Logs
Zimbra	Firewall and use login username and password with htaccess.
Drupal	The Drupal has database abstraction layer provides placeholder mechanisms to prevent SQL Injection vulnerabilities.

need for well-organized computer security management is clear. To help keep the cost down, the risk assessment should occur whenever an information asset is classified, purchased or a new project is developed. A security review is often helpful. For SYNCC to derive the benefits of computer security management, it must first develop and implement security policy covering all aspects of Information and Communication Technology components. By studying risk management, SYNCC should develop a computer security plan, implement acceptable policy and educate users on computer security.

References

[1] T. Aven, "On some recent definitions and analysis frameworks for risk, vulnerability, and resilience," *Risk Analysis*, vol. 31, no. 4, pp. 515–522, 2011.

[2] T. Aven and E. Zio, "Some considerations on the treatment of uncertainties in risk assessment for practical decision making," *Reliability Engineering & System Safety*, vol. 96, no. 1, pp. 64–74, 2011.

[3] K. M. Brown, "The role of internal and external factors in the discontinuation of off-campus students," *Distance education*, vol. 17, no. 1, pp. 44–71, 1996.

[4] A. Ekelhart, S. Fenz, and T. Neubauer, "Aurum: A framework for information security risk management," in *42nd Hawaii International Conference on System Sciences (HICSS'09)*, pp. 1–10, 2009.

[5] T. Ernawati, D. R. Nugroho, *et al.*, "It risk management framework based on ISO 31000: 2009," in *International Conference on System Engineering and Technology (ICSET'12)*, pp. 1–8, 2012.

[6] S. Fenz and A. Ekelhart, "Verification, validation, and evaluation in information security risk management," *IEEE Security & Privacy*, vol. 9, no. 2, pp. 58–65, 2011.

[7] P. Hopkin, *Fundamentals of Risk Management: Understanding, Evaluating and Implementing Effective Risk Management*, Kogan Page Publishers, 2017. ISBN-13: 978-0749483074.

[8] D. W. Hubbard, *The Failure of Risk Management: Why It's Broken and How to Fix It*, John Wiley & Sons, 2009. ISBN: 978-0-470-38795-5.

[9] IEC, "ISO 31010: 2009-11," *Risk Management–Risk Assessment Techniques*, 2009. (`https://www.iso.org/standard/51073.html`)

[10] ISO, "31000: 2009 risk management–principles and guidelines," *International Organization for Standardization, Geneva, Switzerland*, 2009.(`https://www.iso.org/iso-31000-risk-management.html`)

[11] S. R. Iyer, D. A. Rogers, B. J. Simkins, and J. Fraser, "Academic research on enterprise risk management," *Enterprise Risk Management*, pp. 419–439, 2010.

[12] C. Lalonde and O. Boiral, "Managing risks through ISO 31000: A critical analysis," *Risk Management*, vol. 14, no. 4, pp. 272–300, 2012.

[13] J. H. Lambert, Y. Y. Haimes, D. Li, R. M. Schooff, and V. Tulsiani, "Identification, ranking, and management of risks in a major system acquisition," *Reliability Engineering & System Safety*, vol. 72, no. 3, pp. 315–325, 2001.

[14] T. P. Layton, *Information Security: Design, Implementation, Measurement, and Compliance*, CRC Press, 2016. ISBN: 9781420013412 .

[15] M. Leitch, "ISO 31000: 2009—the new international standard on risk management," *Risk Analysis*, vol. 30, no. 6, pp. 887–892, 2010.

[16] S. S. Lim, T. Vos, A. D. Flaxman, G. Danaei, K. Shibuya, H. Adair-Rohani, M. A. AlMazroa, M. Amann, H. R. Anderson, K. G. Andrews, *et al.*, "A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: A systematic analysis for the global burden of disease study 2010," *The Lancet*, vol. 380, no. 9859, pp. 2224–2260, 2013.

[17] A. Liuksiala, *The Use of the Risk Management Standard ISO 31000 in Finnish Organizations*, 2012. (http://tampub.uta.fi/bitstream/handle/10024/84249/gradu06462.pdf;sequence=1)

[18] Microsoft, *Microsoft Esecurity Guide for Small Business*, 2004. (www.professorsteve.com/FACT_Sheets/MicrosofteSecurityGuideforSmallBusiness.pdf)

[19] T. R. Peltier, *Information Security Risk Analysis*, CRC press, 2005. (https://www.taylorfrancis.com/books/9781439839577)

[20] T. R. Peltier, *Information Security Policies, Procedures, and Standards: Guidelines for Effective Information Security Management*, CRC Press, 2016. (https://www.taylorfrancis.com/books/9780849390326)

[21] G. Purdy, "ISO 31000: 2009— setting a new standard for risk management," *Risk Analysis*, vol. 30, no. 6, pp. 881–886, 2010.

[22] O. Renn, "Three decades of risk research: Accomplishments and new challenges," *Journal of Risk Research*, vol. 1, no. 1, pp. 49–71, 1998.

[23] J. Shortreed, "Enterprise risk management and ISO 31000," *The Journal of Policy Engagement*, vol. 2, no. 3, p. 9, 2010.

[24] M, Sumner, "Risk factors in enterprise-wide/erp projects," *Journal of Information Technology*, vol. 15, no. 4, pp. 317–327, 2000.

[25] S. Q. Wang, M. F. Dulaimi, and M. Y. Aguria, "Risk management framework for construction projects in developing countries," *Construction Management and Economics*, vol. 22, no. 3, pp. 237–252, 2004.

[26] P. Woodman, "Business continuity management," *Chartered Management Institute, Savoy Court, Strand, London*, pp. 8–12, 2007.

Biography

Abdullah Abdulrahman graduated from Al-Ahgaff University Hathramout in 2009. He works in communtiy college from 2009. He enrolled to study Master in University Kebangsaan Malaysia (UKM) in Computer Sceince (Network Technology) in 2015. His research interests are in Software Defined Network (SDN) and IP Security (IPSec), Computer Security.

Dr. Mohammed A. Hassan is assistant Professor at Department of Information Systems, Seiyun Community College, Yemen. He received the B.S. degree in Mathematics and Computer Science from the University of Al-Ahgaff, Yemen, in 2002, the M.S. degree in Computer Science from The University of Hamdard, India, in 2008 and Ph.D. degree in Computer Science from The Central University of Hyderabad, India, in 2014. His research interests include human visual system models for solving image and video processing problems.

On the Security of a Certificateless Proxy Signature Scheme in the Standard Model

Caixue Zhou, Xiwei Dong, Lihua Wang, and Tao Li
(Corresponding author: Caixue Zhou)

School of Information Science and Technology, Jiujiang University
551 Qianjin Donglu, Jiujiang 332005, China
(Email: charlesjjjx@126.com)

(Received Dec. 10, 2017; Revised and Accepted Apr. 12, 2018; First Online Mar. 2, 2019)

Abstract

Certificateless cryptosystem can overcome the costly certificate management in the traditional public key cryptosystem, and meanwhile it does not have the private key escrow problem in the identity-based cryptosystem. Proxy signature can allow a proxy signer authorized by an original signer to sign messages on behalf of the latter. In this paper, we show that a recently proposed certificateless proxy signature scheme in the standard model is vulnerable to the public key replacement attack. Through this kind of attack, a malicious original signer or proxy signer can forge a valid proxy signature. We analyse the reasons for the success of the attack and point out the flaw in the proof of the original scheme.

Keywords: Bilinear Pairing; Certificateless Proxy Signature; Public Key Replacement Attack; Standard Model

1 Introduction

Certificateless cryptosystem [17] can simplify the costly certificate management in the traditional public key cryptosystem, and meanwhile to eliminate the private key escrow problem in the identity-based cryptosystem [16]. It has attracted a lot of attention since its introduction.

Proxy signature allows an original signer to delegate his/her signing power to a proxy signer [4,8–12,14,18,23,24]. Then the latter can sign messages on behalf of the former when the former is absent. It has been widely used in practice since its introduction.

By combining the certificateless cryptosystem and proxy signature, Li *et al.* [15] proposed the first certificateless proxy signature scheme by using bilinear pairings in 2005. But unfortunately, Yap *et al.* [25] pointed out that Li *et al.*'s scheme is vulnerable to the public key replacement attack in 2007. In the same year, Lu *et al.* [19] and Choi *et al.* [3] further gave an improvement to Li *et al.*'s scheme, respectively. However, neither of them gave the security proof of their schemes. In the aspects of provably secure certificateless proxy signature schemes, Chen *et al.* [2] gave a security model of certificateless proxy signature for the first time and a concrete provably secure scheme in 2009. Later, many provably secure certificateless proxy signature schemes [7,13,22] were proposed.

Considering the random oracle model [6] and the standard model [21], Canetti *et al.* [1] showed that security in the random oracle model cannot guarantee the security in the real world. Thus, it is very important to work out schemes that are secure in the standard model. Eslami *et al.* [5] took the first step in this respect. They proposed a certificateless proxy signature scheme in the standard model for the first time in 2012. But unfortunately, Lu *et al.* [20] pointed out that Eslami *et al.*'s scheme is vulnerable to the public key replacement attack and malicious KGC (Key Generation Center) attack in 2016. Lu *et al.* further proposed a new scheme and proved their scheme to be secure under the Squ-CDH assumption in the standard model. To the best of the authors' knowledge, only the above two certificateless proxy signature schemes have been proposed in the standard model till now. In this paper, we point out that Lu *et al.*'s scheme is still insecure. We give two public key replacement attacks to their scheme. We analyse the reasons for the success of this kind of attack and point out the flaw in the proof of the original scheme. Thus, designing a provably secure certificateless proxy signature scheme in the standard model is still an open problem.

The rest of the paper is organized as follows. In Section 2, we review Lu *et al.*'s scheme. In Section 3, we give two public key replacement attacks to their scheme. Then we analyse the reasons for the success of the attack and point out the flaw in the proof of the original scheme. We conclude the paper in Section 4.

2 Lu *et al.*'s Scheme

Setup: Given a security parameter 1^k, the KGC chooses two cyclic groups G_1 and G_2 of prime order q, a random generator g of G_1, a bilinear map $e : G_1 \times G_1 \rightarrow$

G_2 and three hash functions $H_0 : \{0,1\}^* \to \{0,1\}^n$, $H_1, H_2 : \{0,1\}^* \to Z_q^*$. He/she randomly selects $\alpha \in Z_q^*$ and $g_2, u', u_1, \cdots, u_n, v_0, v_1, m_0, m_1 \in G_1$, and sets $g_1 = g^\alpha$. He/she defines a function: $F_u(id) = u' \prod_{j=1}^{n} u_j^{i_j}$, where $id = i_1 i_2 \cdots i_n$ is a bit string. Let $Q \in G_1$, and he/she also defines another function $f(Q)$. If the x-coordinate of Q is odd, then $f(Q) = 1$; else $f(Q) = 0$. The public parameters are

$$\begin{aligned} Params &= \{G_1, G_2, e, q, g_1, g_2, u', u_1, \cdots, u_n, v_0, v_1, \\ &\quad m_0, m_1, H_0, H_1, H_2, F_u, f\}, \end{aligned}$$

and the master private key is $msk = g_1^\alpha$.

Partial-Private-Key-Gen: Given a user U's identity ID_U, the KGC randomly selects $r_U \in Z_q^*$, and computes the user's partial private key as

$$psk_U = (psk_{U,1}, psk_{U,2}) = (g_1^\alpha \cdot F_u(id_U)^{r_U}, g^{r_U}),$$

where $id_U = H_0(ID_U)$.

Set-Secret-Value: The user U randomly selects $x_U \in Z_q^*$ as his/her secret value.

Set-Public-Key: The user U computes his/her public key as

$$\begin{aligned} PK_U &= (PK_{U,1}, PK_{U,2}, PK_{U,3}) \\ &= (g_1^{x_U}, g_2^{1/x_U}, e(g_1, g_1)^{x_U^2}). \end{aligned}$$

The public key can be verified by checking the following equations:

$$\begin{aligned} e(PK_{U,1}, PK_{U,2}) &= e(g_1, g_2) \text{ and } e(PK_{U,1}, PK_{U,1}) \\ &= PK_{U,3}. \end{aligned}$$

Set-Private-Key: The user U randomly selects $r'_U \in Z_q^*$, and computes his/her private key as

$$\begin{aligned} SK_U &= (SK_{U,1}, SK_{U,2}) \\ &= (psk_{U,1}^{x_U^2} \cdot F_u(id_U)^{r'_U}, psk_{U,2}^{x_U^2} \cdot g^{r'_U}), \end{aligned}$$

where $id_U = H_0(ID_U)$.

Delegation-Gen: The original signer O produces a warrant m_w. Then he/she randomly selects $s \in Z_q^*$ and computes the delegation as

$$\begin{aligned} DC_{OP} &= (DC_{OP,1}, DC_{OP,2}, DC_{OP,3}) \\ &= (g^s, SK_{O,2}, SK_{O,1} \cdot (PK_{O,2}^\gamma \cdot v_\lambda)^s), \end{aligned}$$

where

$$\begin{aligned} \lambda &= f(DC_{OP,2}), \\ \gamma &= H_1(DC_{OP,1}, DC_{OP,2}, ID_O, PK_O, m_w, v_\lambda). \end{aligned}$$

Delegation-Verify: The proxy signer P computes

$$\begin{aligned} id_O &= H_0(ID_O), \\ \lambda &= f(DC_{OP,2}) \end{aligned}$$

and

$$\gamma = H_1(DC_{OP,1}, DC_{OP,2}, ID_O, PK_O, m_w, v_\lambda),$$

and checks whether

$$\begin{aligned} e(DC_{OP,3}, g) &= e(PK_{O,1}, PK_{O,1}) \cdot \\ &\quad e(F_u(id_O), DC_{OP,2}) \cdot \\ &\quad e(PK_{O,2}^\gamma \cdot v_\lambda, DC_{OP,1}) \end{aligned}$$

holds. If it does, he/she accepts the delegation. Otherwise, the proxy signer asks the original signer to produce the delegation again.

Proxy-Sign: Let $M \in \{0,1\}^*$. The proxy signer P randomly selects $s', t \in Z_q^*$, and computes the proxy signature as

$$\begin{aligned} \sigma &= (\sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5) \\ &= (g^t, SK_{P,2}, DC_{OP,1} \cdot g^{s'}, DC_{OP,2}, DC_{OP,3} \cdot \\ &\quad (PK_{O,2}^\gamma \cdot v_\lambda)^{s'} \cdot SK_{P,1} \cdot (PK_{P,2}^\eta \cdot m_\mu)^t), \end{aligned}$$

where

$$\begin{aligned} \lambda &= f(\sigma_4), \\ \gamma &= H_1(DC_{OP,1}, \sigma_4, ID_O, PK_O, m_w, v_\lambda), \\ \mu &= f(\sigma_2) \end{aligned}$$

and

$$\eta = H_2(\sigma_1, \sigma_2, \sigma_3, \sigma_4, ID_O, PK_O, ID_P, PK_P, M, m_\mu).$$

Proxy-Verify: Given a proxy signature

$$\sigma = (\sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5)$$

on (M, m_w), a verifier first checks whether m is suitable for the warrant m_w. If it does, he/she computes

$$\begin{aligned} id_O &= H_0(ID_O), \\ id_P &= H_0(ID_P), \\ \lambda &= f(\sigma_4), \\ \gamma &= H_1(DC_{OP,1}, \sigma_4, ID_O, PK_O, m_w, v_\lambda), \\ \mu &= f(\sigma_2) \end{aligned}$$

and

$$\eta = H_2(\sigma_1, \sigma_2, \sigma_3, \sigma_4, ID_O, PK_O, ID_P, PK_P, M, m_\mu),$$

and checks whether the following equation holds:

$$\begin{aligned} e(\sigma_5, g) &= PK_{O,3} \cdot PK_{P,3} \cdot e(F_u(id_O), \sigma_4) \cdot e(PK_{O,2}^\gamma \cdot \\ &\quad v_\lambda, \sigma_3) \cdot e(F_u(id_P), \sigma_2) \cdot e(PK_{P,2}^\eta \cdot m_\mu, \sigma_1). \end{aligned}$$

If it does, he/she accepts the proxy signature.

Note: There is a clerical error in Lu *et al.*'s scheme. In order to verify the proxy signature, the verifier must know the value of $DC_{OP,1}$ to compute $\gamma = H_1(DC_{OP,1}, \sigma_4, ID_O, PK_O, m_w, v_\lambda)$ correctly. Thus, the proxy signer must transmit it with the proxy signature, which makes the proxy signature longer than the original scheme. In fact, we can replace $DC_{OP,1}$ with σ_3 in the computation of γ, and the length of proxy signature will not be added.

3 The Weakness of Lu *et al.*'s Scheme

In this section, we will show that Lu *et al.*'s scheme is vulnerable to the public key replacement attack of Type-I adversary. Then we point out the flaw in Lu *et al.*'s security proof. The formal definition and security model of certificateless proxy signature can be found in Lu *et al.*'s paper.

3.1 Public Key Replacement Attack

1) A malicious original signer O can forge a valid proxy signature without the private key of the proxy signer P.

According to the game of Definition 2 in Lu *et al.*'s paper, in the Queries stage, the malicious original signer O first randomly chooses $x'_P \in Z_q^*$ and computes

$$\begin{aligned} PK'_P &= (PK'_{P,1}, PK'_{P,2}, PK'_{P,3}) \\ &= (g_1^{x'_P}, g_2^{1/x'_P}, e(g_1, g_1)^{(x'_P)^2}). \end{aligned}$$

Subsequently, he/she makes a ReplacePublicKey oracle query to replace the public key of the proxy signer P with the new value PK'_P. In the Forgery stage, he/she randomly chooses $t, s, s', r_P \in Z_q^*$, an arbitrary message M and a warrant m_w. Then he/she computes

$$\begin{aligned} \sigma_1 &= g^t, \sigma_2 = g^{r_P}, \sigma_3 = g^{s+s'}, \sigma_4 = psk_{O,2}^{x_O^2+(x'_P)^2}, \\ \lambda &= f(\sigma_4), \\ \mu &= f(\sigma_2), \\ \gamma &= H_1(g^s, \sigma_4, ID_O, PK_O, m_w, v_\lambda), \\ \eta &= H_2(\sigma_1, \sigma_2, \sigma_3, \sigma_4, ID_O, PK_O, ID_P, \\ &\qquad PK'_P, M, m_\mu), \\ id_P &= H_0(ID_P), \\ id_O &= H_0(ID_O), \\ \sigma_5 &= (psk_{O,1})^{x_O^2} \cdot (psk_{O,1})^{(x'_P)^2} \cdot (PK_{O,2}^\gamma \cdot v_\lambda)^{s+s'} \cdot \\ &\quad F_u(id_P)^{r_P} \cdot ((PK'_{P,2})^\eta \cdot m_\mu)^t. \end{aligned}$$

It can be verified that $\sigma = (\sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5)$ is a valid proxy signature.

$$\begin{aligned} & e(\sigma_5, g) \\ = \; & e((psk_{O,1})^{x_O^2} \cdot (psk_{O,1})^{(x'_P)^2} \cdot (PK_{O,2}^\gamma \cdot v_\lambda)^{s+s'} \\ & \cdot F_u(id_P)^{r_P} \cdot ((PK'_{P,2})^\eta \cdot m_\mu)^t, g) \\ = \; & e((psk_{O,1})^{x_O^2}, g) \cdot e((psk_{O,1})^{(x'_P)^2}, g) \cdot \\ & e((PK_{O,2}^\gamma \cdot v_\lambda)^{s+s'}, g) \cdot e(F_u(id_P)^{r_P}, g) \\ & \cdot e(((PK'_{P,2})^\eta \cdot m_\mu)^t, g) \\ = \; & e((g_1^\alpha \cdot F_u(id_O)^{r_O})^{x_O^2}, g) \cdot \\ & e((g_1^\alpha \cdot F_u(id_O)^{r_O})^{(x'_P)^2}, g) \\ & \cdot e((PK_{O,2}^\gamma \cdot v_\lambda)^{s+s'}, g) \cdot \\ & e(F_u(id_P)^{r_P}, g) \cdot e(((PK'_{P,2})^\eta \cdot m_\mu)^t, g) \\ = \; & PK_{O,3} \cdot PK'_{P,3} \cdot e(F_u(id_O), g^{r_O(x_O^2+(x'_P)^2)}) \cdot \\ & e((PK_{O,2}^\gamma \cdot v_\lambda)^{s+s'}, g) \cdot \\ & e(F_u(id_P)^{r_P}, g) \cdot e(((PK'_{P,2})^\eta \cdot m_\mu)^t, g) \\ = \; & PK_{O,3} \cdot PK'_{P,3} \cdot e(F_u(id_O), psk_{O,2}^{(x_O^2+(x'_P)^2)}) \cdot \\ & e((PK_{O,2}^\gamma \cdot v_\lambda)^{s+s'}, g) \cdot \\ & e(F_u(id_P)^{r_P}, g) \cdot e(((PK'_{P,2})^\eta \cdot m_\mu)^t, g) \\ = \; & PK_{O,3} \cdot PK'_{P,3} \cdot e(F_u(id_O), \sigma_4) \cdot \\ & e(PK_{O,2}^\gamma \cdot v_\lambda, \sigma_3) \cdot e(F_u(id_P), \sigma_2) \cdot \\ & e((PK'_{P,2})^\eta \cdot m_\mu), \sigma_1). \end{aligned}$$

Therefore, the malicious original signer O wins the game with probability 1.

2) A malicious proxy signer P can forge a valid proxy signature without the authorization of the original signer O.

According to the game of Definition 2 in Lu *et al.*'s paper, in the Queries stage, the malicious proxy signer P first randomly chooses $x'_O \in Z_q^*$ and computes

$$\begin{aligned} PK'_O &= (PK'_{O,1}, PK'_{O,2}, PK'_{O,3}) \\ &= (g_1^{x'_O}, g_2^{1/x'_O}, e(g_1, g_1)^{(x'_O)^2}). \end{aligned}$$

Subsequently, he/she makes a ReplacePublicKey oracle query to replace the public key of the original signer O with the new value PK'_O. In the Forgery stage, he/she randomly chooses $t, s, s', r_O \in Z_q^*$, an arbitrary message M and a warrant m_w. Then he/she computes

$$\begin{aligned} \sigma_1 &= g^t, \sigma_2 = psk_{P,2}^{(x'_O)^2+x_P^2}, \sigma_3 = g^{s+s'}, \sigma_4 = g^{r_O}, \\ \lambda &= f(\sigma_4), \\ \mu &= f(\sigma_2), \\ \gamma &= H_1(g^s, \sigma_4, ID_O, PK'_O, m_w, v_\lambda), \\ \eta &= H_2(\sigma_1, \sigma_2, \sigma_3, \sigma_4, ID_O, PK'_O, ID_P, \\ &\qquad PK_P, M, m_\mu), \end{aligned}$$

$$
\begin{aligned}
id_P &= H_0(ID_P),\\
id_O &= H_0(ID_O),\\
\sigma_5 &= (psk_{P,1})^{(x'_O)^2}\cdot(psk_{P,1})^{x_P^2}\\
&\quad\cdot((PK'_{O,2})^{\gamma}\cdot v_\lambda)^{s+s'}\\
&\quad\cdot F_u(id_O)^{r_O}\cdot(PK^{\eta}_{P,2}\cdot m_\mu)^t.
\end{aligned}
$$

It can be verified that $\sigma = (\sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5)$ is a valid proxy signature.

$$
\begin{aligned}
& e(\sigma_5, g)\\
=\ & e((psk_{P,1})^{(x'_O)^2}\cdot(psk_{P,1})^{x_P^2}\cdot((PK'_{O,2})^{\gamma}\\
& \cdot v_\lambda)^{s+s'}\cdot F_u(id_O)^{r_O}\cdot(PK^{\eta}_{P,2}\cdot m_\mu)^t, g)\\
=\ & e((psk_{P,1})^{(x'_O)^2}, g)\cdot e((psk_{P,1})^{x_P^2}, g)\\
& \cdot e(((PK'_{O,2})^{\gamma}\cdot v_\lambda)^{s+s'}, g)\\
& \cdot e(F_u(id_O)^{r_O}, g)\cdot e((PK^{\eta}_{P,2}\cdot m_\mu)^t, g)\\
=\ & e((g_1^{\alpha}\cdot F_u(id_P)^{r_P})^{(x'_O)^2}, g)\\
& \cdot e((g_1^{\alpha}\cdot F_u(id_P)^{r_P})^{x_P^2}, g)\cdot e(((PK'_{O,2})^{\gamma}\\
& \cdot v_\lambda)^{s+s'}, g)\cdot e(F_u(id_O)^{r_O}, g)\\
& \cdot e((PK^{\eta}_{P,2}\cdot m_\mu)^t, g)\\
=\ & PK'_{O,3}\cdot PK_{P,3}\cdot e(F_u(id_P), g^{r_P((x'_O)^2+x_P^2)})\\
& \cdot e(((PK'_{O,2})^{\gamma}\cdot v_\lambda)^{s+s'}, g)\cdot e(F_u(id_O)^{r_O}, g)\\
& \cdot e((PK^{\eta}_{P,2}\cdot m_\mu)^t, g)\\
=\ & PK'_{O,3}\cdot PK_{P,3}\cdot e(F_u(id_P), psk_{P,2}^{((x'_O)^2+x_P^2)})\\
& \cdot e(((PK'_{O,2})^{\gamma}\cdot v_\lambda)^{s+s'}, g)\cdot e(F_u(id_O)^{r_O}, g)\\
& \cdot e((PK^{\eta}_{P,2}\cdot m_\mu)^t, g)\\
=\ & PK'_{O,3}\cdot PK_{P,3}\cdot e(F_u(id_P), \sigma_2)\\
& \cdot e((PK'_{O,2})^{\gamma}\cdot v_\lambda, \sigma_3)\cdot e(F_u(id_O), \sigma_4)\\
& \cdot e(PK^{\eta}_{P,2}\cdot m_\mu, \sigma_1).
\end{aligned}
$$

Therefore, the malicious proxy signer P wins the game with probability 1.

3.2 The Flaw in Lu *et al.*'s Security Proof

First, let's see the σ_5 in a proxy signature.

$$
\begin{aligned}
\sigma_5 &= DC_{OP,3}\cdot(PK^{\gamma}_{O,2}\cdot v_\lambda)^{s'}\cdot SK_{P,1}\cdot(PK^{\eta}_{P,2}\cdot m_\mu)^t\\
&= SK_{O,1}\cdot SK_{P,1}\cdot(PK^{\gamma}_{O,2}\cdot v_\lambda)^{s+s'}\cdot(PK^{\eta}_{P,2}\cdot m_\mu)^t\\
&= g_1^{\alpha(x_O^2+x_P^2)}\cdot F_u(id_O)^{(r_O x_O^2+r'_O)}\cdot F_u(id_P)^{(r_P x_P^2+r'_P)}\\
&\quad\cdot(PK^{\gamma}_{O,2}\cdot v_\lambda)^{s+s'}\cdot(PK^{\eta}_{P,2}\cdot m_\mu)^t.
\end{aligned}
$$

To the original signer O, all the ephemeral variables can be randomly chosen by himself/herself. Therefore, only the master private key g_1^{α} and the proxy signer's secret value x_P are unknown to him/her. In addition, σ_1, σ_2, σ_3 and σ_4 can also be computed by himself/herself by randomly choosing all the ephemeral variables. Therefore, to forge a proxy signature, he/she just needs to know $g_1^{\alpha(x_O^2+x_P^2)}$ and the proxy signer's secret value x_P.

Through public key replacement attack, O can choose x'_P as the secret value of proxy signer P. Therefore, he/she just needs to compute $g_1^{\alpha(x_O^2+x_P^2)}$ to forge a proxy signature.

Holding the partial private key $psk_{O,1}$, he/she can compute

$$
\begin{aligned}
& (psk_{O,1})^{x_O^2}\cdot(psk_{O,1})^{(x'_P)^2}\\
=\ & g_1^{\alpha(x_O^2+(x'_P)^2)}\cdot F_u(id_O)^{r_O(x_O^2+(x'_P)^2)},
\end{aligned}
$$

which includes the $g_1^{\alpha(x_O^2+(x'_P)^2)}$. Therefore, computing σ_5 now becomes very simple and he/she can compute

$$
\begin{aligned}
\sigma_5 &= (psk_{O,1})^{x_O^2}\cdot(psk_{O,1})^{(x'_P)^2}\cdot(PK^{\gamma}_{O,2}\cdot\\
&\quad v_\lambda)^{s+s'}\cdot F_u(id_P)^{r_P}\cdot((PK'_{P,2})^{\eta}\cdot m_\mu)^t\\
&= g_1^{\alpha(x_O^2+(x'_P)^2)}\cdot F_u(id_O)^{r_O(x_O^2+(x'_P)^2)}\\
&\quad\cdot(PK^{\gamma}_{O,2}\cdot v_\lambda)^{s+s'}\cdot F_u(id_P)^{r_P}\cdot((PK'_{P,2})^{\eta}\cdot m_\mu)^t
\end{aligned}
$$

by randomly choosing $t, s, s', r_P \in Z_q^*$.

By setting

$$
\begin{aligned}
\sigma_2 &= g^{r_P}\\
\sigma_4 &= psk_{O,2}^{(x_O^2+(x'_P)^2)}\\
&= g_1^{r_O(x_O^2+(x'_P)^2)}.
\end{aligned}
$$

The original signer O can forge a valid proxy signature successfully.

Now, let's look at the proof of Theorem 1 in Lu *et al.*'s scheme. In Case 2 and Case 3 of the Forgery stage, Lu *et al.* proved that if a malicious original signer can output a forgery of a valid proxy signature, Challenger will solve the Squ-CDH problem with a non-negligible advantage.

Note that there is a condition when the above Theorem holds − during the forging of a proxy signature, there must be an unknown part (which is generally a private key) in computing the forged proxy signature. While in our attack, all parts are known to the original signer O in forging a valid proxy signature. Therefore the condition is not held, the proof is certainly wrong.

4 Conclusions

In this paper, we give two public key replacement attacks to a recently proposed certificateless proxy signature scheme in the standard model. Then we analyze the reasons for the success of the attack and point out the flaw in the proof of the original scheme. Therefore, designing a provably secure certificateless proxy signature scheme in the standard model is still an open problem.

Acknowledgments

This work was supported by the National Natural Science Foundation of China [grant numbers 61462048, 61562047 and 61662039]. We would like to present our thanks to Ms. Yan Di, who checked our manuscript.

References

[1] R. Canetti, O. Goldreich, and S. Halevi, "The random oracle methodology, revisited," *Journal of the ACM*, vol. 51, no. 4, pp. 557–594, 2004.

[2] H. Chen, F. T. Zhang, and R. S. Song, "Certificateless proxy signature scheme with provable security," *Journal of Software*, vol. 20, no. 3, pp. 692–701, 2009.

[3] K. Choi and D. Lee, "Certificateless proxy signature scheme," in *Proceedings of the 3rd International Conference on Multimedia, Information Technology and its Applications (MITA'07)*, pp. 437–440, Aug. 2007.

[4] L. Z. Deng, H. W. Huang, and Y. Y. Qu, "Identity based proxy signature from rsa," *International Journal of Network Security*, vol. 19, no. 2, pp. 229–235, 2017.

[5] Z. Eslami and N. Pakniat, "A certificateless proxy signature scheme secure in standard model," in *Proceedings of 2012 International Conference on Latest Computational Technologies (ICLCT'12)*, pp. 81–84, Mar. 2012.

[6] M. Hassouna, B. Barry, and E. Bashier, "A new level 3 trust hierarchal certificateless public key cryptography scheme in the random oracle model," *International Journal of Network Security*, vol. 19, no. 4, pp. 551–558, 2017.

[7] D. B. He, Y. T. Chen, and J. H. Chen, "An efficient certificateless proxy signature scheme," *Mathematical and Computer Modelling*, vol. 57, no. 9-10, pp. 2510–2518, 2013.

[8] M. S. Hwang, C. C. Lee, S. F. Tzeng, "A new proxy signature scheme for a specified group of verifiers," *Information Sciences*, vol. 227, pp. 102–115, 2013.

[9] M. S. Hwang, I. C. Lin, E. J. L. Lu, "A secure nonrepudiable threshold proxy signature scheme with known signers", *Informatica*, vol. 11, no. 2, pp. 1-8, Apr. 2000.

[10] M. S. Hwang, S. F. Tzeng and S. F. Chiou, "A nonrepudiable multi-proxy multi-signature scheme", *Innovative Computing, Information and Control Express Letters*, vol. 3, no. 3, pp. 259–264, 2009.

[11] M. S. Hwang, S. F. Tzeng, C. S. Tsai, "Generalization of proxy signature based on elliptic curves", *Computer Standards & Interfaces*, vol. 26, no. 2, pp. 73–84, 2004.

[12] C. C. Lee, T. C. Lin, S. F. Tzeng and M. S. Hwang, "Generalization of proxy signature based on factorization", *International Journal of Innovative Computing, Information and Control*, vol. 7, no. 3, pp. 1039–1054, 2011.

[13] J. G. Li, Y. Q. Li, and Y. C. Zhang, "Provably secure forward secure certificateless proxy signature scheme," *KSII Transactions on Internet and Information Systems*, vol. 7, no. 8, pp. 1972–1988, 2013.

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

[15] X. Li, K. Chen, and L. Sun, "Certificateless signature and proxy signature schemes from bilinear pairings," *Lithuanian Mathematical Journal*, vol. 45, no. 1, pp. 76–83, 2005.

[16] D. Liu, S. Zhang, H. Zhong, R. H. Shi, and Y. M. Wang, "An efficient identity-based online/offline signature scheme without key escrow," *International Journal of Network Security*, vol. 19, no. 1, pp. 127–137, 2017.

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

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

[19] R. B. Lu, D. K. He, and C. J. Wang, "Cryptanalysis and improvement of a certificateless proxy signature scheme from bilinear pairings," in *Proceedings of the 8th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD'07)*, pp. 285–290, July 2007.

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

[21] 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.

[22] S. Padhye and N. Tiwari, "Ecdlp-based certificateless proxy signature scheme with message recovery," *Transactions on Emerging Telecommunications Technologies*, vol. 26, no. 3, pp. 346–354, 2015.

[23] S. F. Tzeng, C. C. Lee, and M. S. Hwang, "A batch verification for multiple proxy signature", *Parallel Processing Letters*, vol. 21, no. 1, pp. 77–84, 2011.

[24] C. Y. Yang, S. F. Tzeng, M. S. Hwang, "On the efficiency of nonrepudiable threshold proxy signature scheme with known signers", Journal of Systems and Software, vol. 73, no. 3, pp. 507–514, 2004.

[25] W. S. Yap, S. H. Heng, and B. M. Goi, "Cryptanalysis of some proxy signature schemes without certificates," in *Proceedings of the 1st Workshop on Information Security Theory and Practices Smart Cards, Mobile and Ubiquitous Computing Systems (WISTP'07)*, pp. 115–126, May 2007.

Biography

Caixue Zhou received BS degree in Computer Science Department from Fudan University in 1988, Shanghai, China and MS degree in Space College of Beijing University of Aeronautics and Astronautics in 1991, Beijing, China. He is an Associate Professor in the School of Information Science and Technology, Jiujiang University,

Jiujiang, China since 2007. He is a member of the CCF (China Computer Federation) and a member of CACR(Chinese Association for Cryptologic Research). His research interests include applied cryptography, security of computer networks.

Xiwei Dong received BS degree in School of Computer Science and Technology from Shandong University of Technology in 2005, Zibo, China and MS degree in School of Computer Science and Technology from Dalian University of Technology in 2010, Dalian, China. He is currently a PhD candidate in the School of Computer Science and Technology at the Nanjing University of Posts and Telecommunications. He is a lecturer in the School of Information Science and Technology, Jiujiang University, Jiujiang, China since 2011. He is a member of the CCF (China Computer Federation). His research interests include applied cryptography, security of computer networks and pattern recognition.

Lihua Wang received BS degree in School of Computer and Control from Harbin University of Science and Technology in 2003, Harbin, China and MS degree in School of Computer Science and Technology from Huazhong University of Science and Technology in 2008, Wuhan, China. She is a lecturer in the School of Information Science and Technology, Jiujiang University, Jiujiang, China since 2009. Her research interests include applied cryptography and network security.

Tao Li received MS degree in School of Computer Science and Technology from Anhui University of Science and Technology in 2009, Anhui, China and PhD degree in School of Computer Science and Technology from Hohai University in 2016, Nanjing, China. Now he is a lecturer in the School of Information Science and Technology, Jiujiang University. Jiujiang, China. His research interests include cryptography, information security, wireless network security and vehicular network security.

A Dynamic ID Based Authenticated Group Key Agreement Protocol from Pairing

Shruti Nathani[1], B. P. Tripathi[1], and Shaheena Khatoon[2]
(Corresponding author: B. P. Tripathi)
Department of Mathematics, Govt. N. P. G. College of Science[1]
Raipur-492010, Chhattisgarh, India
S. O. S. in Mathematics, Pt. Ravi Shankar Shukla University[2]
Raipur-492010, Chhattisgarh, India
(Email: bhanu.tripathi@gmail.com)
(Received Jan. 26, 2018; revised and accepted June 3, 2018; First Online Apr. 21, 2019)

Abstract

In this paper we present an identity (ID) based dynamic authenticated group key agreement protocol. Our protocol satisfies all the required security attributes and also provide forward and backward confidentiality. The security of our protocol is based on the bilinear Diffie-Hellman(DH) assumption. We extend Lee *et al.* ID based authenticated key agreement protocol from two party to a group of users by using bilinear pairing.

Keywords: Backward Confidentiality; Bilinear Pairing; Dynamic; Forward Confidentiality; Group Key Agreement Protocol

1 Introduction

The most striking development in the history of cryptography was happened in 1976, when Diffie *et al.* [12] proposed their revolutionary concept of two party key agreement protocol whose security was based on the discrete logarithm problem. But this protocol was not suitable for group of users. Then in 1982, Ingemarsson *et al.* [16] proposed the first group key exchange protocol, but both of these schemes were vulnerable to the man in the middle attack because they did not authenticate the involved parties.

A key agreement protocol is said to provide key authentication, if each entity involved in the exchange is assured that no other entity can learn the shared secret key. A key agreement protocol which provides such a property is called an authenticated key agreement protocol (AKE) [21].

An authentication protocol allows a sender to send messages to a receiver through an insecure communication channel in such a way that the receiver can be convinced that the messages are indeed coming from the intended sender and them messages have not been modified by any adversary sitting in the middle of the communication channel. In short the aim of this type of protocols is to establish an authenticated link from the sender to the receiver. Authentication is a term which is used in a very broad sense. It is a service related to identification [21].

In 1984, Shamir [26] suggested the concept of Identity based cryptosystems where user's identities(such as email address, phone numbers, office location etc.), could be used as the public keys. Since then many identity based key agreement protocols [6, 11, 27, 29, 30, 34] have been proposed.

In the history of key agreement, a major breakthrough was happened, in 2000 when Joux [17] introduced his simple and elegant single round tripartite non-identibased key agreement protocol which makes use of bilinear pairing on elliptic curves. This was the first positive application of pairings in cryptography [13].

In 2001, Bohen *et al.* [3] proposed, a first identity based encryption scheme using weil pairing. Since then many ID based cryptographic scheme using pairing have been proposed in cryptography and is currently an area of very active research [13].

1.1 Literature Review

Based on weil and Tate pairing techniques, Smart [30] in 2002, Chen *et al.* [6] in 2003, Scott [27] in 2002, Shim [29] in 2003, Cullagh [11] in 2004, Lee *et al.* [20] in 2005 designed identity based and authenticated two party key agreement protocols. Cheng *et al.* [8] pointed out that Chen *et al.* [6] protocol is not secure against unknown key share attack. The protocol of Scott [27] is not secure against man in the middle attack. Sun *et al.* [33] showed that the protocol of Shim [29] is insecure against key compromise impersonation attack or man-in-the-middle attack. Also Choo [10] showed that protocol of Cullagh *et al.* [11] is insecure against key revealing attacks.

Since the protocol of Joux [17] was a unauthenticated

key agreement for three party using pairing on elliptic curve. So later in 2002, Nalla *et al.* [23], proposed an authenticated tripartite ID-based key agreement scheme. But this scheme of Nalla *et al.* [23] was soon cryptanalyzed by chen [5] and Shim [28]. Then again in 2002, Zhang *et al.* [37], gave an ID- based one round authenticated three party key agreement protocol the authenticity of which is assured by the Id base signature scheme of Hess [15]. Another direction of research on key agreement is to generalize the two party key agreement to multi party setting and consider the dynamic scenario where participants may join or leave a multi-cast group at any given time.

1.2 Group Key Agreement

A group key agreement is a protocol allows a group of users to exchange information over public and insecure network to agree upon a common secret key which a group session key can be derived. As, the increased popularity of group oriented applications, such as e-learning, e-conference, video-conferencing etc, the design of an efficient authenticated group key agreement protocol has recently received much attention in the current research literature.

In 1995, Burmester *et al.* [4] gave a much more efficient two round key agreement protocol in multiparty setting. In 1996, Steiner *et al.* [31] gave a group key agreement protocol based on the natural extension of the DH key agreement protocol. Later, in 1998, Steiner *et al.* [32] gave a new approach to group key agreement. They studied the problem of key agreement in dynamic peer groups(DPG).

Also, Bresson *et al.* [2] formalized the first security model for group key agreement protocol extending the group key agreement between two or three parties [25]. Then in 2002, Nalla *et al.* [22] extends the ID based two party single round authenticated protocol of Smart [30] to multiparty ID-based key agreement in a tree based setting [14]. Later, in 2003, Barua *et al.* [1], extend the basic three party protocol of Joux [17] to multiparty setting by giving a ternary tree based unauthenticated key agreement protocol. Another group key agreement protocol which is a bilinear version of BD [4] protocol, was proposed by Choi *et al.* [9], in 2004. Later in 2005, a dynamic group key agreement protocol with two constant round was propose by Dutta *et al.* [14].

Many attempts have been performed to extend the Diffie *et al.* [12] two party protocol and the Joux [17] protocol for three party to $n-$participants that means to a group key exchange. Also we seen that in the current research literature of key agreement many ID based dynamic group key agreement schemes [7, 19, 35] by using bilinear pairing have been proposed.

Above we have summerized two and three party identity based key agreement protocols employing pairing operations. Many protocols of this type were proposed [11, 22,27,30,37] analyzed and some broken [5,10,28,29,33]. In this paper we focus on the Lee *et al.* [20] two party authenticated key agreement protocol and extend this two party protocol into a dynamic ID based authenticated group key agreement(DAGKA) using bilinear pairing.

2 Preliminaries

In this section, we briefly describe the notations, definitions, preliminary concepts and properties i.e. bilinear maps, computational problems, efficiency criteria and security attributes that we used later in the paper.

2.1 Bilinear Maps

Let G_1 be an additive group of prime order l and G_2 be a multiplicative group of the same order l. We assume discrete log problems in G_1 and G_2 are hard. We consider a pairing map $e : G_1 \times G_1 \to G_2$ satisfying the following properties [20].

Bilinearity. $e(aP, bQ) = e(P, Q)^{ab}$for all $P, Q \in G_1$, and for all $a, b \in \mathbb{Z}_l^*$.

Non-degeneracy. The map does not send all pairs in $G_1 \times G_1$ to the identity in G_2. Observe that since G_1 and G_2 are groups of prime order this implies that if P is a generator of G_1, then $e(P, P)$ is a generator of G_2.

Computability. Given $P, Q \in G_1$, $e(P, Q)$ can be efficiently computable.

A bilinear map satisfying the three properties above is said to be an admissible bilinear map. We note that the Weil and Tate pairings associated with supersingular elliptic curves or abelian varieties can be modified to create such bilinear maps. We consider G_1 to be an additive abelian group defined on elliptic curves.

We consider an admissible bilinear map $e : G_1 \times G_1 \to G_2$ defined as above. Let P be a generator of G_1.

Bilinear Diffie Hellman Problem (BDHP):
The BDH problem in $< G_1, G_2, e >$ is as follows. Given $P, aP, bP, cP \in G_1$, compute $e(P, P)^{abc} \in G_2$ where a, b, c are randomly chosen from $\mathbb{Z}_l^*$. An algorithm is said to solve the BDH problem with an advantage of ϵ if

$$Pr[\mathcal{A}(P, aP, bP, cP) = e(P, P)^{abc}] \geq \epsilon$$

where the probability is over the random choice of $a, b, c \in \mathbb{Z}_l^*$, the random choice of $P \in G_1^*$, and the random bits of $\mathcal{A}$. We assume that BDHP is hard, in other words, there is no polynomial time algorithm to solve BDHP with non-negligible probability.

2.2 Security Attributes

Now we give the desirable security attributes of the key agreement protocols:

Known-key security. Each run of the protocol should result in a unique secret session key. The compromise of one session key should not com- promise other session keys.

Perfect Forward secrecy. If long-term private keys of all entities are compromised, the secrecy of previously established session keys should not be affected.

Key compromise impersonation. The compromise of an entity A's long- term private key will allow an adversary to impersonate A, but it should not be able the adversary to impersonate other entities to A.

Unknown-key share. An entity A ends up believing she shares a key with B and although this is in fact the case, B mistakenly believes the key is instead shared with an entity $E \neq A$.

Message confidentiality is one of the most important feature in secure group communication. Message confidentiality ensures that the sender confidential data which can be read only by an authorized and intended receiver. Specially in DGKA protocols message confidentiality is achieved mainly by the following two components [19]:

Forward confidentiality. While a group user leaves from the current group, he should not be able to calculate the new session key.

Backward confidentiality. While a new user joins into the current group, he should not be able to calculate the previous session key.

3 Lee *et al.'s* ID Based Key Agreement

In this section, we will introduce Lee *et al.'s* [20] two party ID based key agreement.

Initialization. Let G_1 and G_2 be two groups of prime order l, where G_1 is an additive group and G_2 is a multiplicative group. The discrete logarithm problems (DLP) in both G_1 and G_2 are assumed to be hard. Let P be a generator of G_1, and $H : \{0,1\}^* \rightarrow \mathbb{Z}_l^*$ be a cryptographic hash function. The key generation center (KGC) chooses a random number $s \in \mathbb{Z}_l^*$ and set $P_{pub} = sP$. The center publishes system parameters $Params =< G_1, G_2, l, e, P, P_{pub}, H >$ and keep s as the master key, which is known only by itself.

In addition to the system initialization, KGC performs the following private key issuing process.

Private key extraction. Let A and B be the two entities who are going to agree to some session keys. The identities of A and B are ID_A and ID_B, respectively. Their public keys and private keys are as follows: $A's$ public key is $P_A = H(ID_A)$ and the private key is $S_A = sP_A$. $B's$ public key is $P_B = H(ID_B)$ and the private key is $S_B = sP_B$. The pairs (P_ID, S_ID) for A and B serve as their static public/private key pairs.

Suppose two users A and B want to share a common secret. A and B have static private keys $S_A = sP_A$ and $S_B = sP_B$ obtained from KGC. Let $kdf : G_2 \times G_1 \times G_1 \rightarrow \{0,1\}^*$ be a key derivation function which can be readily found in a number of standard documents. A and B generate ephemeral private keys a and b, respectively. The corresponding ephemeral public keys are (V_A, W_A) and (V_B, W_B) where $V_A = aP_B$, $W_A = aS_A$, $V_B = bP_A$, $W_B = bS_B$. These are the data flow between A and B.

$$\begin{aligned} A &\Rightarrow B : (V_A, W_A); \\ B &\Rightarrow A : (V_B, W_B). \end{aligned}$$

User A computes $k_A = e(aP_A + V_B, W_B)^a$. User B computes $k_B = e(bP_B + V_A, W_A)^b$. Then the shared common secret between A and B is $K = kdf(k_A, P_A, P_B) = kdf(k_B, P_A, P_B) = kdf(e(P_A, P_B)^{(a+b)abs}, P_A, P_B)$.

4 Proposed Protocol

Let $U_0 = \{U_1, U_2, \cdots, U_n\}$ be the initial set of participants that want to generate a common key. Where U_n is the group leader. And

$$ID_0 = ID_{u_1} \parallel ID_{u_2} \parallel ..... \parallel ID_{u_n}.$$

4.1 Setup

Let G_1 and G_2 be two groups of prime order l where G_1 is an additive group and G_2 is a multiplicative group. The DLP in both G_1 and G_2 are assumed to be hard. Let P be a generator of G_1 and $H : \{0,1\}^* \rightarrow \mathbb{Z}_l^*$ be a cryptographic hash function. The user U_n randomly picks a value $s \in \mathbb{Z}_l^*$ and keeps s as master private key. The user U_n computes the master public key $P_{pub} = sP$ and publishes the system parameters. $param = \{G_1, G_2, l, e, P, P_{pub}, H\}$.

Private Key Extraction. For a given user U with identity string ID, the user U_n computes the public key $PK_{ID} = Q_{ID} = H(ID)$ and distributes the corresponding static private key $SK_{ID} = sQ_{ID}$ to the user via a secure channel. Thus user $U's$ public/private key pair is defined as PK_{ID}/SK_{ID}.

Round 1.

Step 1: The group leader U_n:

1.1 chooses his ephemeral private keys $a_n \leftarrow^R \mathbb{Z}_n$.;

1.2 ephemeral public keys are (V_n, W_n)sends: where $V_n = a_n P_{pub}$ and $W_n = a_n SK_{U_n}$;

1.3 broadcast in the group:
$U_n \rightarrow$: $(\{U_1, U_2, \cdots, U_{n-1}\}, V_n, W_n)$.

Step 2: User U_1 :

2.1 chooses his ephemeral private keys $a_1 \leftarrow^R \mathbb{Z}_n$;

2.2 computes $K_1 = sQ_{U_1} + a_1$ and $M_1 = h_1(U_1, ..., U_{n-1}, a_1)$;

2.3 also ephemeral public keys are (V_1, W_1) : where $V_1 = a_1 P_{pub}$ and $W_1 = a_1 SK_{U_1}$;

2.4 sends a request:

$$U_1 \rightarrow U_n : (U_1, K_1, M_1, V_1, W_1).$$

Step 3: The user U_n :

3.1 computes $a_1 = K_1 - sQ_{U_1}$;

3.2 checks if $M_1 = h_1(U_1, ..., U_{n-1}, a_1)$; if the equality does not hold, he quits;

3.3 broadcasts:
$U_n \rightarrow^*$: (V_1, W_1).

Step 4: Each User $U_i, i = 2, ..., n-1$:

4.1 chooses his ephemeral private keys $a_i \leftarrow^R \mathbb{Z}_n$;

4.2 computes $K_i = sQ_{U_i} + a_i$ and $M_i = h_1(U_1, ..., U_{n-1}, a_i)$;

4.3 ephemeral public keys are (V_i, W_i)sends: where $V_i = a_i P_{pub}$ and $W_i = a_i SK_{u_i}$;

4.4 sends

$$U_i \rightarrow U_n : (U_i, K_i, M_i, V_i, W_i).$$

Step 5: The User U_n:$i = 2, \cdots, n-1$:

5.1 computes $a_i = K_i - sQ_{U_i}$;

5.2 checks if $M_i = h_1(U_1, ..., U_{n-1}, a_i)$; if atleast one equality does not hold, he quits;

5.3 broadcasts:
$U_n \rightarrow$: (V_i, W_i).

Round 2.

User U_1 computes

$$\begin{aligned} K_1 &= e((V_2 \times V_3 \times \cdots \times V_n)P_{pub}, \\ &\quad (W_2 \times W_3 \times \cdots \times W_n)SK_{U_1})^{a_1}. \end{aligned}$$

User U_2 computes

$$\begin{aligned} K_2 &= e((V_1 \times V_3 \times \cdots \times V_n)P_{pub}, \\ &\quad (W_1 \times W_3 \times \cdots \times W_n)SK_{U_2})^{a_2}. \\ &\vdots \end{aligned}$$

User U_{n-1} computes

$$\begin{aligned} K_{n-1} &= e((V_1 \times \cdots \times V_{n-2} \times V_n)P_{pub}, \\ &\quad (W_1 \times \cdots \times W_{n-2} \times W_n)SK_{U_{n-1}})^{a_{n-1}}. \end{aligned}$$

Key Computation. Each user $U_i, i = 1, 2, \cdots, n-1$. Let $kdf : G_2 \times \underbrace{G_1 \times G_1 \times \cdots \times G_1}_{ntimes} \rightarrow \{0,1\}^*$ be a key derivation function which can be readily found in a number of standard documents. Thus the shared common group session key,

$$\begin{aligned} K &= kdf(K_1, Q_{U_1}, Q_{U_2}, \cdots, Q_{U_n}) \\ &= kdf(K_2, Q_{U_1}, Q_{U_2}, \cdots, Q_{U_n}) \\ &\vdots \\ &= kdf(K_{(n-1)}, Q_{U_1}, Q_{U_2}, \cdots, Q_{U_n}). \end{aligned}$$

For user U_1,

$$\begin{aligned} K &= kdf(K_1, Q_{U_1}, Q_{U_2}, \cdots, Q_{U_n}) \\ &= kdf(e((V_2 \times V_3 \times \cdots \times V_n)P_{pub}, \\ &\quad (W_2 \times W_3 \times \cdots \times W_n)SK_{U_1})^{a_1}, Q_{U_1}, \\ &\quad Q_{U_2}, \cdots, Q_{U_n}) \\ &= kdf(e((a_2 P_{pub} \times a_3 P_{pub} \times \cdots \times a_n P_{pub})P_{pub}, \\ &\quad (a_2 sQ_{U_2} \times a_3 sQ_{U_3} \times \cdots \times a_n sQ_{U_n})sQ_{U_1})^{a_1}, \\ &\quad Q_{U_1}, Q_{U_2}, \cdots, Q_{U_n}) \\ &= kdf(e((a_2 \times a_3 \times \cdots \times a_n)P_{pub}^n, \\ &\quad (a_2 \times a_3 \times \cdots \times a_n)s^n \\ &\quad (Q_{U_1} \times Q_{U_2} \times Q_{U_3} \times \cdots \times Q_{U_n}))^{a_1}, \\ &\quad Q_{U_1}, Q_{U_2}, \cdots, Q_{U_n}) \\ &= kdf(e((sP)^n, (Q_{U_1} \times \cdots \times Q_{U_n})s^n)^{(a_1 \times a_2 \times \cdots \times a_n)}, \\ &\quad Q_{U_1}, Q_{U_2}, \cdots, Q_{U_n}) \\ &= kdf(e(P^n, (Q_{U_1} \times ... \times Q_{U_n}))^{s^n(a_1 \times a_2 \times \cdots \times a_n)}, \\ &\quad Q_{U_1}, Q_{U_2}, \cdots, Q_{U_n}) \end{aligned}$$

For user U_2,

$$\begin{aligned} K &= kdf(K_2, Q_{U_1}, Q_{U_2}, \cdots, Q_{U_n}) \\ &= kdf(e((V_1 \times V_3 \times \cdots \times V_n)P_{pub}, \\ &\quad (W_1 \times W_3 \times \cdots \times W_n)SK_{U_2})^{a_2}, \\ &\quad Q_{U_1}, Q_{U_2}, \cdots, Q_{U_n}) \\ &= kdf(e((a_1 P_{pub} \times a_3 P_{pub} \times \cdots \times a_n P_{pub})P_{pub}, \\ &\quad (a_1 sQ_{U_1} \times a_3 sQ_{U_3} \times \cdots \times a_n sQ_{U_n})sQ_{U_2})^{a_2}, \\ &\quad Q_{U_1}, Q_{U_2}, \cdots, Q_{U_n}) \\ &= kdf(e((a_1 \times a_3 \times \cdots \times a_n)P_{pub}^n, \\ &\quad (a_1 \times a_3 \times \cdots \times a_n)s^n \\ &\quad (Q_{U_1} \times Q_{U_2} \times Q_{U_3} \times, ... \times Q_{U_n}))^{a_2}, \\ &\quad Q_{U_1}, Q_{U_2}, \cdots, Q_{U_n}) \\ &= kdf(e((sP)^n, (Q_{U_1} \times \cdots \\ &\quad \times Q_{U_n})s^n)^{(a_1 \times a_2 \times \cdots \times a_n)}, \\ &\quad Q_{U_1}, Q_{U_2}, \cdots, Q_{U_n}) \\ &= kdf(e(P^n, (Q_{U_1} \times \cdots \\ &\quad \times Q_{U_n}))^{s^n(a_1 \times a_2 \times \cdots \times a_n)}, \\ &\quad Q_{U_1}, Q_{U_2}, \cdots, Q_{U_n}). \end{aligned}$$

For user U_{n-1},

$$\begin{aligned}
K &= kdf(K_{n-1}, Q_{U_1}, Q_{U_2}, \cdots, Q_{U_n}) \\
&= kdf(e((V_1 \times \cdots \times V_{n-2} \times V_n)P_{pub}, \\
&\quad (W_1 \times \cdots \times W_{n-2} \times W_n)SK_{U_{n-1}})^{a_{n-1}}, \\
&\quad Q_{U_1}, Q_{U_2}, \cdots, Q_{U_n}) \\
&= kdf(e((a_1 P_{pub} \times \cdots \times a_{n-2}P_{pub} \\
&\quad \times a_n P_{pub})P_{pub}, \\
&\quad (a_1 sQ_{U_1} \times \cdots \times a_{n-2}sQ_{U_{n-2}} \\
&\quad \times a_n sQ_{U_n})sQ_{U_{n-1}})^{a_{n-1}}, \\
&\quad Q_{U_1}, Q_{U_2}, \cdots, Q_{U_n}) \\
&= kdf(e((a_1 \times \cdots \times a_{n-2} \times a_n)P_{pub}^n, \\
&\quad (a_1 \times \cdots \times a_{n-2} \times a_n)s^n \\
&\quad (Q_{U_1} \times Q_{U_2} \times \cdots \times Q_{U_n}))^{a_{n-1}}, \\
&\quad Q_{U_1}, Q_{U_2}, \cdots, Q_{U_n}) \\
&= kdf(e((sP)^n, (Q_{U_1} \times \cdots \\
&\quad \times Q_{U_n})s^n)^{(a_1 \times a_2 \times \cdots \times a_n)}, \\
&\quad Q_{U_1}, Q_{U_2}, \cdots, Q_{U_n}) \\
&= kdf(e(P^n, (Q_{U_1} \times cdots \\
&\quad \times Q_{U_n}))^{s^n(a_1 \times a_2 \times \cdots \times a_n)}, \\
&\quad Q_{U_1}, Q_{U_2}, \cdots, Q_{U_n}).
\end{aligned}$$

User n can compute the session key directly.

4.2 Join Algorithm

Let $U_{n+1}, U_{n+2}, \cdots, U_{n+m}$ be the set of users who will join the initial group U_0 , $U_j = U_1, \cdots, U_{n+m}$.

$$ID_j = ID_{U_1} \| \cdots \| ID_{U_{n+m}}.$$

As in the above protocol the user U_n is the group leader of this new group U_j also. When a new user joins the group it register itself to the group leader U_n by sending its identity $ID_{U_{n+i}}$. Then the join algorithm is executed in the following way:

Private Key Extraction. For each new registered user U_{n+i} the group leader U_n computes the public key $PK_{U_{n+i}} = Q_{U_{n+i}} = H(U_{n+i})$ and distributes the corresponding static private key $SK_{U_{n+i}} = sQ_{U_{n+i}}$ to the new joined users via a secure channel.

Round 1.

Step 1: The group leader U_n broadcasts in the group:

$$U_n \rightarrow: (\{U_1, \cdots, U_{n-1}, U_{n+1}, \cdots, U_{n+m}\}, V_i, W_i),$$

where $i = 1, 2, ..., n$.

Step 2: Each user $U_{n+i}, i = 1, ..., m$

2.1 Choose his ephemeral private keys $a_{n+i} \leftarrow^R \mathbb{Z}_n$;

2.2 Computes $K_{n+i} = sQ_{U_{n+i}} + a_{n+i}$ and $M_{n+i} = h_1(U_1, \cdots, U_{n-1}, U_{n+1}, \cdots, U_{n+m}, a_{n+i})$;

2.3 Computes $V_{n+i} = a_{n+i}P_{pub}$ and $W_{n+i} = a_{n+i}SK_{U_{n+i}}$;

2.4 Sends: $U_{n+i} \rightarrow U_n$: $(U_{n+i}, K_{n+i}, M_{n+i}, V_{n+i}, W_{n+i})$.

Step 3: The User U_n: $i = i, ..., m$

3.1 Computes $a_{n+i} = K_{n+i} - sQ_{U_{n+i}}$;

3.2 Checks if $M_{n+i} = h_1(U_1, \cdots, U_{n-1}, U_{n+1}, \cdots, U_{n+m}, a_{n+i})$;
if atleast one equality does not hold, he quits;

3.3 Broadcasts:
$U_n \rightarrow: (U_1, \cdots, U_{n-1}, U_{n+1}, \cdots, U_{n+m}, V_{n+i}, W_{n+i})$.

Round 2.

User U_1 computes $K_1 = e((V_2 \times V_3 \times \cdots \times V_{n+m})P_{pub}, (W_2 \times W_3 \times \cdots \times W_{n+m})SK_{U_1})^{a_1}$

$\vdots$

User U_{n-1} computes $K_{n-1} = e((V_1 \times \cdots \times V_{n-2} \times V_n \times \cdots \times V_{n+m})P_{pub}, (W_1 \times \cdots \times W_{n-2} \times W_n \times \cdots \times W_{n+m})SK_{U_{n-1}})^{a_{n-1}}$.

User U_{n+1} computes $K_{n+1} = e((V_1 \times \cdots \times V_n \times V_{n+2} \times \cdots \times V_{n+m})P_{pub}, (W_1 \times \cdots \times W_n \times W_{n+2} \times \cdots \times W_{n+m})SK_{U_{n+1}})^{a_{n+1}}$

$\vdots$

User U_{n+m} computes $K_{n+m} = e((V_1 \times V_2 \times \cdots \times V_{n+m-1})P_{pub}, (W_1 \times W_2 \times \cdots \times W_{n+m-1})SK_{U_{n+m}})^{a_{n+m}}$.

Key Computation. Each user $U_i, i = 1, 2, ..., n+m-1$. Let $kdf : G_2 \times \underbrace{G_1 \times G_1 \times \cdots \times G_1}_{(n+m)times} \rightarrow \{0,1\}^*$ be a key derivation function which can be readily found in a number of standard documents. Thus the shared common group session key,

$$\begin{aligned}
K &= kdf(K_1, Q_{U_1}, Q_{U_2}, ..., Q_{U_{n+m}}) \\
&= kdf(K_{n-1}, Q_{U_1}, Q_{U_2}, \cdots, Q_{U_{n+m}}) \\
&= kdf(K_{n+1}, Q_{U_1}, Q_{U_2}, \cdots, Q_{U_{n+m}}) \\
&\quad \vdots \\
&= kdf(K_{n+m}, Q_{U_1}, Q_{U_2}, ..., Q_{U_{n+m}}).
\end{aligned}$$

For user U_1,

$$\begin{aligned}
K &= kdf(K_1, Q_{U_1}, Q_{U_2}, \cdots, Q_{U_{n+m}}) \\
&= kdf(e((V_2 \times V_3 \times \cdots \times V_{n+m})P_{pub}, \\
&\quad (W_2 \times W_3 \times \cdots \times W_{n+m})SK_{U_1})^{a_1}, \\
&\quad Q_{U_1}, \cdots, Q_{U_{n+m}}) \\
&= kdf(e((a_2P_{pub} \times a_3P_{pub} \times \cdots \\
&\quad \times a_{n+m}P_{pub})P_{pub}, (a_2sQ_{U_2} \times a_3sQ_{U_3} \times \cdots \\
&\quad \times a_{n+m}sQ_{U_{n+m}})sQ_{U_1})^{a_1}, Q_{U_1}, \cdots, Q_{U_{n+m}}) \\
&= kdf(e((a_2 \times a_3 \times \cdots \times a_{n+m})P_{pub}^{n+m}, \\
&\quad (a_2 \times a_3 \times \cdots \times a_{n+m})s^{n+m}Q_{U_1} \times Q_{U_2} \times \\
&\quad \cdots \times Q_{U_{n+m}}))^{a_1}, Q_{U_1}, \cdots, Q_{U_{n+m}}) \\
&= kdf(e((sP)^{n+m}, (Q_{U_1} \times \cdots \\
&\quad \times Q_{U_{n+m}})s^{n+m})^{(a_1 \times a_2 \times \cdots \times a_{n+m})}, \\
&\quad Q_{U_1}, Q_{U_2}, \cdots, Q_{U_{n+m}}) \\
&= kdf(e(P^{n+m}, (Q_{U_1} \times \cdots \\
&\quad \times Q_{U_{n+m}}))^{s^{n+m}(a_1 \times a_2 \times \cdots \times a_{n+m})}, \\
&\quad Q_{U_1}, Q_{U_2}, \cdots, Q_{U_{n+m}}).
\end{aligned}$$

For user U_{n-1},

$$\begin{aligned}
K &= kdf(K_{n-1}, Q_{U_1}, Q_{U_2}, \cdots, Q_{U_{n+m}}) \\
&= kdf(e((V_1 \times \cdots \times V_{n-2} \times V_n \times \\
&\quad \cdots \times V_{n+m})P_{pub}, (W_1 \times \cdots \times W_{n-2} \times W_n \\
&\quad \times \cdots \times W_{n+m})SK_{U_{n-1}})^{a_{n-1}}, Q_{U_1}, \\
&\quad \cdots, Q_{U_{n+m}}) \\
&= kdf(e((a_1P_{pub} \times \cdots \times a_{n-2}P_{pub} \times a_nP_{pub} \\
&\quad \times \cdots \times a_{n+m}P_{pub})P_{pub}, (a_1sQ_{U_1} \times \cdots \\
&\quad \times a_{n-2}sQ_{U_{n-2}} \times a_nsQ_{U_n} \times \cdots \\
&\quad \times a_{n+m}sQ_{U_{n+m}})sQ_{U_{n-1}})^{a_{n-1}}, Q_{U_1}, \\
&\quad \cdots, Q_{U_{n+m}}) \\
&= kdf(e((a_1 \times \cdots \times a_{n-2} \times a_n \times \cdots \\
&\quad \times a_{n+m})P_{pub}^{n+m}, (a_1 \times \cdots \times a_{n-2} \times a_n \times \cdots \\
&\quad \times a_{n+m})s^{n+m}(Q_{U_1} \times \cdots \times Q_{U_{n-2}} \times Q_{U_n} \times \\
&\quad \cdots \times Q_{U_{n+m}})Q_{U_{n-1}})^{a_{n-1}}, Q_{U_1}, \cdots, Q_{U_{n+m}}) \\
&= kdf(e((sP)^{n+m}, (Q_{U_1} \times \cdots \times Q_{U_{n-2}} \times Q_{U_n} \times \\
&\quad \cdots \times Q_{U_{n+m}})s^{n+m}Q_{U_{n-1}})^{a_1 \times a_2 \times \cdots \times a_{n+m}}, \\
&\quad Q_{U_1}, \cdots, Q_{U_{n+m}}) \\
&= kdf(e(P^{n+m}, (Q_{U_1} \times \cdots \\
&\quad \times Q_{U_{n+m}}))^{s^{n+m}(a_1 \times a_2 \times \cdots \times a_{n+m})}, Q_{U_1}, \cdots, \\
&\quad Q_{U_{n+m}}).
\end{aligned}$$

For user U_{n+1},

$$\begin{aligned}
K &= kdf(K_{n+1}, Q_{U_1}, Q_{U_2}, \cdots, Q_{U_{n+m}}) \\
&= kdf(e((V_1 \times \cdots \times V_n \times V_{n+2} \cdots \\
&\quad \times V_{n+m})P_{pub}, (W_1 \times \cdots \times W_n \times W_{n+2} \times \cdots \\
&\quad \times W_{n+m})SK_{U_{n+1}})^{a_{n+1}}, Q_{U_1}, \cdots, Q_{U_{n+m}}) \\
&= kdf(e((a_1P_{pub} \times \cdots \times a_nP_{pub} \times a_{n+2}P_{pub} \times \cdots \\
&\quad \times a_{n+m}P_{pub})P_{pub}, (a_1sQ_{U_1} \times \cdots \times a_nsQ_{U_n} \\
&\quad \times a_{n+2}sQ_{U_{n+2}} \times \cdots \times a_{n+m}sQ_{U_{n+m}})sQ_{U_{n+1}})^{a_{n+1}}, \\
&\quad Q_{U_1}, \cdots, Q_{U_{n+m}}) \\
&= kdf(e((a_1 \times \cdots \times a_n \times a_{n+2} \times \cdots \times a_{n+m})P_{pub}^{n+m}, \\
&\quad (a_1 \times \cdots \times a_n \times a_{n+2} \times \cdots \times a_{n+m})s^{n+m}(Q_{U_1} \\
&\quad \times \cdots \times Q_{U_n} \times Q_{U_{n+2}} \times \cdots \times Q_{U_{n+m}})Q_{U_{n+1}})^{a_{n+1}}, \\
&\quad Q_{U_1}, \cdots, Q_{U_{n+m}}) \\
&= kdf(e((sP)^{n+m}, (Q_{U_1} \times \cdots \times Q_{U_n} \times Q_{U_{n+2}} \times \\
&\quad \cdots \times Q_{U_{n+m}})s^{n+m}Q_{U_{n+1}})^{a_1 \times a_2 \times \cdots \times a_{n+m}}, Q_{U_1}, \\
&\quad \cdots, Q_{U_{n+m}}) \\
&= kdf(e(P^{n+m}, (Q_{U_1} \times \cdots \\
&\quad \times Q_{U_{n+m}}))^{s^{n+m}(a_1 \times a_2 \times \cdots \times a_{n+m})}, Q_{U_1}, \\
&\quad \cdots, Q_{U_{n+m}}).
\end{aligned}$$

For userU_{n+m},

$$\begin{aligned}
K &= kdf(K_{n+m}, Q_{U_1}, Q_{U_2}, \cdots, Q_{U_{n+m}}) \\
&= kdf(e((V_1 \times V_2 \times \cdots \times V_{n+m-1})P_{pub}, (W_1 \\
&\quad \times W_2 \times \cdots \times W_{n+m-1})SK_{U_{n+m}})^{a_{n+m}}, \\
&\quad Q_{U_1}, \cdots, Q_{U_{n+m}}) \\
&= kdf(e((a_1P_{pub} \times a_2P_{pub} \times \cdots \times a_{n+m-1}P_{pub})P_{pub}, \\
&\quad (a_1sQ_{U_1} \times a_2sQ_{U_2} \times \cdots \\
&\quad \times a_{n+m-1}sQ_{U_{n+m-1}})sQ_{U_{n+m}})^{a_{n+m}}, \\
&\quad Q_{U_1}, \cdots, Q_{U_{n+m}}) \\
&= kdf(e((a_1 \times a_2 \times \cdots \times a_{n+m-1})P_{pub}^{n+m}, \\
&\quad (a_1 \times a_2 \times \cdots \times a_{n+m-1})s^{n+m}(Q_{U_1} \times Q_{U_2} \\
&\quad \times \cdots \times Q_{U_{n+m}}))^{a_{n+m}}, Q_{U_1}, \cdots, Q_{U_{n+m}}) \\
&= kdf(e((sP)^{n+m}, (Q_{U_1} \times Q_{U_2} \times \cdots \\
&\quad \times Q_{U_{n+m}})s^{n+m})^{a_1 \times a_2 \times \cdots \times a_{n+m}}, Q_{U_1}, \cdots, Q_{U_{n+m}}) \\
&= kdf(e(P^{n+m}, (Q_{U_1} \times \cdots \\
&\quad \times Q_{U_{n+m}}))^{s^{n+m}(a_1 \times a_2 \times \cdots \times a_{n+m})}, Q_{U_1}, \\
&\quad \cdots, Q_{U_{n+m}}).
\end{aligned}$$

4.3 Leave Algorithm

Without loss of generality, we assume that $U_{v-1} = \{U_1, U_2, \cdots, U_n\}$ is the current group that $L = \{U_1, \cdots, U_m\}$ is the set of leaving users. Then

$$\begin{aligned}
U_v &= \{U_{m+1}, \cdots, U_{m+n}, U_n\} \\
ID_v &= ID_{U_{m+1}} \| \cdots \| ID_{U_{m+n}} \| ID_{U_m}.
\end{aligned}$$

Then the leave algorithm is executed in the following way.

Round 1:

Step 1: The group leader U_n broadcasts in the group:

$$U_n \rightarrow: (\{U_{m+1}, \cdots, U_{m+n}, U_n\}).$$

Step 2: Each user $U_{m+i}, i = 1, ..., n$

2.1 Choose his ephemeral private keys $a_{m+i} \leftarrow^R \mathbb{Z}_n$;

2.2 Computes $K_{m+i} = sQ_{U_{m+i}} + a_{m+i}$ and

$$M_{m+i} = h_1(U_{m+1}, \cdots, U_{m+n}, a_{m+i});$$

2.3 Computes $V_{m+i} = a_{m+i}P_{pub}$ and $W_{m+i} = a_{m+i}SK_{U_{m+i}}$;

2.4 Sends:

$$U_{m+i} \rightarrow U_n : (U_{m+i}, K_{m+i}, M_{m+i}, V_{m+i}, W_{m+i}).$$

Step 3: The User U_n: $i = i, ..., n$

3.1 Computes $a_{m+i} = K_{m+i} - sQ_{U_{m+i}}$;

3.2 Checks if $M_{m+i} = h_1(U_{m+1}, \cdots, U_{m+n}, a_{m+i})$; if atleast one equality does not hold, he quits;

3.3 Broadcasts: $U_n \rightarrow$: $(\{U_{m+1}, \cdots, U_{m+n}\}, V_{m+i}, W_{m+i})$.

Round 2:

User U_{m+1} computes $K_{m+1} = e((V_{m+2} \times \cdots \times V_{m+n} \times V_n)P_{pub}, (W_{m+2} \times \cdots \times W_{m+n} \times W_n)SK_{U_{m+1}})^{a_{m+1}}$

$\vdots$

User U_{m+n} computes $K_{m+n} = e((V_{m+1} \times \cdots \times V_{m+n-1} \times V_n)P_{pub}, (W_{m+1} \times \cdots \times W_{m+n-1} \times W_n)SK_{U_{m+n}})^{a_{m+n}}$

Key Computation: Each user $U_{m+i}, i = 1, 2, ..., n$. Let $kdf : G_2 \times \underbrace{G_1 \times G_1 \times \cdots \times G_1}_{(m+n+1)times} \rightarrow \{0,1\}^*$ be a key derivation function which can be readily found in a number of standard documents. Thus the shared common group session key,

$$\begin{aligned} K &= kdf(K_{m+1}, Q_{U_{m+1}}, ..., Q_{U_{m+n}}, Q_{U_n}) \\ &\vdots \\ &= kdf(K_{(m+n)}, Q_{U_{m+1}}, ..., Q_{U_{m+n}}, Q_{U_n}). \end{aligned}$$

For user U_{m+1},

$$\begin{aligned} K &= kdf(K_{m+1}, Q_{U_{m+1}}, ..., Q_{U_{m+n}}, Q_{U_n}) \\ &= kdf(e((V_{m+2} \times \cdots \times V_{m+n} \times V_n)P_{pub}, \\ &\quad (W_{m+2} \times \cdots \times W_{m+n} \times W_n)SK_{U_{m+1}})^{a_{m+1}}, \\ &\quad Q_{U_{m+1}}, \cdots, Q_{U_{m+n}}, Q_{U_n}) \\ &= kdf(e((a_{m+2}P_{pub} \times \cdots \times a_{m+n}P_{pub} \\ &\quad \times a_nP_{pub})P_{pub}, (a_{m+2}sQ_{U_{m+2}} \times \cdots \\ &\quad \times a_{m+n}sQ_{U_{m+n}} \times a_nsQ_{U_n})sQ_{U_{m+1}})^{a_{m+1}}, \\ &\quad Q_{U_{m+1}}, \cdots, Q_{U_{m+n}}, Q_{U_n}) \\ &= kdf(e((a_{m+2} \times \cdots \times a_{m+n} \times a_n)P_{pub}^{m+n+1}, \\ &\quad (a_{m+2} \times \cdots \times a_{m+n} \times a_n)s^{m+n+1}(Q_{U_{m+1}} \\ &\quad \times \cdots \times Q_{U_{m+n}} \times Q_{U_n}))^{a_{m+1}}, Q_{U_{m+1}}, \cdots, \\ &\quad Q_{U_{m+n}}, Q_{U_n}) \\ &= kdf(e((sP)^{m+n+1}, (Q_{U_{m+1}} \times \cdots \times Q_{U_{m+n}} \\ &\quad \times Q_{U_n})s^{m+n+1})^{a_{m+1} \times \cdots \times a_{m+n} \times a_n}, Q_{U_{m+1}}, \\ &\quad \cdots, Q_{U_{m+n}}, Q_{U_n}) \\ &= kdf(e(P^{m+n+1}, (Q_{U_{m+1}} \times \cdots \times Q_{U_{m+n}} \\ &\quad \times Q_{U_n}))^{s^{m+n+1}(a_{m+1} \times \cdots \times a_{m+n} \times a_n)}, \\ &\quad Q_{U_{m+1}}, \cdots, Q_{U_{m+n}}, Q_{U_n}). \end{aligned}$$

For user U_{m+n},

$$\begin{aligned} K &= kdf(K_{m+n}, Q_{U_{m+1}}, ..., Q_{U_{m+n}}, Q_{U_n}) \\ &= kdf(e((V_{m+1} \times \cdots \times V_{m+n-1} \times V_n)P_{pub}, \\ &\quad (W_{m+1} \times \cdots \times W_{m+n-1} \times W_n)SK_{U_{m+n}})^{a_{m+n}}, \\ &\quad Q_{U_{m+1}}, \cdots, Q_{U_{m+n}}, Q_{U_n}) \\ &= kdf(e((a_{m+1}P_{pub} \times \cdots \times a_{m+n-1}P_{pub} \\ &\quad \times a_nP_{pub})P_{pub}, (a_{m+1}sQ_{U_{m+1}} \times \cdots \\ &\quad \times a_{m+n-1}sQ_{U_{m+n-1}} \times a_nsQ_{U_n})sQ_{U_{m+n}})^{a_{m+n}}, \\ &\quad Q_{U_{m+1}}, \cdots, Q_{U_{m+n}}, Q_{U_n}) \\ &= kdf(e((a_{m+1} \times \cdots \times a_{m+n-1} \times a_n)P_{pub}^{m+n+1}, \\ &\quad (a_{m+1} \times \cdots \times a_{m+n-1} \times a_n)s^{m+n+1}(Q_{U_{m+1}} \\ &\quad \times \cdots \times Q_{U_{m+n}} \times Q_{U_n}))^{a_{m+n}}, Q_{U_{m+1}}, \cdots, \\ &\quad Q_{U_{m+n}}, Q_{U_n}) \\ &= kdf(e((sP)^{m+n+1}, (Q_{U_{m+1}} \times \cdots \times Q_{U_{m+n}} \\ &\quad \times Q_{U_n})s^{m+n+1})^{a_{m+1} \times \cdots \times a_{m+n} \times a_n}, Q_{U_{m+1}}, \\ &\quad \cdots, Q_{U_{m+n}}, Q_{U_n}) \\ &= kdf(e(P^{m+n+1}, (Q_{U_{m+1}} \times \cdots \times Q_{U_{m+n}} \\ &\quad \times Q_{U_n}))^{s^{m+n+1}(a_{m+1} \times \cdots \times a_{m+n} \times a_n)}, \\ &\quad Q_{U_{m+1}}, \cdots, Q_{U_{m+n}}, Q_{U_n}). \end{aligned}$$

User n can compute the session key directly.

5 Security Analysis

5.1 Known Key Security

From the randomness of $a_i's$ in our proposed group key agreement protocol, the session keys in different key agreements are independent of each other. The knowledge of the previous session keys does not help an adversary to derive any future session key. Hence our proposed group key agreement protocol provides known key security.

5.2 Forward Secrecy

Even if a long term private key $SK_{ID}(= sQ_{ID})$ of our proposed group key agreement is compromised , the data protected with the previous session key K is still secure because the derivation of K requires the knowledge of previous random values $a_i's$. Therefore our group key agreement protocol has the property of (perfect) forward secrecy.

5.3 Trivial Attack

An attacker may directly try to compute the group key K from the transmitted message $[U_i \rightarrow U_n :$

$(U_i, \{U_1, K_i, M_i, V_i, W_i), i = 1, ..., n-1]$ but due to difficulties of the DLP and onewayness of hash function the trivial attack is not possible in the proposed protocol.

5.4 Key Compromise Impersonation

Suppose that an adversary who know's user $U_1's$ long term private key SQ_{U_1} and want to masquerade with the group leader U_n. Then first he chooses a random value $a_1^{'} \in^R \mathbb{Z}_n$ and calculate

$$K_1 = sQ_{U_1} + a_1^{'}$$

but to verify the correspondence of his guessed random value $a_1^{'}$. He has to compute

$$M_1 = h_1(U_1, U_2, \cdots, U_{n-1}, a_1)$$

which is impossible since he requires the value of a_1 which is the ephemeral private key of the user U_1. Hence U_n will found this un-equality. So this type of attacks are also not possible.

5.5 Unknown Key Share

In our proposed GKA protocol consider the special case (i.e.for $n = 2$), the shared secret $S_{1,2} = S_{2,1} = kdf(e(P^2, (Q_{U_1} \times Q_{U_2}))^{s^{2(a_1 \times a_2)}})$, between U_1 and U_2 involves both members long term private and public keys. This ensures that only U_1 and U_2 who own the corresponding long-term private keys can obtain the same group key and can compute valid key confirmations. Any other entity cannot obtain the same group key. It is impossible that U_1 ends up believing that she/he shares a key with U_2 and although this is in fact the case, while U_2 mistakenly believes that the key is instead shared with another entity E.

5.6 Message Confidentiality

In our proposed scheme the size of shared common group session key is totally depends on the number of users in the current group and their ephemeral private keys. So when a group user want to leave or a new user want to join the group the session key size is obviously change. Also in our proposed scheme in join or leave algorithm the joining and leaving members can not know the number of participant in previous or subsequent group and they also don't know their private keys .

Hence the joining member can not compute previous session keys and leaving member can not compute the subsequent session keys .

6 Comparison

We now compare our protocol with another dynamic group key agreement protocols [18, 19, 35]. We will use the following notations.

1) *Round*: The total number of rounds.
2) Mul: The total number of scalar multiplications and modular multiplication .
3) $Msize$: The maximum number of messages sent by per user.
4) P/E: The total number of pairing computations and modular exponentiations.

Table 1: Setup algorithm -A set of users $U_{[1,...,n]}$

Protocol	Round	M size	Mul	P/E
[18]	$O(n)$	$O(n^2)$	0	$O(n^2)$
[35]	2	$O(n)$	$O(n^2)$	$O(n^2)$
[19]	1	$O(n^2)$	$O(n^2)$	O(n)
Ours	2	$O(5n)$	$O(3n)$	$O(n)$

We observe from Table 1, in our protocol the message size is $O(5n)$ which is linear as compare to [18] and [19]. Similarly, the total number of scalar multiplication is of quadratic order *i.e.* $O(n^2)$ in [35] and [19]. But in our proposed scheme Mul is $O(3n)$ which is again linear. Also in [18] and [35] the pairing computation P/E is again quadratic in order.

We observe from Table 2 the total number of users is $(n+m)$. So in our proposed scheme M size is $O(5(n+m))$, total number of scalar multiplication is $O(3(n+m))$ and the pairing computation is $O(n)$. Hence in join algorithm of our proposed scheme all cases are liner in order as compare to other recent protocols [18, 19, 35].

In Table 3 of leave algorithm the size of the resulting set of users is $(n-m)$. The total number of scalar multiplication in [35], [19] and pairing computation in [35] is of quadratic order $O(n-m)^2$. But in this table, we see that in case of our proposed scheme the M size, total number of scalar multiplication and P/E all are linear in order.

7 Conclusion

With the increasing need of authenticated and secure communication, ID based two round DAGKA protocol is presented here, which resist to all the known attacks. Our protocol also provides forward and backward confidentiality which is the important feature in case of dynamic key agreements. In the last we have given the comparison of our protocol with other recent dynamic group key agreement protocols.

References

[1] R. Barua, R. Dutta, and P. Sarkar, "Extending Joux protocol to multi party key agreement", in *Proceedings of Indocrypt 2003*, LNCS 2904, pp. 205–217, Springer-Verlag, 2003.

Table 2: Join algorithm -A set of users $U_{[n+1,\ldots,n+m]}$ join the set of users $U_{[1,\ldots,n]}$ resulting a set of size $n+m$

Protocol	Round	M size	Mul	P/E
[18]	$O(n+m)$	$O(n+m)^2$	0	$O(m(n+m))$
[35]	2	$O(m)$	$O(m(n+m))$	$O(m(n+m))$
[19]	1	$O(m(n+m))$	$O(m(n+m))$	$O(n+m)$
Ours	2	$O(5(n+m))$	$O(3(n+m))$	$O(n+m)$

Table 3: Leave algorithm -A set of users $U_{[l_1,\ldots,l_m]}$ leave the set of users $U_{[1,\ldots,n]}$ resulting a set of size $n-m$

Protocol	Round	M size	Mul	P/E
[18]	1	$O(n-m)$	0	$O(n-m)$
[35]	2	$O(n-m)$	$O((n-m)^2)$	$O((n-m)^2)$
[19]	0	0	$O((n-m)^2)$	$O(n-m)$
Ours	2	$O(5(n-m))$	$O(3(n-m))$	$O(n-m)$

[2] E. Bresson, O. Chevassut, and D. Pointcheval, "Provably authenticated group diffie-hellman key exchange - The dynamic case", In *Proceedings of Asiacrypt 2001*, LNCS 2248, pp. 290–309, Springer-Verlag, 2001.

[3] D. Bohen and M. Franklin, "Identity based encryption from the Weil pairing", *SIAM Journal of Computing*, vol. 32, no. 3, pp. 586–615, 2001.

[4] M. Burmester and Y. Desmedt, " A secure and efficient conference key distribution system", In *Proceedings of Eurocrypt 1994*, LNCS 950, pp. 275–286, Springer-Verlag, 1995.

[5] Z. Chen, "Security analysis on Nalla-Reddy's ID-based tripartite authenticated key agreement protocol", *Cryptology ePrint Archive*, Report 2003/103, 2003. (`http://eprint.iacr.org/2003/103`)

[6] L. Chen and C. Kudla, "Identity based authenticated key agreement protocols from pairing", in *16th IEEE Security Fudation Wrkshop* , pp. 219—233, 2003.

[7] Q. Cheng and C. Tang, "Cryptanalysis of an ID-based authenticated dynamic group key agreement with optimal round", *Internatioal Journal of Network Security*, vol. 17, no. 6, pp. 678–682, Nov. 2015.

[8] Z. Cheng and L. Chen, "On the security proof of McCullagh-Barreto's key agreement protocol and its variants", *Internatioal Journal of Security and Networks*, Special Issue on Cryptography in Network.

[9] K. Y. Choi, J. Y. Hwang, and D. H. Lee, "Efficient ID-based group key agreement with bilinear maps", in *Proceedings of PKC 2004*, LNCS 2947, pp. 130–144, Springer-Verlag, 2004.

[10] K. Choo, "Revisit of McCullagh-Barreto two party ID-based authentication key agreement protocols", 2019. (`http://eprint.iacr.org/2004/343.pdf`)

[11] N. M. Cullagh and P. Barreto, "A new two-party identity-based authenticated key agreement", Cryptology ePrint Archive, Report, 2004/122, 2004. (`http://eprint.iacr.org/2004/122.pdf`)

[12] W. Diffie and M. Hellman, " New directions in cryptography", *IEEE Transactions on Information Theory*, vol. 6, pp. 644–654, 1976.

[13] R. Dutta and R. Barua, "Overview of key agreement protocols", *IACR Cryptology ePrint Archive*, 2005.

[14] R. Dutta and R. Barua, "Constant round dynamic group key agreement", in *Proceedings of ISC 2005*, 2005. (`http://eprint.iacr.org/2005/221`)

[15] F. Hess, "Efficient identity based signature schemes based on pairings", in *Proceedings of SAC 2002*, LNCS 2595, pp. 310–324, Springer-Verlag, 2002.

[16] I. Ingemarsson, D. Tang, and C. Wang, "A conference key distribution system", *IEEE Transactions on Information Theory*, vol. 28, pp. 714–720, 1982.

[17] A. Joux, "A one round protocol for tripartite Diffie-Hellman", in *Proceedings of ANTS 4*, LNCS 1838, pp. 385–394, 2000.

[18] H. J. Kim, S. M. Lee, and D. H. Lee, "Constant-round authenticated group key exchange for dynamic groups", in *Proceedings of Asiacrypt 2004*, LNCS 3329, pp. 245–259, 2004.

[19] F. Li, D. Xie, W. Gao, J. Yan, and X. A. Wang, "Round-optimal ID-based dynamic authenticated group key agreement", *International Journal of High Performance Systems Architecture*, vol. 6, no. 3, pp. 153, 2016.

[20] H. S. Lee and Y.R.Lee, "Identity based authenticated key agreement from pairings", *Commun. Korean Math.*, Soc. 20, no. 4, pp. 849–859, 2005.

[21] A. Menezes, P. V. Oorschot, and S. Vanstone, *Handbook of Applied Cryptography*, CRC Press, Boca Raton, FL, 1996.

[22] D. Nalla and K. C. Reddy, "Identity based authenticated group key agreement protocol", in *Proceedings of Indocrypt 2002*, LNCS 2551, pp. 215–233, Springer-Verlag, 2002.

[23] D. Nalla and K. C. Reddy, "ID-based tripartite authenticated key agreement protocols from pairings",

Cryptology ePrint Archive, Report 2003/004, 2003. (http://eprint.iacr.org/2003/004)

[24] E. Okamoto, "Proposal for identity-based key distribution system", *Electronics Letters*, vol. 22, pp. 1283–1284, 1986.

[25] R. S. Ranjani, D. L. Bhaskari, and P. S. Avadhani, "An extended ID based authenticated asymmetric group key agreement protocol", *International Journal of Network Security*, vol. 17, no. 5, pp. 510–516, Sept. 2015.

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

[27] M. Scott, "Authenticated ID-based key exchange and remote log-in with insecure token and PIN number", 2002. (http://eprint.iacr.org/2002/164.pdf)

[28] K. Shim, "Cryptanalysis of ID-based tripartite authenticated key agreement protocol", *Cryptology ePrint Archive*, Report 2003/115, 2003. (http://eprint.iacr.org/2003/115)

[29] K. Shim, "Efficient ID-based authenticated key agreement protocol based on the Weil pairing", *Electronics Letters*, vol. 39, no. 8, pp. 653–654, 2003.

[30] N. P. Smart, "Identity-based authenticated key agreement protocol based on Weil pairing", *Cryptology ePrint Archive*, Report 2001/111, 2001. (http://eprint.iacr.org/)

[31] M. Steiner, G. Tsudik, and M. Waidner, "Diffie-Hellman key distribution extended to group communication", in *Proceedings of ACM CCS 1996*, pp. 31-37, 1996.

[32] M. Steiner, G. Tsudik and M.Waidner, "Cliques: A new approach to group key agreement", in *IEEE Conference on Distributed Computing Systems*, pp. 380–380, May 1998.

[33] S. Sun and B. Hsieh, "Security analysis of Shim's authenticated key agreement protocols from pairing", 2003. (http://eprint.iacr.org/2003/113.pdf)

[34] K. Tanaka and E. Okamoto, "Key distribution system for mail systems using ID-related information directory", *Computers and Security*, vol. 10, pp. 25–33, 1991.

[35] J. K. Teng, C. K. Wu and C. M. Tang, "An ID-based authenticated dynamic group key agreement with optimal round", *Science China Information Sciences*, vol. 55, no. 11, pp. 2542–2554, 2012.

[36] S.B. Wilson and A. Menezes, "Unknown key share attacks on the station-to-staion (STS) protocol", in *Proceedings of Second International Workshop on Practice and Theory in Public Key Cryptography (PKC'99)*, LNCS 1560, pp. 154–170, 1999.

[37] F. Zhang, S. Liu, and K. Kim, "ID-based one round authenticated tripartite key agreement protocol with pairings", *Crytology ePrint Archive*, 2002. (http://eprint.iacr.org/2002/122.46)

Biography

Shruti Nathani received the B.Sc., M.Sc. and M.Phil degrees in Mathematics from Pt. Ravishankar Shukla University, Raipur, Chhattisgarh, India in 2008, 2010 and 2011 respectively. She is currently a PhD candidate at the Department of Mathematics in Govt. N.PG. College of Science affiliated from Pt. Ravishankar Shukla University, Raipur, Chhattisgarh, India. Her main research interests include public key cryptography, especially in group oriented cryptography and group key establishment protocols.

B. P. Tripathi, Assistant Professor, Deptt. of Mathematics, Govt. N.PG. college of Science Raipur. The institute is affiliated to Pt. Ravishankar Shukla University, Raipur, Chhattisgarh, India. His field of interest are Non-linear Analysis, Fixed point theory and Public Key Cryptography. He has teaching experience of 24 years of undergraduate and postgraduate classes. He has written 2 books and published 35 research papers in various National and International journals. Two scholar has awarded Ph.D. degree and recently four scholars are pursuing their research work under the supervision of Dr. Tripathi.

Shaheena Khatoon received the B.Sc., M.Sc. and M.Phil degrees in Mathematics from Pt. Ravishankar Shukla University, Raipur, Chhattisgarh, India in 2005, 2007 and 2009 respectively. She joined School of studies in Mathematics, Pt. Ravi Shankar Shukla University, Raipur, Chhattisgarh, India for her research work.

Unidirectional FHPRE Scheme from Lattice for Cloud Computing

Juyan Li[1,2], Chunguang Ma[1,3], Lei Zhang[1,4], and Qi Yuan[1,5]
(Corresponding author: Chunguang Ma)

College of Computer Science and Technology, Harbin Engineering University[1]
Harbin 150001, P.R. China
College of Data Science and Technology, Heilongjiang University[2]
Harbin 150080, P.R. China
State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences[3]
Beijing 100093, P.R. China
College of Information and Electronic Technology, Jiamusi University[4]
Jiamusi 154007, P.R. China
College of Communication and Electronic Engineering, Qiqihar University[5]
Qiqihar, 161006, P.R. China
(Email: machunguang@hrbeu.edu.cn)
(Received Dec. 20, 2017; Revised and Accepted Apr. 21, 2018; First Online Feb. 26, 2019)

Abstract

With the emerging of new types of network forms, services and cloud computation, the situation has transformed from one party to many parties at least one of both communication ends, that is "one-to-many," "many-to-one," and "many-to-many" situations. Most of the existing fully homomorphic encryption schemes only allow one party to encrypt the plaintext and another party to decrypt the ciphertext without the decryption keys. This form of cryptography loses efficiency under the demands of "one-to-many," "many-to-one," and "many-to-many" scenarios. In this paper, we combine the fully homomorphic encryption with proxy re-encryption to propose the fully homomorphic proxy re-encryption scheme which can be applied to "many-to-one" scenario, that is the fully homomorphic proxy re-encryption scheme allows one party to compute arbitrary functions over encrypted data for many parties without the decryption keys. Finally, IND-CPA, KP-CPA and master secret security proof of our proposal are given.

Keywords: FHPRE; Key Privacy; Many-to-One; STP-Binary-LWE

1 Introduction

Proxy Re-Encryption (PRE), which is an extension of public key encryption, was introduced by Bleumer *et al.* at Eurocrypt 1998 [4]. A PRE scheme allows proxy (semi trusted) to transform a ciphertext for Alice (delegator) into a ciphertext for Bob (delegatee) without knowing the message. The interesting property makes PRE more applicable in many scenarios, such as encrypted email forwarding [4], vehicular ad hoc network, outsourced filtering of encrypted spam , the distributed file system [3, 9]. Fully-homomorphic encryption (FHE) marks another milestone in the history of modern cryptography. A FHE scheme allows one party to compute arbitrary functions over encrypted data for another party without the decryption key. FHE has many applications in cloud computation, such as private queries to a search engine, searching on encrypted data [8, 10, 14].

The existing FHE schemes are mostly in the form of "one-to-one" deployment situations. With the emerging of new types of network forms, services and cloud computation, the situation has transformed from one party to many parties at least one of both communication ends, that is "one-to-many," "many-to-one," and "many-to-many" situations. It's interesting to combine the concept of FHE and PRE to construct a fully homomorphic proxy re-encryption (FHPRE), which allows one party to compute arbitrary functions over encrypted data for many parties without the decryption keys, satisfying the many-to-one situation. The application of FHPRE in the cloud computation can see [13, 21, 24].

Xagawa [22] constructed the first bidirectional PRE scheme based on lattices, which is CPA secure. Aono *et al.* [1] proposed a unidirectional key-private PRE (KP-PRE) scheme based on lattices, which is CPA secure. A unidirectional scheme permits user Alice to delegate to user Bob, without permitting Alice to decrypt user

Bob's ciphertexts. A unidirectional proxy re-encryption is said to be key privacy if any adversary cannot distinguish a real re-encryption key from a random re-encryption key even if the adversary is allowed to access to the re-encryption key oracle and the re-encryption oracle which re-encrypts input ciphertexts by using the real re-encryption key [2, 18]. Ateniese *et al.* [3] introduced master secret security as another security requirement for unidirectional PRE based on lattices. Master secret security demands that it is hard for the coalition of the proxy and Bob to compute Alice's secret key.

Singh *et al.* [20] showed [1,22] is not secure under master secret security model and constructed a unidirectional multi-use PRE which is secure under master secret security model. Nishimak *et al.* [18] proposed two unidirectional KP-PRE schemes from LWE assumptions, which are CPA secure. Jiang *et al.* [11] constructed a multi-use unidirectional PRE scheme based on lattices, which is CPA secure and master secret secure. Kirshanova *et al.* [12] proposed a unidirectional proxy re-encryption scheme based on LWE problem and showed it is CCA-1 secure in the selective model. Zhang *et al.* [23] proposed Unidirectional IBPRE scheme from lattice for cloud computation, which is CPA secure.

Recently, FHE from learning with errors (LWE) assumption has attracted many attentions due to their average-case to worst-case equivalence and their conjectured resistance to quantum attacks [19]. The efficiency of FHE is one of the most concerned problems. A number of techniques are proposed and used to improve the efficiency of FHE, such as re-linearization technique, dimension modulus reduction technique [5], modulus switching technique [6]. In 2012, Brakerski [7] constructed a scale-invariant fully homomorphic encryption scheme, whose noise only grows linearly with every multiplication (before refreshing). Ma *et al.* [15] proved that STP-binary-LWE is hard when LWE is hard, and modified the scale-invariant fully homomorphic encryption scheme [7] based on STP-Binary-LWE so that it is more efficient. Furthermore, Ma *et al.* [15] can encrypt several messages at a time and achieve a balance between security and efficiency in the hierarchical encryption systems.

Unfortunately, all of the above FHE schemes are not applicable to the many-to-one situation. Zhong *et al.* [24] constructed a "many-to-one" homomorphic encryption scheme based on approximate GCD problem, which is not lattice-based scheme. The essence of the scheme [24] is a PRE scheme, and needs the trusted third party to distribute the key. Ma *et al.* [16,17] constructed a homomorphic proxy re-encryption scheme based on LWE which can only encrypt one message at a time.

In this paper, we construct a unidirectional FHPRE scheme from lattices which can be used in the "many-to-one" situation and only needs semi trusted third party. The FHPRE can encrypt two messages at a time. At last, we prove that our FHPRE is indistinguishable against chosen-plaintext attacks, and key privacy secure.

The rest of this paper is organized as follows. Section 2 is preliminaries. Section 3 describes the constructed FHPRE scheme and proves the security of FHPRE. At last, the conclusion will be given in Section 4.

2 Preliminaries

2.1 Notation

All scalars, column vectors and matrices will be denoted in the form of plain (*e.g.* x), bold lowercase (*e.g.* $\vec{x}$) and uppercase (*e.g.* X), respectively. For a real number x ($x \geq 0$), $\lceil x \rceil$, $\lfloor x \rfloor$, $\lfloor x \rceil$ denoted rounding up or down, rounding to the nearest integer. We denote $\eta = \lceil \log q \rceil$, $[x]_q = x \bmod q$, $\mathbb{Z}_q = (-\frac{q}{2}, \frac{q}{2}] \cap \mathbb{Z}$, $[k] = \{1, 2, \cdots, k\}$. The l_i norm of a vector $\vec{v}$ is denoted by $||\vec{v}||_i$. k-dimensional identity matrix is denoted by I_k. Inner product, tensor product and semitensor product are denoted by $< \vec{v}, \vec{u} >$, $P \otimes Q$, $P_{r \times kl} \ltimes Q_{l \times t} = (P(Q \otimes I_k))_{r \times kt}$, respectively.

$[X|Y] \in \mathbb{Z}_q^{m \times (n+l)}$ is the concatenation of the columns of $X \in \mathbb{Z}_q^{m \times n}, Y \in \mathbb{Z}_q^{m \times l}$. $[X;Y] \in \mathbb{Z}_q^{(n+l) \times m}$ is the concatenation of the rows of $X \in \mathbb{Z}_q^{n \times m}, Y \in \mathbb{Z}_q^{l \times m}$. We set

$$\begin{aligned} BD\left(\vec{x}^T\right) &= \left(\vec{u}_1^T | \cdots | \vec{u}_\eta^T\right) \in \{0,1\}^{n\eta}; \\ P2\left(\vec{x}\right) &= \left(1, 2, \cdots, 2^{\eta-1}\right)^T \otimes \vec{x} \\ &= \left(1\vec{x}; 2\vec{x}; \cdots; 2^{\eta-1}\vec{x}\right)^T \in \mathbb{Z}_q^{n\eta}, \end{aligned}$$

where $\vec{x} \in \mathbb{Z}_q^n$, $\vec{x}^T = \sum_{k=1}^{\eta} 2^{k-1} \vec{u}_k^T$. When A is a matrix, let $P2(A), BD(A)$ be the matrix formed by applying the operation to each column of A.

Concerning a probability distribution D, we record it as $\vec{x} \leftarrow D$, which means that $\vec{x}$ is sampled according to D. So for a set S, we record it as $y \leftarrow S$, which means that y is sampled uniformly from S. Two random variables X and Y are said to be statistically (and computationally) indistinguishable, denoted by $X \approx_s Y$ ($X \approx_c Y$).

2.2 $STP - Binary - LWE_{n,q,\chi^k}$ and Key Switching

Ma *et al.* [15] proved that STP-binary-LWE is hard and showed the Key Switching functions by semitensor product.

Theorem 1. *([15]) For an integer $q = q(n) \geq 2$ and a distribution χ on $\mathbb{Z}_q$, an integer dimension $n = n' log(logn') \in \mathbb{Z}^+$, where n' is the dimension of LWE problem. The $STP-Binary-LWE_{n,q,\chi^k}$ problem, which is to distinguish the following two distributions: In the first distribution, one samples $(\vec{a}; b_1, \cdots, b_k)$ uniformly from $\mathbb{Z}_q^{n+k}$. In the second distribution, one first draws $\vec{s} \leftarrow \mathbb{Z}_2^{n/k}$ and then samples $(\vec{a}, b_1, \cdots, b_k) \in \mathbb{Z}_q^{n+k}$ by independently sampling $\vec{a} \leftarrow \mathbb{Z}_q^n$, $e_i \leftarrow \chi, i \in [k]$, and setting $(b_1, \cdots, b_k) = \vec{a}^T \ltimes \vec{s} + (e_1, \cdots, e_k)$, is hard.*

In the following, we can without loss of generality let that $k = 2$. We show the Key Switching functions which

can switch ciphertexts under S into ciphertexts under $(1;\vec{t})$. Let q be an integer and χ be a distribution over $\mathbb{Z}$.

- SwitchKeyGen$_q(S,\vec{t})$: Input $S \in \mathbb{Z}^{n_s\times 2}$, $\vec{t} \in \mathbb{Z}^{\frac{n_t}{2}}$, $A_{s:t} \leftarrow \mathbb{Z}_q^{\hat{n}_s\times n_t}$ and $X \leftarrow \chi^{\hat{n}_s\times 2}$, where $\hat{n}_s = n_s \cdot \lceil logq \rceil$. Output $P_{s:t} = [B_{s:t}||-A_{s:t}] \in \mathbb{Z}_q^{\hat{n}_s\times(n_t+2)}$, where $B_{s:t} := [A_{s:t} \ltimes \vec{t} + X_{s:t} + PowersOf2_q(S)]_q \in \mathbb{Z}_q^{\hat{n}_s\times 2}$.
- SwitchKey$_q(P_{s:t},\vec{c_s})$: Input $P_{s:t}$ and ciphertext $\vec{c_s}$ under S. Output ciphertext $\vec{c_t} := [P_{s:t}^T \cdot BitDecomp_q(\vec{c_s})]_q$ under $(1;\vec{t})$.

Lemma 1. *([15]) (correctness). Let $S \in \mathbb{Z}^{n_s\times 2}, \vec{t} \in \mathbb{Z}^{n_t/2}$ and $\vec{c_s} \in \mathbb{Z}_q^{n_s}$ be any vectors. Let $P_{s:t} \leftarrow SwitchKeyGen_q(S,\vec{t})$ and set $\vec{c_t} \leftarrow SwitchKey_q(P_{s:t},\vec{c_s})$. Then*

$$\vec{c_s}^T \ltimes S = \vec{c_t} \ltimes (1;\vec{t}) - BitDecomp_q(\vec{c_s})^T X_{s:t}(modq).$$

Lemma 2. *([15]) (security). Let $S \in Z^{n_s\times 2}$ be any vector, $\vec{t} \leftarrow \mathbb{Z}^{n_t/2}$, $P_{s:t} \leftarrow SwitchKeyGen(S,\vec{t})$, then P is computationally indistinguishable from uniform over$\mathbb{Z}_q^{\hat{n}_s\times(n_t+2)}$, assuming STP-Binary-DLWE$_{n,q,\chi^k}$.*

2.3 Syntax of FHPRE and Security Model

The FHPRE compromises FHE and PRE, the Syntax of FHPRE is as follows.

Definition 1. *(Unidirectional FHPRE Scheme)*
A single-hop unidirectional FHPRE scheme consists of the following 7 algorithms:

1) *Setup$(1^k,1^L)\rightarrow pp$: Given the security parameter k, the upper bound on the maximal multiplicative depth $L \in \mathbb{N}$ that the scheme can homomorphically evaluate, output the public parameters pp.*
2) *Gen$(pp,i,L) \rightarrow (ek^i,dk^i,evk^i)$: Given pp, L and a user identity i, output an encryption/decryption key pair (ek^i,dk^i), eval keys $evk^i = \{evk^i_{(l-1):l}\}_{l\in[L]}$, and decryption keys dk_l^i at level l of the circuit, $l \in [L]$.*
3) *Enc$(pp,ek^i,\mu)\rightarrow ct$: Given pp, ek^i and a message μ , output a ciphertext ct_0^i at level 0 of the circuit.*
4) *Eval$(pp,evk^i_{(l-1):l},c^i_{l-1,1},c^i_{l-1,2}) \rightarrow c_l^i$: Given $pp, evk^i_{(l-1):l}$, and ciphertexts $c^i_{l-1,1},c^i_{l-1,2}$ at level $l-1$ of the circuit, output a ciphertext c_l^i at level l of the circuit, $l \in [L]$.*
5) *Dec$(pp,dk^i,ct_L^i)\rightarrow \mu$: Given dk^i and ct_L^i at level L of the circuit, output a plaintext μ or an error symbol $\perp$.*
6) *Rekey $(pp,dk_l^i,\mathrm{ek}^j) \rightarrow rk_{l\rightarrow 0}^{i\rightarrow j}$: Given a decryption key dk_l^i of user i at level l of the circuit and ek^j of user j, output a re-encryption key $rk_{l\rightarrow 0}^{i\rightarrow j}$, $l = 0,1,\cdots,L$.*
7) *ReEnc$\left(pp,rk_{l\rightarrow 0}^{i\rightarrow j},ct_l^i\right) \rightarrow ct_0^j$: Given the re-encryption key $rk_{l\rightarrow 0}^{i\rightarrow j}$ and ct_l^i for the user i at level l of the circuit, output a ciphertext ct_0^j for the user j at level 0 of the circuit.*

Correctness: Three requirements are needed:

$$\begin{aligned} Dec\left(pp,dk_l^i,ct_l^i\right) &= \mu;\\ Dec\left(pp,dk^i,ct_L^i\right) &= \mu;\\ \mathrm{Dec}\left(pp,dk_i^j,ReEnc\left(pp,rk_{l\rightarrow 0}^{i\rightarrow j},ct_l^i\right)\right) &= \mu, \end{aligned}$$

where $l \in [L]$. Now we define the security model of an FHPRE scheme.

Definition 2. *(IND-CPA security) Let UniFH-PRE=(Setup, Gen, Enc, Eval, Dec, ReKey, ReEnc) be a single-hop, unidirectional PRE Scheme, k a security parameter. Suppose that there exists a PPT algorithm RandEnc which takes pp as input and outputs a random ciphertext at output side. Let $H = H(k)$ and $C = C(k)$ be polynomials of k, which stands for the number of honest users and corrupted users, respectively. Consider the following game, denoted by $Expt^{IND-CPA}_{A,UniFHPRE}(k)$, between challenger and adversary.*

Initialization: *Given security parameter k and coin $b \in \{0,1\}$, run $pp \leftarrow Setup(1^k,1^L)$. Initialize $CU \leftarrow \{H+1,\cdots,H+C\}$, which denote the set of corrupted users. For $i = 0,\cdots,H+C$, generate key pairs $(ek^i,dk^i,evk^i) \leftarrow \mathrm{Gen}\,(\mathrm{pp},1^i,1^L)$. Run the adversary on input pp, key pairs of corrupted users $\{(ek^i,dk^i,evk^i)\}_{i=H+1,\cdots,H+C}$, and public keys of honest users $\{(ek^i,evk^i)\}_{i=0,\cdots,H}$.*

Learning Phase: *For $\forall l \in [L]\cup\{0\}$, the adversary could issue queries to the following oracles in any order and many times:*

Oracle REKEY receives two indices $i,j \in \{0,1,\cdots,H+C\}$. If $i = j$ then it returns $\perp$; if $(i=0)\cap(j\in CU)$ then the oracle returns $\perp$; otherwise, returns $rk_{l\rightarrow 0}^{i\rightarrow j} \leftarrow Rekey\,(pp,dk_l^i,\mathrm{ek}^j)$.

Oracle REENC receives two indices $i,j \in \{0,1,\cdots,H+C\}$ and ciphertext ct_l^i. If $i = j$ then returns $\perp$; if $(i=0)\cap(j\in CU)$ then the oracle returns $\perp$; otherwise, it queries (i,j) to REKEY, obtains $rk_{l\rightarrow 0}^{i\rightarrow j}$, and returns $ct_0^j \leftarrow ReEnc\left(pp,rk_{l\rightarrow 0}^{i\rightarrow j},ct_l^i\right)$.

Oracle CHALLENGE, which can be queried only once, receives μ. If $(b=0)$, it returns $ct \leftarrow RandEnc(pp)$. If $(b=1)$, it returns $ct \leftarrow Enc(pp,ek^0,\mu)$.

Eventually. The adversary halts after it and outputs its decision $b' \in \{0,1\}$.

Finalization: *Output 1 if $b' = b$. Otherwise, output 0.*

We define the advantage of the adversary as

$$Adv^{Ind-CPA}_{A,UniFHPRE}(k) = \left| \begin{array}{l} \Pr\left[Expt^{Ind-CPA}_{A,UniFHPRE}(k) \to 1 \,|b=1\right] \\ -\Pr\left[Expt^{Ind-CPA}_{A,UniFHPRE}(k) \to 1 \,|b=0\right] \end{array} \right|$$

We say that UniFHPRE is IND-CPA secure if $Adv^{Ind-CPA}_{A,UnFHiPRE}(\cdot)$ *is negligible for every PPT adversary.*

Definition 3. *(KP-CPA security) Let UniFHPRE=(Setup, Gen, Enc, Eval, Dec, ReKey, ReEnc) be a single-hop, unidirectional FHPRE Scheme, k a security parameter. Suppose that there exists a PPT algorithm RandRekey which takes pp as input and outputs a random re-encryption key rk. Let* $H = H(k)$ *and* $C = C(k)$ *be polynomials of k, which stands for the number of honest users and corrupted users, respectively. Consider the following game, denoted by* $Expt^{KP-CPA}_{A,UniFHPRE}(k)$*, between challenger and adversary.*

Initialization: *Given security parameter k and coin* $b \in \{0,1\}$*, run* $pp \leftarrow Setup\left(1^k, 1^L\right)$*. Initialize* $T \leftarrow \phi$ *which is a table containing the re-encryption keys and shared among oracles. For* $i = -1, 0, \cdots, H+C$*, generate key pairs* $\left(ek^i, dk^i, evk^i\right) \leftarrow \text{Gen}\left(\text{pp}, 1^i, 1^L\right)$*. Run adversary with pp, the public keys and eval keys of honest users* $\{(ek^i, evk^i)\}_{i=0,\cdots,H}$*, the key pairs of corrupted users* $\{(ek^i, dk^i, evk^i)\}_{i=H+1,\cdots,H+C}$*.*

Learning Phase: *For* $\forall l \in L$*, adversary could issue queries to the following oracles in any order and many times except for the constraint in oracle CHALLENGE.*

Oracle REKEY receives two indices $i, j \in \{-1, 0, \cdots, H+C\}$*. If* $i = j$ *then it returns* $\perp$*; if* $(i,j) = (0,-1)$*, then it returns* $\perp$*; if there already exists the re-encryption key from user i at level l of the circuit to user j,* i.e. $\left(i, l, j, rk^{i\to j}_{l\to 0}\right) \in T$*, then it returns* $rk^{i\to j}_{l\to 0}$*, otherwise, it generates* $rk^{i\to j}_{l\to 0} \leftarrow \text{Rekey}\left(pp, dk^i_l, \text{ek}^j\right)$*, updates* $T \leftarrow T \cup \left\{\left(i, l, j, rk^{i\to j}_{l\to 0}\right)\right\}$*, and returns* $rk^{i\to j}_{l\to 0}$*.*

Oracle REENC receives two indices $i, j \in \{-1, 0, \cdots, H+C\}$ *and a ciphertext* ct^i_l*. if* $i = j$ *then it returns* $\perp$*; if there exists no re-encryption key* $rk^{i\to j}_{l\to 0}$ *in the table T, it generates* $rk^{i\to j}_{l\to 0} \leftarrow Rekey\left(pp, dk^i_l, ek^j\right)$*, and updates* $T \leftarrow T \cup \left\{\left(i, j, rk^{i\to j}_{l\to 0}\right)\right\}$*, it finally returns* $ct^j_0 \leftarrow ReEnc\left(pp, rk^{i\to j}_{l\to 0}, ct^i_l\right)$*.*

Oracle CHALLENGE can be queried only once. On the query, the oracle searches the table T for $\left(0, l, -1, rk^{0\to -1}_{l\to 0}\right)$*, if such key does not exist, it generates* $rk^{0\to -1}_{l\to 0} \leftarrow \text{ReKey}\left(pp, dk^0_l, ek^{-1}\right)$ *and updates* $T \leftarrow T \cup \left\{\left(0, l, -1, rk^{0\to -1}_{l\to 0}\right)\right\}$*. If* $b = 0$ *then it returns a random re-encryption key* $rk \leftarrow FakeReKey\,(pp)$*, which is not contained in T. If* $b = 1$*, then it returns the real re-encryption key* $rk^{0\to -1}_{l\to 0}$ *contained in T.*

Eventually. Adversary halts after it outputs its decision $b' \in \{0,1\}$*.*

Finalization: *Output 1 if* $b' = b$*. Otherwise, output 0.*

The advantage of Adversary is

$$Adv^{KP-CPA}_{A,UniFHPRE}(k) = \left| \begin{array}{l} \Pr\left[Expt^{KP-CPA}_{A,UniFHPRE}(k) \to 1 \,|b=1\right] \\ -\Pr\left[Expt^{KP-CPA}_{A,UniFHPRE}(k) \to 1 \,|b=0\right] \end{array} \right|$$

We say that UniFHPRE is KP-CPA secure if $Adv^{KP-CPA}_{A,UniFHPRE}(\cdot)$ *is negligible for every polynomial-time adversary.*

3 Unidirectional FHPRE Scheme

In this section, we constructed a single-hop unidirectional FHPRE scheme based on [15] and proved the scheme is IND-CPA and KP-CPA security.

3.1 Our Construction

A single-hop unidirectional FHPRE scheme consists of the following 7 algorithms.

1) Setup$(1^k, 1^L)$: Sample $A \leftarrow \mathbb{Z}_q^{N\times n}$, where $N \triangleq (n+2)\cdot(logq + O(1))$, $n = n' log(logn') \in \mathbb{Z}^+$, n' is the dimension of LWE problem. Output $pp = \left(1^k, 1^n, q, \chi, L, A\right)$.

2) Gen(pp, i): Sample $s^i_l, t^i_l \leftarrow \mathbb{Z}_2^{n/2}$, $l = 0, 1, \cdots, L$, and compute $B^i_0 = [A \ltimes \vec{s}^i_0 + X^i_0]_q$, where $X^i_0 \leftarrow \chi^{N\times 2}$. Let $P^i_0 = [B^i_0 \parallel -A] \in \mathbb{Z}_q^{N\times(n+2)}$. For $\forall l \in [L]$, define

$$\widetilde{S}^i_{l-1} = (\alpha || \beta) \in \mathbb{Z}_2^{(n+2)^2\lceil logq\rceil^2 \times 2},$$

where

$$\alpha = BD((1; \vec{s}^i_{l-1}) \otimes (1;0)) \otimes BD((1; \vec{s}^i_{l-1}) \otimes (1;0)),$$

$$\beta = BD((1; \vec{s}^i_{l-1}) \otimes (0;1)) \otimes BD((1; \vec{s}^i_{l-1}) \otimes (0;1)),$$

and compute $P^i_{(l-1):l} \leftarrow SwitchKeyGen(\widetilde{S}^i_{l-1}, \vec{s}^i_{l-1})$. Output

$$\begin{aligned} \left(ek^i, dk^i\right) &= \left(P^i_0, \vec{s}^i_L\right) \\ dk^i_l &= \vec{s}^i_l, l \in [L] \\ evk^i &= \{evk^i_{(l-1):l}\}_{l\in[L]} \\ &= \{P^i_{(l-1):l}\}_{l\in[L]}. \end{aligned}$$

3) Enc$(pp, ek^i = P^i_0, (m_1, m_2))$: Compute

$$\vec{c}^i_0 = \left[P^{i\,T}_0 \cdot \vec{r} + \left\lfloor \frac{q}{2} \right\rfloor \vec{m}\right]_q \in \mathbb{Z}_q^{(n+2)},$$

where $\vec{r} \leftarrow \{0,1\}^N$, $\vec{m} = (m_1, m_2, 0\cdots, 0)^T \in \mathbb{Z}_2^{(n+2)}$. Output $ct^i_0 = \vec{c}^i_0$.

4) Eval($\bullet$): Suppose the homomorphic addition and multiplication over GF(2) be enable to evaluate depth L arithmetic circuits in a gate-by-gate manner. For any $i \in [L]$, a gate at level i of the circuit is that the operand ciphertexts can be decrypted using $\vec{s}_{i-1}$, and the output of the homomorphic operation can be decrypted using $\vec{s}_i$.

- Add$(pp, evk^i_{(l-1):l}, c^i_{l-1,1}, c^i_{l-1,2})$: Input ciphertexts $c^i_{l-1,1} = \vec{c}^i_{l-1,1}$, $c^i_{l-1,2} = \vec{c}^i_{l-1,2}$ under secret key $\vec{s}^i_{l-1}$, and compute

$$\vec{c}^i_{l-1,add} = P2(\vec{c}^i_{l-1,1} + \vec{c}^i_{l-1,2}) \otimes P2(1, 1, 0, \cdots, 0),$$

$$\vec{c}^i_{l,add} \leftarrow Switchkey(P^i_{(l-1):l}, \vec{c}^i_{l-1,add}) \in \mathbb{Z}_q^{n+2}.$$

Output $c^i_{add,l} = \vec{c}^i_{l,add}$.

- Mult$(pp, evk^i_{(l-1):l}, c^i_{l-1,1}, c^i_{l-1,2})$: Input ciphertexts $c^i_{l-1,1} = \vec{c}^i_{l-1,1}, c^i_{l-1,2} = \vec{c}^i_{l-1,2}$ under secret key $\vec{s}^i_{l-1}$, and compute

$$\tilde{\vec{c}}^j_{l-1,mult} = \lfloor \frac{2}{q}(P2(\vec{c}_1) \otimes P2(\vec{c}_2)) \rceil,$$

$$\vec{c}^i_{l,mult} \leftarrow SwitchKey(P^i_{(l-1):l}, \tilde{\vec{c}}^j_{l-1,mult}) \in \mathbb{Z}_q^{n+2}.$$

Output $c^i_{mult,l} = \vec{c}^i_{l,mult}$.

5) Dec$\left(pp, dk^i = \vec{t}^i_L, ct^i_L = \vec{c}^i_L\right)$: Input ciphertext ct^i_L under secret key $dk^i (= \vec{s}^i_L)$ and $\vec{s}^i_L$. Output

$$(m_1, m_2) = \left[\left\lfloor 2 \cdot \frac{\left[c_L^T \ltimes (1; \vec{s}_L)\right]_q}{q} \right\rceil\right]_2$$

6) Rekey$\left(pp, dk^i_{l-1} = \vec{s}^i_{l-1}, \mathrm{ek}^j = P^j_0\right)$: Compute

$$\begin{aligned} M^{i\to j}_{l\to 0} &\in \mathbb{Z}_q^{(n+2)\lceil \log q \rceil \times (n+2)} \\ &\leftarrow R^{i\to j}_{l\to 0} P^j_0 + P2\left((1; \vec{s}^i_l) \otimes I_2 || 0\right) \\ N^j_0 &\in Z_q^{N\times(n+2)} \leftarrow R^j_0 P^j_0, \end{aligned}$$

where $0 \in \{0\}^{(n+2)\times n}$, $R^{i\to j}_{l\to 0} \in Z_2^{(n+2)\lceil \log q \rceil \times N}$, $R^j_0 \in Z_2^{N\times N}$. Output $rk^{i\to j}_{l\to 0} = (M^{i\to j}_{l\to 0}, N^j_0)$.

7) ReEnc$\left(pp, rk^{i\to j}_{l\to 0} = (M^{i\to j}_{l\to 0}, N^j_0), ct^i_l = \vec{c}^i_l\right)$: Output

$$ct^j_0 = \vec{c}^j_0 = SwitchKey_q\left(M^{i\to j}_{l\to 0}, \vec{c}^i_l\right) + {N^j_0}^T \vec{r}^j_0,$$

where $\vec{r}^j_0 \in \mathbb{Z}_2^N$.

We show the correctness of the FHPRE scheme below.

Lemma 3. *([15]) Let $\vec{s} \in Z_2^{n/2}$, $\vec{c} \in Z_q^{n+2}$ be such that $\vec{c}^T \ltimes (1, \vec{s}) = \lfloor \frac{q}{2} \rfloor \cdot (m_1, m_2) + X(modq)$, where $m_1, m_2 \in \{0, 1\}$ and $||X||_\infty \le \lfloor \frac{q}{2} \rfloor /2$. Then $Dec(\vec{c}) = (m_1, m_2)$.*

proposition 1. *Let $q, n, |\chi| \le B, L$ be parameters for FHPRE, and let ciphertexts $c^i_l = \vec{c}^i_l$ and secret key $\vec{s}^i_l$ be such that*

$$\vec{c}^{iT}_l \ltimes (1; \vec{s}^i_l) = \left\lfloor \frac{q}{2} \right\rfloor (m_1, m_2) + X^i_l (modq),$$

where $m_1, m_2 \in \{0, 1\}$ and $||X^i_l||_\infty \le E < \lfloor \frac{q}{2} \rfloor /2$. Define $\vec{c}^j_0 \leftarrow ReEnc\left(pp, rk^{i\to j}_{l\to 0}, \vec{c}^i_l\right)$. Then

$$\vec{c}^{jT}_0 \ltimes \left(1; \vec{s}^j_0\right) = \left\lfloor \frac{q}{2} \right\rfloor (m_1, m_2) + X(modq),$$

where $||X||_\infty \le E + N(n+2)\lceil \log q \rceil B^2 + N^2 B$.

Proof. Suppose $\vec{c}^{iT}_l \ltimes (1; \vec{s}^i_l) = \lfloor \frac{q}{2} \rfloor (m_1, m_2) + X^i_l (modq)$, where $|X^i_l||_\infty \le E < \lfloor \frac{q}{2} \rfloor /2$. To decrypt the re-encrypted ciphertext $ct^j_0 = \vec{c}^j_0 = SwitchKey\left(M^{i\to j}_{l\to 0}, \vec{c}^i_l\right) + {N^j_0}^T \vec{r}^j_0$ with $\left(1; \vec{s}^j_0\right)$, where $\vec{r}^j_0 \in \mathbb{Z}_2^N$, $M^{i\to j}_{l\to 0} = R^{i\to j}_{l\to 0} Q^j_0 + P2\left((1; s^i_l) \otimes I_2 || 0\right)$, $R^{i\to j}_{l\to 0} \in \mathbb{Z}_2^{(n+2)\lceil \log q \rceil \times N}$, $N^j_0 = R^j_0 Q^j_0$, $R^j_0 \in \mathbb{Z}_2^{N\times N}$, one computes

$$\begin{aligned} & \vec{c}^{jT}_0 \ltimes \left(1; \vec{s}^j_0\right) \\ = \; & SwitchKey\left(M^{i\to j}_{l\to 0}, \vec{c}^i_l\right)^T \ltimes \left(1; \vec{s}^j_0\right) \\ & + {N^j_0}^T \vec{r}^j_0 \ltimes \left(1; \vec{s}^j_0\right) (\mathrm{mod} q) \\ = \; & BD(\vec{c}^i_l)^T R^{i\to j}_{l\to 0} Q^j_0 \ltimes \left(1; \vec{s}^j_0\right) \\ & + BD(\vec{c}^i_l)^T P2\left((1; \vec{s}^i_l) \otimes I_2 || 0\right) \ltimes \left(1; \vec{s}^j_0\right) \\ & + \vec{r}^{jT}_0 R^j_0 Q^j_0 \ltimes \left(1; \vec{s}^j_0\right) (\mathrm{mod} q) \\ = \; & \left\lfloor \frac{q}{2} \right\rfloor (m_1, m_2) + X^i_l + BD(\vec{c}^i_l)^T R^{i\to j}_{l\to 0} Y^j_0 \\ & + \vec{r}^{jT}_0 R^j_0 Y^j_0 (\mathrm{mod} q). \end{aligned}$$

Let $X = X^i_l + BD(\vec{c}^i_l)^T R^{i\to j}_{l\to 0} Y^j_0 + \vec{r}^{jT}_0 R^j_0 Y^j_0$, we have

$$\begin{aligned} & \left\| X^i_l + BD(\vec{c}^i_l)^T R^{i\to j}_{l\to 0} Y^j_0 + \vec{r}^{jT}_0 R^j_0 Y^j_0 \right\|_\infty \\ \le \; & \left\| X^i_l \right\|_\infty + \left\| BD(\vec{c}^i_l)^T R^{i\to j}_{l\to 0} Y^j_0 \right\|_\infty + \left\| \vec{r}^{jT}_0 R^j_0 Y^j_0 \right\|_\infty \\ < \; & E + N(n+2)\lceil \log q \rceil B^2 + N^2 B. \end{aligned}$$

□

Lemma 4. *([15]) Let $q, n, |\chi| \le B, L$ be parameters for FHPRE, and let $(pk, evk, dk) \leftarrow Gen\ (1^L, 1^n)$. Let $\vec{c}_1, \vec{c}_2$ be such that*

$$\vec{c}^T_1 \ltimes (1, \vec{s}_{i-1}) = \left\lfloor \frac{q}{2} \right\rfloor (m_1, m_2) + X_1 (\mathrm{mod}\ q),$$

$$\vec{c}^T_2 \ltimes (1, \vec{s}_{i-1}) = \left\lfloor \frac{q}{2} \right\rfloor (m'_1, m'_2) + X_2 (\mathrm{mod}\ q),$$

with $||X_1||_\infty, ||X_2||_\infty \le E \le \lfloor \frac{q}{2} \rfloor /2$. Define

$$\vec{c}_{add} \leftarrow HE.Add_{evk}(\vec{c}_1, \vec{c}_2),$$

$$\vec{c}_{mult} \leftarrow HE.Mult_{evk}(\vec{c}_1, \vec{c}_2).$$

Then

$$\vec{c}_{add}^{T} \ltimes (1, \vec{s}_i) = \left\lfloor \frac{q}{2} \right\rfloor [(m_1 + m_1^{'}, m_2 + m_2^{'})]_2 + X_{add} (\bmod\ q),$$

$$\vec{c}_{mult}^{T} \ltimes (1, \vec{s}_i) = \left\lfloor \frac{q}{2} \right\rfloor (m_1 m_1^{'}, m_2 m_2^{'}) + X_{mult} (\bmod\ q),$$

where $||X_{add}||_\infty, ||X_{mult}||_\infty \leq O(n) \cdot max\{E, nlog^3 q \cdot B\}$.

Theorem 2. *([15]) The scheme HE with parameters* $n, q, |\chi| \leq B, L$ *for which* $q/B \geq (O(n))^{L+O(1)}$, *is L-homomorphic.*

3.2 Security

We show the security of the FHPRE scheme in this section which includes IND-CPA and KP-CPA security.

proposition 2. *Under the* $STP - Binary - LWE_{n,q,\chi^k}$ *assumption, the FHPRE scheme is IND-CPA secure.*

Proof. We consider the following games for $b \in \{0, 1\}$.

Game$_0^b$**:** This is the real game $Expt_{A,UniFHPRE}^{Ind-CPA,I}(k)$ with b. Suppose the target public key is $ek^0 = P_0^0$, where $P_0^0 = [B_0^0 \parallel -A]$, $B_0^0 = [A \ltimes \vec{s}_0^0 + X_0^0]_q$, $X_0^0 \leftarrow \chi^{N \times 2}$. The other public keys of honest users are $\{ek^i\}_{i=1,\cdots,H} = \{P_0^i\}_{i=1,\cdots,H}$, where $P_0^i = [B_0^i \parallel -A]$, $B_0^i = [A \ltimes \vec{s}_0^i + X_0]_q$, $X_0^i \leftarrow \chi^{N \times 2}$. The challenger computes the re-encryption key from user 0 at level l to user $i \in [H]$ at level 0 of the circuit as $M_{l\to 0}^{0\to i} \leftarrow R_{l\to 0}^{0\to i} P_0^i + P2\left((1; \vec{s}_l^0) \otimes I_2 || 0\right)$, $N_0^i \leftarrow R_0^i P_0^i$, where $0 \in \{0\}^{(n+2)\times n}$, $R_{l\to 0}^{0\to i} \in \mathbb{Z}_2^{(n+2)\lceil \log q \rceil \times N}$, $R_0^i \in \mathbb{Z}_2^{N\times N}$. The challenger computes the target ciphertext on query (m_1, m_2) as follows:

- If $(b = 0)$, it returns $ct \leftarrow \mathbb{Z}_q^{n+2}$.
- If $(b = 1)$, it returns $ct \leftarrow \left[P_0^{0T} \cdot \vec{r} + \lfloor \frac{q}{2} \rfloor \vec{m}\right]_q \in \mathbb{Z}_q^{(n+2)}$, where $\vec{r} \leftarrow \{0,1\}^N$, $\vec{m} = (m_1, m_2, 0 \cdots, 0)^T \in \mathbb{Z}_2^{(n+2)}$.

The adversary finally outputs its guess $b' \in \{0, 1\}$.

Game$_1^b$**:** We replace P_0^i, $P_{(l-1):l}^i$ with ${P_0^i}^+ \leftarrow \mathbb{Z}_q^{N\times 2}$, ${P_{(l-1):l}^i}^+ \leftarrow \mathbb{Z}_q^{(n+2)^2\lceil log q\rceil^3 \times (n+2)}$ for $i \in [H]$. The challenger computes a re-encryption key from user 0 at level l to user i ($i \in [H]$) at level 0 of the circuit by using $\mathbf{s}_l^0$ and ${P_0^i}^+$ as Game$_0^b$. The others are the same as in Game$_0^b$.

Since in the two games, the challenger does not require the secret $\vec{s}_0^i$, there is $P_0^i \approx_c {P_0^i}^+$ under the $STP - Binary - LWE_{n,q,\chi^k}$ assumption. It follows from lemma 2, we have $P_{(l-1):l}^0 \approx_c {P_{(l-1):l}^0}^+$. Furthermore, Game$_0^b \approx_c$Game$_1^b$.

Game$_2^b$**:** We replace $M_{l\to 0}^{0\to i}$, N_0^i with $M_{l\to 0}^{0\to i^+} \leftarrow \mathbb{Z}_q^{(n+2)\lceil \log q\rceil \times (n+2)}$, $N_0^{i+} \leftarrow \mathbb{Z}_q^{N\times(n+2)}$. The others are the same as in Game$_1^b$.

It follows from the leftover hash lemma, we have $M_{l\to 0}^{0\to i} \approx_s M_{l\to 0}^{0\to i^+}$ and $N_0^i \approx_s N_0^{i+}$. Furthermore, Game$_1^b \approx_s$Game$_2^b$.

Game$_3^b$**:** We replace $ct_0^j \leftarrow ReEnc\left(pp, rk_{l\to 0}^{i\to j}, ct_l^i\right)$ with $ct_0^{j+} \leftarrow \mathbb{Z}_q^{n+2}$. The others are the same as in Game$_2^b$.

It follows from the leftover hash lemma, we have $ct_0^{j+} \approx_s ct_0^j$. Furthermore,

$$\text{Game}_2^b \approx_s \text{Game}_3^b$$

.

Finally, we have that Game$_3^0 \approx_s$Game$_3^1$ from the leftover hash lemma. Combining the above indistinguishability, we have shown that Game$_0^0 \approx_c$Game$_0^1$. This completes the proof. □

Theorem 3. *Under the* $STP - Binary - LWE_{n,q,\chi^k}$ *assumption, the homomorphic PRE scheme is KP-CPA secure.*

Proof. We start with the original game with b = 1.

Game$_0$**:** This is the game $Expt_{A,UniFHPRE}^{KP-CPA}(k)$ with b = 1. The challenger runs the adversary with input *pp*, public keys $\{ek^i\}_{i=0,\cdots,H}$ and eval keys $\{evk^i\}_{i=0,\cdots,H}$ for honest users and key pairs $\{ek^i, dk^i\}_{i=H+1,\cdots,H+C}$, $\{evk^i\}_{i=H+1,\cdots,H+C}$ for corrupted users. The challenger generates the real re-encryption key $M_{l\to 0}^{0\to -1} \leftarrow R_{l\to 0}^{0\to -1} P_0^{-1} + P2\left((1; \vec{s}_l^0) \otimes I_2 || 0\right)$, $N_0^{-1} \leftarrow R_0^{-1} P_0^{-1}$, where $0 \in \{0\}^{(n+2)\times n}$, $R_{l\to 0}^{0\to -1} \in \mathbb{Z}_2^{(n+2)\lceil \log q\rceil \times N}$, $R_0^{-1} \in \mathbb{Z}_2^{N\times N}$. On the re-encryption query (0, l, -1, $ct = c_l^o$), it re-encrypts the ciphertext with the real re-encryption key, that is, it returns $ct_0^{-1} = c_0^{-1} = SwitchKey\left(M_{l\to 0}^{0\to -1}, \vec{c}_l^0\right) + {N_0^{-1}}^T \vec{r}_0^{-1}$, where $\vec{r}_0^{-1} \in \mathbb{Z}_2^N$. We summarize the input and the answers to the adversary as follows:

RealPK: P_0^{-1};
Challenge: $M_{l\to 0}^{0\to -1}, N_0^{-1}$;
Table: $M_{l\to 0}^{0\to -1}, N_0^{-1}$;
ReEnc: $ct_0^{-1} = \vec{c}_0^{-1} = SwitchKey_q\left(M_{l\to 0}^{0\to -1}, \vec{c}_l^0\right) + {N_0^{-1}}^T \vec{r}_0^{-1}$.

After the learning phase, the adversary outputs its guess $b' \in \{0, 1\}$.

Game$_1$**:** The challenger replaces P_0^{-1} with $P_0^{-1+} \leftarrow \mathbb{Z}_q^{N\times(n+2)}$, and the re-encryption keys in challenge and the table is constructed from P_0^{-1+} and $\vec{s}_l^0$. The other parts are the same as *Game*$_0$. The challenger re-encrypts a given ciphertext with the re-encryption key in the table. The challenger answers the queries from user 0 at level l to user -1 at level 0 as follows:

RealPK: P_0^{-1+};

Challenge: $M_{l\to 0}^{0\to -1}, N_0^{-1}$;

Table: $M_{l\to 0}^{0\to -1}, N_0^{-1}$;

ReEnc: $ct_0^{-1} = \vec{c}_0^{-1} = SwitchKey_q\left(M_{l\to 0}^{0\to -1}, \vec{c}_l^{0}\right) + {N_0^{-1}}^T \vec{r}_0^{-1}$.

It is easy to verify that $P_0^{-1} \approx_c P_0^{-1+}$ under the $STP - Binary - LWE_{n,q,\chi^k}$ assumption, since we do not need to know $\vec{s}_0^{-1}$. Furthermore, we have $Game_0 \approx_c Game_1$ by the leftover hash lemma.

Game$_2$**:** The challenger replaces $M_{l\to 0}^{0\to -1}, N_0^{-1}$ with $M_{l\to 0}^{0\to -1+} \leftarrow \mathbb{Z}_q^{(n+2)\lceil \log q\rceil \times (n+2)}$, $N_0^{-1+} \leftarrow Z_q^{N\times(n+2)}$. The other parts are not changed from the previous game: the challenger re-encrypts a given ciphertext with the random re-encryption key in the table. The challenger answers the queries from user 0 at level l to user -1 at level 0 as follows:

RealPK: P_0^{-1+};

Challenge: $M_{l\to 0}^{0\to -1+}, N_0^{-1+}$;

Table: $M_{l\to 0}^{0\to -1+}, N_0^{-1+}$;

ReEnc: $ct_0^{-1} = \vec{\tilde{c}}_0^{-1} = SwitchKey_q\left(M_{l\to 0}^{0\to -1+}, \vec{c}_l^{0}\right) + {N_0^{-1+}}^T \vec{r}_0^{-1}$.

It follows from the leftover hash lemma, we have $M_{l\to 0}^{0\to -1} \approx_s M_{l\to 0}^{0\to -1+}$ and $N_0^{-1} \approx_s N_0^{-1+}$. Furthermore, $Game_1 \approx_s Game_2$.

Game$_3$**:** If the query is (0,l -1, $ct = \vec{c}_l^{0}$), then it returns $\vec{c}_0^{-1+} \leftarrow \mathbb{Z}^{n+2}$. The other parts are not changed from the previous game: The challenger answers the queries from user 0 to -1 as follows: The challenger answers the queries from user 0 at level l to user -1 at level 0 as follows:

RealPK: P_0^{-1+};

Challenge: $M_{l\to 0}^{0\to -1+}, N_0^{-1+}$;

Table: $M_{l\to 0}^{0\to -1+}, N_0^{-1+}$;

ReEnc: $ct_0^{-1} = \vec{c}_0^{-1+}$.

It follows from the leftover hash lemma, we have $\vec{c}_0^{-1+} \approx_s \vec{\tilde{c}}_0^{-1}$. Furthermore, $Game_2 \approx_s Game_3$.

Game$_4$**:** The challenger additionally generates another random re-encryption key $M_{l\to 0}^{0\to -1++} \leftarrow \mathbb{Z}_q^{(n+2)\lceil \log q\rceil \times (n+2)}$, $N_0^{-1++} \leftarrow \mathbb{Z}_q^{N\times(n+2)}$ and uses it in the re-encryption oracle. The other parts are not changed from the previous game: As a summary, the challenger answers the queries from user 0 at level l to user -1 at level 0 as follows:

RealPK: P_0^{-1+};

Challenge: $M_{l\to 0}^{0\to -1+}, N_0^{-1+}$;

Table: $M_{l\to 0}^{0\to -1++}, N_0^{-1++}$;

ReEnc:

$$\begin{aligned} ct_0^{-1} &= \vec{c}_0^{-1++} \\ &= SwitchKe_qy\left(M_{l\to 0}^{0\to -1++}, \vec{c}_l^{0}\right) \\ &\quad + {N_0^{-1++}}^T \vec{r}_0^{-1}. \end{aligned}$$

We note that the adversary does not know the alternative fake re-encryption key $M_{l\to 0}^{0\to -1++}, N_0^{-1++}$, directly. Even if the adversary knows the alternative, it cannot distinguish the two games since the re-encrypted ciphertext, which is almost uniformly at random in the ciphertext space from the leftover hash lemma. Hence, we have $Game_3 \approx_s Game_4$.

Game$_5$**:** We again modify the re-encryption key in the table and the re-encryption oracle. The challenger additionally generates a fake re-encryption key $M_{l\to 0}^{0\to -1*} \leftarrow R_{l\to 0}^{0\to -1*} P_0^{-1+} + P2\left(\left(1; \vec{s}_l^{0}\right) \otimes I_2 || 0\right)$, $N_0^{-1*} \leftarrow R_0^{-1*} P_0^{-1+}$, where $0 \in \{0\}^{(n+2)\times n}$, $R_{l\to 0}^{0\to -1*} \in \mathbb{Z}_2^{(n+2)\lceil \log q\rceil \times N}$, $R_0^{-1*} \in \mathbb{Z}_2^{N\times N}$ In the re-encryption oracle, the oracle uses the additional fake-re-encryption key. The other parts are not changed from the previous game: As a summary, the challenger answers the queries from user 0 at level l to user -1 at level 0 as follows:

RealPK: P_0^{-1+};

Challenge: $M_{l\to 0}^{0\to -1+}, N_0^{-1+}$;

Table: $M_{l\to 0}^{0\to -1*}, N_0^{-1*}$;

ReEnc: $ct_0^{-1} = \vec{c}_0^{-1*} SwitchKey_q\left(M_{l\to 0}^{0\to -1*}, \vec{c}_l^{0}\right) + {N_0^{-1*}}^T \vec{r}_0^{-1}$.

It follows from the leftover hash lemma, we have $M_{l\to 0}^{0\to -1+} \approx_s M_{l\to 0}^{0\to -1*}$, $N_0^{-1+} \approx_s N_0^{-1*}$, $\vec{c}_0^{-1++} \approx_s \vec{c}_0^{-1*}$. Furthermore, $Game_4 \approx_s Game_5$.

Game$_6$**:** This is a final game. We replace the fake public key $P_0^{-1+}$with the real public key P_0^{-1} . The other parts are not changed from the previous game: As a summary, the challenger answers the queries from user 0 at level l to user -1 at level 0 as follows:

RealPK: P_0^{-1};

Challenge: $M_{l\to 0}^{0\to -1+}, N_0^{-1+}$;

Table: $M_{l\to 0}^{0\to -1*}, N_0^{-1*}$;

ReEnc: $ct_0^{-1} = \vec{c}_0^{-1*} SwitchKey_q\left(M_{l\to 0}^{0\to -1*}, \vec{c}_l^{0}\right) + {N_0^{-1*}}^T \vec{r}_0^{-1}$.

Since $M_{l\to 0}^{0\to -1+}, N_0^{-1+}$ is distributed uniformly at random, this game is equivalent to $Expt_{A,UniFHPRE}^{KP-CPA}(k)$ with $b = 0$. It follows from the $STP - Binary - LWE_{n,q,\chi^k}$ assumption, we have $P_0^{-1} \approx_c P_0^{-1+}$. Furthermore, $Game_5 \approx_c Game_6$.

Above all, we know $Game_0 \approx_c Game_6$, that is $Expt_{A,UniFHPRE}^{KP-CPA}(k)$ with $b = 0$ and $Expt_{A,UniFHPRE}^{KP-CPA}(k)$ with $b = 1$ are computationally indistinguishable under $STP - Binary - LWE_{n,q,\chi^k}$ assumption. This completes the proof. □

3.3 Comparison

Compared with the homomorphic proxy re-encryption scheme of Ma *et al.* [16, 17], our scheme can encrypt two messages at a time under the same computation complexity , and has the same security of IND-CPA and KP-CPA under LWE. The comparison results in Table 1.

4 Conclusion

In this paper, we adopt the scheme of Ma *et al.* to construct a FHPRE scheme which allows one party to compute arbitrary functions over encrypted data for many parties without the decryption keys. That is, the FHPRE scheme satisfies the "many-to-one" situation. We also prove that our FHPRE scheme is IND-CPA, KP-CPA and master secret secure. We will be devoted to improving the computation efficiency in our future work, so as to make our FHPRE schemes more practical.

Acknowledgments

The authors thank the anonymous referees for their helpful comments. This work was supported by the National Natural Science Foundation of China (61472097) and the Open Fund of the State Key Laboratory of Information Security(2016-MS-10).

References

[1] Y. Aono, X. Boyen, L. Wang, "Key-private proxy re-encryption under LWE," in *The 14th International Conference on Cryptology*, pp. 1-18, 2013.

[2] G. Ateniese, K. Benson, S. Hohenberger, "Key-private proxy re-encryption," in *Cryptographers Track at the RSA Conference*, pp. 279-294, 2009.

[3] G. Ateniese, K. Fu, M. Green, S. Hohenberger, "Improved proxy re-encryption schemes with applications to secure distributed storage," *ACM Transactions on Information and System Security (TISSEC'06)*, vol.9, no. 1, pp.1-30, 2006.

[4] M. Blaze, G. Bleumer, and M. Strauss, "Divertible protocols and atomic proxy cryptography," in *Advances in Cryptology (EUROCRYPT'98)*, pp. 127-144, 1998.

[5] Z. Brakerski, V. Vaikuntanathan, "Efficient fully homomorphic encryption from (Standard) LWE," in *The 52nd Annual IEEE Symposium on Foundations of Computer Science (FOCS'11)*, pp. 97-106, 2011.

[6] Z. Brakerski, C. Gentry, V. Vaikuntanathan, "(leveled) Fully homomorphic encryption without bootstrapping," in *The 3rd Innovations in Theoretical Computer Science Conference (ITCS'12)*, pp. 309-325, 2012.

[7] Z. Brakerski, "Fully homomorphic encryption without modulus switching from classical GapSVP," in *The 32nd Annual Cryptology Conference (CRYPTO'12)*, pp. 868-886, 2012.

[8] Z. Cao, L. Liu, Y. Li, "Ruminations on fully homomorphic encryption in client-server computing scenario," *International Journal of Electronics and Information Engineering*, vol. 8, no. 1, pp. 32-39, 2018.

[9] P. Chung, C. Liu, M. Hwang, "A study of attribute-based proxy re-encryption scheme in cloud environments," *International Journal of Network Security*, vol.16, no.1, pp. 1-13, 2014.

[10] C. Gentry, "A fully homomorphic encryption scheme," *ACM Digital Library*, 2009. ISBN: 978-1-109-44450-6

[11] M. M. Jiang, Y. P. Hu, B. C. Wang, *et al.*, "Lattice-based multi-use unidirectional proxy re-encryption," *Security and Communication Networks*, vol. 18, no. 8, pp. 3796-3803, 2015.

[12] E. Kirshanova "Proxy re-encryption from lattices," in *International Workshop on Public Key Cryptography*, pp. 77-94, 2014.

[13] C. Lan, H. Li, S. Yin, *et al.*, "A new security cloud storage data encryption scheme based on identity proxy re-encryption," *International Journal of Network Security*, vol. 19, no. 5, pp. 804-810, 2017.

[14] L. Liu, Z. Cao, "Analysis of two confidentiality-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] C. Ma, J. Li, G. Du, "A flexible fully homomorphic Encryption," *Wireless Personal Communications*, vol. 95, no. 2, pp. 761-772, 2017.

[16] C. Ma, J. Li, W. Ouyang, "A homomorphic proxy re-encryption from Lattices," in *10th International Conference*, pp.353-372, 2016.

[17] C. Ma, J. Li, W. Ouyang, "Lattice-based identity-based homomorphic conditional proxy re-encryption for secure big data computing in cloud environment," *International Journal of Foundations of Computer Science*, vol. 28, no. 6, pp. 645-660, 2017.

[18] R. Nishimaki, K. Xagawa, "Key-private proxy re-encryption from lattices, revisited," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. 98, no. 1, pp. 100-116, 2015.

[19] O. Regev, "On lattices, learning with errors, random linear codes, and cryptography," in *The 37th Annual ACM Symposium on Theory of Computing (STOC'05)*, pp. 84-93, 2005.

[20] K. Singh, C. P. Rangan, A. K. Banerjee, "Cryptanalysis of unidirectional proxy re-encryption scheme," in *Information and Communication Technology-EurAsia Conference*, pp. 564-575, 2014.

[21] Y. Wang, D. Yan, F. Li, X. Hu, "A key-insulated proxy Re-encryption Scheme for data sharing in a cloud environment," *International Journal of Network Security*, vol. 19, no. 4, pp. 623-630, 2017.

[22] K. Xagawa, *Cryptography with Lattices*, PhD thesis, Tokyo Institute of Technology, Tokyo, 2010.

Table 1: Comparison

Cryptosystem	Computation complexity	Message	INC-CPA	KP-CPA	LWE	Many-to-one
The scheme of [16, 17]	$O(n^2)$	1	YES	YES	YES	YES
The proposed scheme	$O(n^2)$	2	YES	YES	YES	YES

[23] M. Zhang, L. Wu, X. Wang, X. Yang, "Unidirectional IBPRE scheme from lattice for cloud computation," *Journal of Ambient Intelligence and Humanized Computing*, vol. 7, pp. 623-631, 2016.

[24] H. Zhong, J. Cui, R. Shi, C. Xia, "Many-to-one homomorphic encryption scheme," *Security and Communication Networks*, vol. 9, pp. 1007-1015, 2015.

Biography

Juyan Li is currently a Ph.D. candidate at Harbin Engineering University, Harbin, China. Currently his researches focus on cryptography and information security. His email address is lijuyan587@163.com

Chunguang Ma is a full professor in the Harbin Engineering University. He received his BSc, MSc and PhD in 1996, 2002 and 2005 respectively. His current research interests include post-quantum cryptography, distributed cryptographic protocol, cloud computing security and privacy, AI and security, block chain technology and application, etc. He published many papers and his research is funded by Natural Science Foundation of China, Natural Science Foundation of Heilongjiang. He is the corresponding author of this work and his email address is machunguang@hrbeu.edu.cn.

Lei Zhang is currently a Ph.D. candidate at Harbin Engineering University, Harbin, China. Currently his researches focus on cryptography and information security.

Qi Yuan is currently a Ph.D. candidate at Harbin Engineering University, Harbin, China. Currently her researches focus on cryptography and information security.

Face Database Security Information Verification Based on Recognition Technology

Shumin Xue

(Corresponding author: Shumin Xue)

School of Computer Science and technology, Baoji University of Arts and Sciences
No.42 mailbox, Baoji University of Arts and Sciences, Baoji, Shaanxi 721016, China
(Email: shuminx84@126.com)

(Received Aug. 21, 2018; revised and accepted Mar. 12, 2019; First Online June 1, 2019)

Abstract

In recent years, with the rapid development of the Internet, information can be transmitted more and more rapidly, and the issue of confidentiality and security has become increasingly important. Compared with traditional information verification methods, biometric identification is more secure and convenient. This study used the MATLAB software to carry out the simulation of information verification performance of face recognition algorithm based on Local Directional Pattern Algorithm (LDP) and Principal Component Analysis (PCA). The face image data were from ORL database. The results showed that the increase of the training set samples could raise the accuracy of security information verification of the two algorithms and took less time, and under the same number of training samples, the algorithm of face recognition based on PCA, compared with face recognition algorithm based on LDP, had higher accuracy and less time consuming. In conclusion, PCA-based face recognition algorithm is more suitable for security information verification.

Keywords: Face Recognition; Local Directional Pattern; ORL Face Database; Principal Component Analysis

1 Introduction

After entering the 21st century, the Internet has been widely used, and the speed of data transmission is getting faster and faster. At the same time, the security of information data [10] is becoming more and more serious. How to ensure the identification and authentication in the process of network communication has become an important problem in the development of the Internet communication [27]. The essential principle of the identity authentication system [3, 12, 18] is to associate an identifier with the identity of the user, which is identification feature identity, in order to achieve the recognition of the identity of the holder [13]. However, in traditional identification systems, identifiers and holders are independent from each other [4,6,14,25,27]. The system will recognize the holder of the identifier as the correct person once confirming the identifier and will not judge whether the person who holds the identifier is the real owner [5, 24, 26, 28]. Therefore, the new biometric recognition technologies [8, 15, 16] are applied to the identity authentication system.

Biometric features mainly refer to voiceprint, fingerprint, face and so on. These features are unique to individuals. Using biometric features as identifiers in identity authentication systems can solve the disadvantages of physical isolation between identifiers and holders in traditional systems, which is because that biometric and its holder is impossible to separate under normal circumstances. No additional account password is required for biometric applications and the certification system will be more convenient. Gilani [11] proposed a model-based 3D face recognition algorithm, and tested the performance of the algorithm with two large common 3D face data sets. The results showed that the method could effectively recognize face with posture and expression change, and the comparison of single data set and composite data set showed that the recognition accuracy decreased as the size of the image library increased.

In order to achieve robust to illumination, posture and facial expression change of unconstrained face recognition, Ding *et al.* [7] proposed a new methods that extracted "multiple layers of double direction patterns" from face image, and the experimental results on Face Recognition Technology (FERET), CAS - pose, expression, accessories, and lighting (PEAL) - R1, Face Recognition Grand Challenge 2.0 (FRGC 2.0) and Labeled Faces in Wild (LFW) database showed that the method in face recognition and face verification tasks were superior to the most advanced local descriptor. In order to establish the connection between Kinect and face recognition research, Rui *et al.* [21] proposed the first publicly available face database based on Kinect sensor, and used the standards of the proposed face recognition methods to benchmark the proposed database, and proved the performance gain by fractional fusion when depth data was integrated

with Red, Green, Blue (RGB) data. In this study, MATLAB software was used to simulate the security information verification performance of two face recognition algorithms based on Local Directional Pattern Algorithm (LDP) and Principal Component Analysis (PCA).

2 Face Detection, Recognition and Matching

As shown in Figure 1, in the security information verification based on face recognition, first of all, the face image of the registrant is collected through the camera, and then face detection is carried out on the image to ensure that there is only face area in the image. Then, the eigenvectors of the image are extracted by the recognition algorithm. After that, information verification is carried out. The camera is used to collect and verify images, and then face detection is carried out on the verified images. Meanwhile, the non-face area is removed. Then, the same recognition algorithm is used to extract the eigenvectors of the verified images and the classifier is applied to compare the registered image and verify whether the eigenvectors of the image can be classified into one category. If they can, the people in the two images will be judged to be the same person, and pass the verification of information, and if not, they will be determined as different persons, and information validation will fail.

3 Face Recognition Algorithm Based on LDP

LDP [23] can extract image features, the principle of which is statistics of directional edge. X is a pixel in the image and will be centered on the pixel gray value in the field of 3×3 to have convolution with Kirsch template N [1] to get the corresponding edge response $|n_i|$, and then edge response will be sorted according to the gradient. The first K is denoted as code 1, and the rest of the record is denoted as code 0. There are 8 types of template N, including

$$N_0 = \begin{bmatrix} -3 & -3 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & 5 \end{bmatrix}$$

$$N_1 = \begin{bmatrix} -3 & 5 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & -3 \end{bmatrix}$$

$$N_2 = \begin{bmatrix} 5 & 5 & 5 \\ -3 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix}$$

$$N_3 = \begin{bmatrix} 5 & 5 & -3 \\ 5 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix}$$

$$N_4 = \begin{bmatrix} 5 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & -3 & -3 \end{bmatrix}$$

$$N_5 = \begin{bmatrix} -3 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & 5 & -3 \end{bmatrix}$$

$$N_6 = \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & -3 \\ 5 & 5 & 5 \end{bmatrix}$$

$$N_7 = \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & 5 \\ -3 & 5 & 5 \end{bmatrix}$$

The formula of LDP coding [2] is:

$$\left\{ \begin{array}{l} n_k = k_{th}(N) \\ N = |n_0, n_1, \cdots, n_7| \\ LDP_k(r,c) = \sum_{i=0}^{7} b_i(n_i - n_k) \times 2^i \\ b_i(n_i - n_k) = \left\{ \begin{array}{ll} 1 & n_i - n_k \geq 0 \\ 0 & n_i - n_k < 0 \end{array} \right\} \end{array} \right\} \quad (1)$$

where n_k is the edge response of K^{th}, N is Kirsch template, $LDP_R(r,c)$ is the LDP code corresponding to center point c, and r is the radius of the field which was set as 3 in this study.

As shown in Figure 2, each pixel in the original face image was converted into LDP code by combining 8 Kirsch templates and Equation (1), and then the LDP coded image of the face was constructed according to the LDP code. After that, the LDP coded image was divided into blocks of number $a \times b$ to extract histogram in each of them. Finally, the histogram of the extracted block was connected end to end to obtain the final eigenvector.

After obtaining the final eigenvector, the classifier was required to classify the collected eigenvector to determine whether the face image was the same face classification. Moreover, different face recognition algorithms had different vector classifiers, and the selection of classifier would directly impact on the recognition effect.

In this study, LDP recognition algorithm adopted nearest neighbor classifier [20] to classify eigenvectors, and distance function was applied to calculate the contiguous degree between samples. The formula of LDP recognition algorithm is as follows:

$$d_{\chi^2}(a,b) = \sum_n \frac{a_n^2 - 2a_n b_n + b_n^2}{a_n + b_n} \quad (2)$$

where a, b are LDP eigenvectors which are corresponding to two face images respectively, and $d_{\chi^2}(a,b)$ is the chi-square distance between the eigenvectors of two images.

4 Face Recognition Algorithm Based on PCA

The basic principle of PCA [19] is to transform the original random vector related to components into random

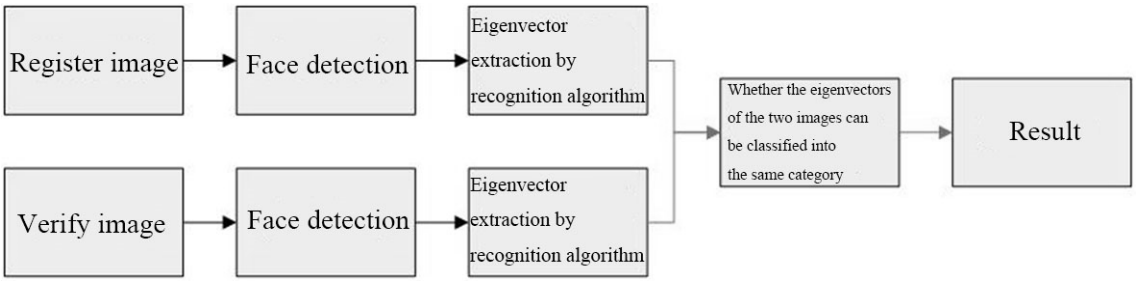


Figure 1: Face detection, recognition and matching

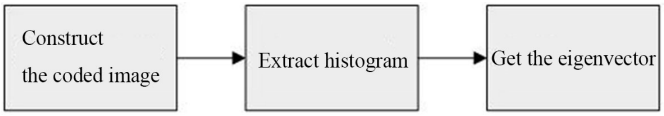


Figure 2: The extraction process of facial features of LDP

vector unrelated to components by means of orthogonal transformation, reduce the dimensionality of the transformed multidimensional variable system, and then transform the low-dimensional variable system into a one-dimensional system through value function. PCA is a statistical method of dimensionality reduction in mathematics, which can greatly reduce the amount of calculation and improve the efficiency of calculation.

The extraction process of eigenvectors based on PCA is as follows. First, it is necessary to calculate the mean vector of all images in the training sample. The calculation formula of the mean vector of all images [29] is as follows:

$$n = \frac{\sum_{i=1}^{Q} A_i}{Q} \tag{3}$$

where n is the mean vector of all training sample images, Q is the sum of training sample images, and A_i is the original eigenvector of the i^{th} training sample image.

Then, the original eigenvector mean value of the single face image in the training sample is calculated, and the calculation formula of the original eigenvector mean value of the single face image is as follows:

$$n_i = \frac{\sum_{j=1}^{L} A_{ij}}{L} \tag{4}$$

where n_i is the average value of the original eigenvector of the i^{th} person's face image, L is the number of training samples of the i^{th} person's face image, and A_{ij} is the original eigenvector of the j^{th} individual face image training sample of the i^{th} individual. Then the population dispersion matrix is calculated:

$$\left\{ \begin{array}{l} S_b = \frac{BB^T}{P} \\ B = [(n_0 - n), (n_1 - n), \cdots, (n_{P-1} - n)] \end{array} \right\} \tag{5}$$

where S_b is the total population scatter matrix, P is the sum of people trained, and B is the matrix of the difference between the vector mean of single face image and the vector mean of all training samples. Then the construction matrix is calculated:

$$\left\{ \begin{array}{l} R = B^T B \\ \sum_{i=1}^{t} \delta_i / \sum_{i=1}^{P} \delta_i \geq \theta \\ U_i = \frac{BN_i}{\sqrt{\delta_i}} \end{array} \right\} \tag{6}$$

where R is the construction matrix, δ_i is the eigenvalue of the i^{th} training sample, V_i is the orthonormalized eigenvector of the i^{th} training sample, and U_i is the orthonormalized eigenvectors of S_b.

To sum up, the projection of the average eigenvector of each person's training sample in the eigensubspace is $C_i = W^T m$, where C_i is the feature subspace for each person and W^T is the dimension reduction matrix.

After the dimensionality reduction of high-dimensional facial image eigenvectors by PCA, it is necessary to select an appropriate classifier to classify the collected eigenvectors. In this study, linear kernel function (SVM) classifier was selected to recognize the eigenvectors.

Firstly, the training data set of some feature space was selected, and then the optimization problem was constructed for the data set: the objective function was:

$$\left\{ \begin{array}{l} \min[\frac{\sum_{i=1}^{M}\sum_{j=1}^{M}\alpha_i\alpha_j y_i y_j (K(x,z)+\frac{\lambda_{ij}}{C})}{2} - \sum_{j=1}^{M}\alpha_j] \\ \lambda_{ij} = \left\{ \begin{array}{ll} 1 & i = j \\ 0 & i \neq j \end{array} \right\} \end{array} \right\} \tag{7}$$

The condition was:

$$\left\{ \begin{array}{l} \sum_{i=1}^{M} \alpha_i y_i = 0 \\ \alpha_i \geq 0 \end{array} \right\} \tag{8}$$

where $K(x, z)$ is kernel function, α is a parameter suitable for C and λ_{ij} is the parameter that determines whether or not $\frac{1}{C}$ exists.

Finally, the decision function was constructed:

$$\left\{ \begin{array}{l} b = y_j(1 - \frac{\alpha_j}{C}) - \sum_{i=1}^{M} \alpha_i y_i K(x, z) \\ 0 < \alpha_j < C \\ f(x) = sign(\sum_{i=1}^{N} \alpha_i y_i K(x, z + b)) \end{array} \right\} \tag{9}$$

where α_j is a positive component of α and b is a parameter involved in the decision.

5 Simulation Experiment

5.1 Experimental Environment

The experiments in this study were carried out on a laboratory server. The server configuration was Windows7 system, I7 processor and 16G memory. MATLAB software [9] was used for algorithm programming.

5.2 Experimental Data

This study adopted the data set of ORL face database [22] that contained 400 positive face images of people distributed in 40 folders which were corresponding to 40 people respectively. Each folder contained 10 positive face images of the same person with different expressions, and the image size was 112×92 pixels.

5.3 Experimental Settings

The experimental procedure of this study is shown in Figure 3. Firstly, the data set was divided into training set and test set. n face images were randomly selected from the folder corresponding to each person as the training set, and the rest as the test set. The selection of n was one, three and five. Then, the training set was used to train LDP recognition algorithm and PCA recognition algorithm respectively for face recognition where block parameters of 7×7 were selected when LDP image histogram was extracted from the LDP recognition algorithm. After extracting LDP eigenvectors, the nearest neighbor classifier was applied to classify features, so as to train LDP recognition algorithm; in the PCA recognition algorithm, the image was dimensionalized by PCA, the principal components of the 20-dimensional vector were obtained, and the SVM classifier classified them. The kernel function in the SVM was the linear kernel function. was set as 1, so as to train the PCA recognition algorithm. At last, the face images in the test set were detected and classified by two recognition algorithms respectively.

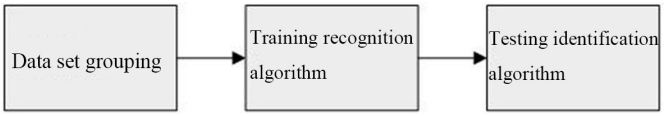


Figure 3: Experimental flow

When the face image in the test set was classified to the corresponding person, it meant that it passed the security information verification. The correct number of test set of each person should be counted, and the accuracy of the two recognition algorithms through the security information verification should be calculated.

5.4 Experiment Results

As shown in Figure 4, when the number of training sets was one face image per person, the accuracy of security information verification of the LDP-based recognition algorithm was 68.45%, and that of the PCA-based recognition algorithm was 70.12%. When the number of training sets was three face images per person, the accuracy of the LDP based recognition algorithm was 78.15%, and that of the PCA-based recognition algorithm was 86.45%. When the number of training sets was five face images per person, the accuracy of the LDP based recognition algorithm was 88.54%, and that of the PCA-based recognition algorithm was 98.41%.

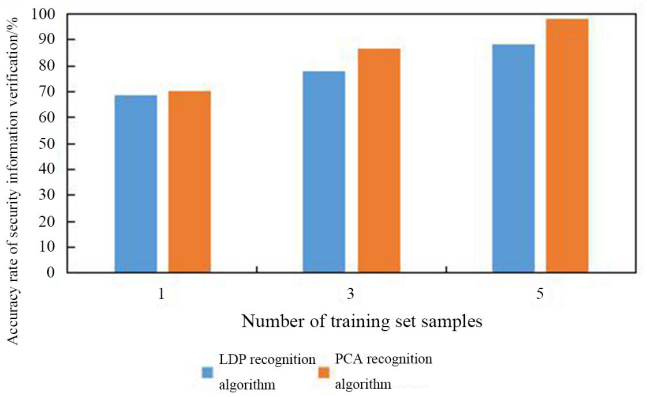


Figure 4: The accuracy rate of security information verification of the two algorithms

It could be seen from Figure 4 that with the increase of the sample number of the training set, the accuracy of the two recognition algorithms also increased. At the same time, with the same number of training samples, the accuracy of the PCA based recognition algorithm was higher than that of the LDP-based recognition algorithm.

As shown in Table 1, when the number of training samples was one face image per person, the security information verification of LDP algorithm took 366 ms, while that of PCA algorithm took 354 ms. When the number of training samples was three face images per person, the security information verification of LDP algorithm took 329 ms, while that of PCA algorithm took 321 ms. When the number of training samples was five face images per person, the security information verification of LDP algorithm took 315 ms, while that of PCA algorithm took 301ms. It could be seen that with the increase of training samples, the time required by the two recognition algorithms to verify the security information of the test set also decreased. At the same time, under the same number of training samples, the PCA recognition algorithm took less time.

6 Conclusion

This paper simply introduced face recognition algorithms based on LDP and PCA, and the MATLAB software information was used to simulate the security information verification performance of two face recognition al-

Table 1: Security information verification time of the two recognition algorithms

Training sample size	The time consumed by LDP algorithm/ms	The time consumed by PCA algorithm/ms
1	366	354
3	329	321
5	315	301

gorithms, the two algorithms were trained by training samples containing one face image per person, three face images per person and five images per person, and then the rest of the image was as a test set. When the training samples were 1, 3 and 5 per person, the accuracy rate of security information verification of LDP recognition algorithm was 68.45%, 78.15% and 88.54%, respectively. The accuracy rate of security information verification of PCA recognition algorithm was 70.12%, 86.45% and 98.41%, respectively. The accuracy of both algorithms increased with the increase of training samples, and the accuracy of PCA algorithm was higher. When the number of training samples was one, three and five per person, the security information verification of LDP recognition algorithm took 366 ms, 329 00 ms and 315 ms, while the security information verification of PCA recognition algorithm took 354 ms, 321 ms and 301 ms. The time of the two algorithms decreased with the increase of training samples, and the time of PCA algorithm was less.

References

[1] S. Chakraborty, S. K. Singh, P. Chakraborty, "Local directional gradient pattern: A local descriptor for face recognition," *Multimedia Tools & Applications*, vol. 76, no. 1, pp. 1201–1216, 2017.

[2] S. Chakraborty, S. K. Singh, P. Chakraborty, "Correction to: Local directional gradient pattern: A local descriptor for face recognition," *Multimedia Tools & Applications*, vol. 77, no. 15, pp. 20269, 2018.

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

[4] T. Y. Chen, C. C. Lee, M. S. Hwang, J. K. Jan, "Towards secure and efficient user authentication scheme using smart card for multi-server environments," *Journal of Supercomputing*, vol. 66, no. 2, pp. 1008–1032, Nov. 2013.

[5] Y. C. Chen, W. L. Wang, M. S. Hwang, "RFID authentication protocol for anti-counterfeiting and privacy protection", in *The 9th International Conference on Advanced Communication Technology*, pp. 255–259, 2007.

[6] M. A. A. Dewan, E. Granger, G. L. Marcialis, *et al.*, "Adaptive appearance model tracking for still-to-video face recognition," *Pattern Recognition*, vol. 49(C), pp. 129–151, 2016.

[7] C. Ding, J. Choi, D. Tao, *et al.*, "Multi-directional multi-level dual-cross patterns for robust face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 3, pp. 518–531, 2016.

[8] C. Ding, D. Tao, "A comprehensive survey on pose-invariant face recognition," *ACM Transactions on Intelligent Systems & Technology*, vol. 7, no. 3, 2016.

[9] A. Fathi, P. Alirezazadeh, F. Abdali-Mohammadi, "A new global-Gabor-Zernike feature descriptor and its application to face recognition," *Journal of Visual Communication & Image Representation*, vol. 38, pp. 65–72, 2016.

[10] J. Galbally, S. Marcel, J. Fierrez, "Biometric antispoofing methods: A survey in face recognition," *IEEE Access*, vol. 2, pp. 1530–1552, 2017.

[11] S. Z. Gilani, A. Mian, "Towards large-scale 3D face recognition," in em International Conference on Digital Image Computing: Techniques & Applications, 2016.

[12] M. S. Hwang, Li-Hua Li, "A New Remote User Authentication Scheme Using Smart Cards", *IEEE Transactions on Consumer Electronics*, vol. 46, no. 1, pp. 28–30, Feb. 2000.

[13] M. S. Hwang, J. W. Lo, S. C. Lin, "An efficient user identification scheme based on ID-based cryptosystem", *Computer Standards & Interfaces*, vol. 26, no. 6, pp. 565–569, 2004.

[14] M. S. Hwang, C. H. Wei, C. Y. Lee, "Privacy and security requirements for RFID applications", *Journal of Computers*, vol. 20, no. 3, pp. 55–60, Oct. 2009.

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

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

[17] C. T. Li, M. S. Hwang, C. Y. Liu, "An electronic voting protocol with deniable authentication for mobile Ad hoc networks," *Computer Communications*, vol. 31, no. 10, pp. 2534–2540, June 2008.

[18] H. Li, L. Zhang, H. Bing, *et al.*, "Sequential three-way decision and granulation for cost-sensitive face

recognition," *Knowledge-Based Systems*vol. 91(C), pp. 241–251, 2016.

[19] Y. Qiang, W. Rong, X. Yang, *et al.*, "Diagonal principal component analysis with non-greedy ℓ_1-norm maximization for face recognition," *Neurocomputing*, vol. 171, pp. 57–62, 2016.

[20] S. P. Ramalingam, "Dimensionality reduced local directional number pattern for face recognition," *Journal of Ambient Intelligence & Humanized Computing*, vol. 9, no. 1, pp. 1–9, 2016.

[21] M. Rui, N. Kose, J. L. Dugelay, "KinectFaceDB: A kinect database for face recognition," *IEEE Transactions on Systems Man & Cybernetics Systems*, vol. 44, no. 11, pp. 1534–1548, 2017.

[22] A. Soula, S. B. Said, R. Ksantini, *et al.* , "A novel kernelized face recognition system," in *4th International Conference on Control Engineering & Information Technology*, pp. 1–5, Hammamet, 2016.

[23] Srinivasa PerumalR.aChandra MouliP.V.S.S.R., "Dimensionality reduced local directional pattern (DR-LDP) for face recognition," *Expert Systems with Applications*, vol. 63, pp. 66–73, 2016.

[24] C. H. Wei, M. S. Hwang, Augustin Y. H. Chin, "A mutual authentication protocol for RFID", *IEEE IT Professional*, vol. 13, no. 2, pp. 20–24, Mar. 2011.

[25] C. H. Wei, M. S. Hwang, Augustin Y. H. Chin, "An authentication protocol for low-cost RFID tags", *International Journal of Mobile Communications*, vol. 9, no. 2, pp. 208–223, 2011.

[26] C. H. Wei, M. S. Hwang, Augustin Y. H. Chin, "An improved authentication protocol for mobile agent device in RFID," *International Journal of Mobile Communications*, vol. 10, no. 5, pp. 508–520, 2012.

[27] C. H. Wei, M. S. Hwang, Augustin Y. H. Chin, "Security analysis of an enhanced mobile agent device for RFID privacy protection," *IETE Technical Review*, vol. 32, no. 3, pp. 183–187, 2015.

[28] C. H. Wei, M. S. Hwang, Augustin Y. H. Chin, "A secure privacy and authentication protocol for passive RFID tags," *International Journal of Mobile Communications*, vol. 15, no. 3, pp. 266–277, 2017.

[29] Z. Zhong, L. Shuang, "Coupled principal component analysis based face recognition in heterogeneous sensor networks," *Signal Processing*, vol. 126, pp. 134–140, 2016.

Biography

Xue Shumin, February 1984, master, lecturer, mainly engaged in computer applications, face recognition and database research. Presided over and completed two university-level key projects, presided over one university-level educational reform project, participated in two industrial key projects of baoji science and technology bureau, participated in one project of shaanxi provincial department of education, and participated in a number of university-level key and general projects. Guided the undergraduate innovation and entrepreneurship project "classroom attendance system based on face recognition", which was rated as a national innovation project; guided the student "Internet +" project "ZhenShiTong – campus intelligent management platform based on face recognition", published many papers, two patents and three software Copyrights.

Design of an Anonymous Lightweight Communication Protocol for Smart Grid and Its Implementation on 8-bit AVR and 32-bit ARM

Dariush Abbasinezhad-Mood, Arezou Ostad-Sharif, and Morteza Nikooghadam

(Corresponding author: Dariush Abbasinezhad-Mood)

The Department of Computer Engineering and Information Technology, Imam Reza International University
Razavi Khorasan Province, Mashhad, Sanabaad, Daneshgah Avenue, Mashhad, Iran
(Email: dariush.abbasinezhad@imamreza.ac.ir.)

(Received Sept. 12, 2017; revised and accepted Oct. 13, 2018; First Online June 1, 2019)

Abstract

Upgrading the conventional electrical grid to smart grid offers more efficiency, resiliency, and reliability. Thus, the smart grid adoption is essential in today's modern countries and the information age. In smart grid, consumption reports are gathered from smart meters and sent to the control center and some control messages are sent vice versa. These bidirectional communications are subject to various security challenges. Because of the constrained resources of smart meters, employing lightweight communication protocols is critical. For this purpose, recently, scholars have proposed several lightweight communication protocols. Nonetheless, most of these protocols are not anonymous or fail to assuage the entire desired security features. Therefore, in this paper, we propose an efficient communication scheme that not only is anonymous, but also can thwart the well-known attacks. Our actual hardware performance analysis, which has been done on both 8-bit AVR and 32-bit ARM microcontrollers, confirms the outperformance of the proposed scheme.

Keywords: ARM; AVR; Lightweight; Secure Communications; Smart Grid Security

1 Introduction

The legacy energy grid cannot fulfil the actual needs of today's modern countries and the information era [15, 20]. Hence, in near future, the adoption of smart grid (SG) will become so critical as it promises to offer more efficiency, resiliency, and reliability [2]. In *SG*, the smart meters (SMs), which are some resource-constrained electronic measurement devices, gather the energy usage of consumers and send them to the control center via some intermediary gateways, such as neighborhood gateways (NGs) [3]. The communications of *SM*s and *NG*s are bidirectional and the *NG*s may also send some commands to *SM*s [4]. As these two-way communications are susceptible to several security threats, proposing secure communication protocols is vital. Evidently, overlooking the security concerns will hamper the wide adoption of *SG*. The security needs to be fully considered from the very beginning of usage reports collection by *SM*s up to their reception at the control and power management center. Further, the constrained resources of *SM*s in terms of flash storage and computational capability should be fully taken into consideration [4, 20].

The security challenges have taken much attention from the academia [11, 17, 26], and same as other fields [9, 12–14, 16, 18, 23–25, 27], for *SG*, many scholars have presented key agreement schemes [2, 5, 6, 10, 21, 22, 29] and secure communication protocols [4, 19, 20, 28]. Nevertheless, careful assessment of the related works shows that the existing protocols cannot totally fulfil the desired security properties. As an example, most of the existing communication schemes cannot provide the anonymity, a feature that helps to better preserve the privacy of consumers.

1.1 Related Work

In 2011, Fouda *et al.* [10] addressed the traffic analysis, denial of service (DoS), *DoS* buffer overflow, spoofing, reconfigure, man in the middle, and replay attacks as the security threats that exist in *SG* communications. Further, they presented a lightweight message authentication protocol and indicated that their scheme can provide semantic-secure shared key, mutual authentication, and an encrypted channel for successive communications. In 2013, Li *et al.* [19] put forward an authenticated communication scheme called *AC* using the Merkle hash tree. Although the presented scheme by Li *et al.* has a proper level of performance, it requires lots of space for storing the generated parameters. In addition, in [19], the authors have indicated that their scheme can thwart the replay, message injection, message analysis, and message

modification attacks. Nonetheless, their scheme cannot resist the pollution or *DoS* attack. In 2016, Liu *et al.* [20] have proposed another lightweight authenticated communication protocol named *LAC* that has a better performance than *AC* in terms of storage space, communication overhead, and computational cost. However, same as *AC*, *LAC* fails to withstand the pollution attack and needs lots of space for storing generated parameters. At the same year, two other schemes have been proposed by Mahmood *et al.* [21] and Uludag *et al.* [28]. Mahmood *et al.* have mainly concentrated on the authentication and key agreement and Uludag *et al.* have proposed a holistic scheme consisting of both key establishment and data collection. Quite recently, to remedy the challenges of Li *et al.*'s scheme [19] and Liu *et al.*'s scheme [20], we have proposed an ultra-lightweight scheme for communications of *SM*s and *NG*s [4]. However, none of these schemes can offer the anonymity. Therefore, in this paper, we try to propose an anonymity-preserving protocol that can withstand the well-known attacks and has a proper level of performance to be executed on the resource-constrained *SM*s.

1.2 Motivation

First, as stated earlier, Li *et al.*'s [19] and Liu *et al.*'s [20] schemes suffer from the pollution attack. Second, as mentioned in [4, 19], the next generation of *SM*s should be able to send the consumption reports in one-minute or less time intervals. In this case, the presented schemes by Li *et al.* [19] and Liu *et al.* [20] would be impractical as the required storage space will exceed the flash storage of most popular low-cost microcontrollers. Third, most of the existing communication protocols are not anonymous or fail to withstand the well-known attacks like the *SM* memory modification attack. Fourth, more efficient the presented communication protocol, more its suitability to be performed on *SM*s. Finally yet significantly, the *ARM* microcontrollers are one of the most cost-effective and energy-efficient *MCU*s in the market and it will be useful to test the performance of different cryptographic operations on them in comparison to their counterparts in *AVR*. These facts motivated us to propose a communication protocol that can assuage the mentioned necessities.

1.3 Contribution

The fourfold contribution of this paper is as follows.

- Presenting a communication scheme, which (a) is secure against the well-known attacks, (b) provides the anonymity, and (c) is much more efficient than several recently published schemes.
- Presenting the details of shared key storage and retrieval both in the *SM* flash storage and the *NG* database.
- Reducing the number of parameters need to be stored in the *SM* flash storage to only two parameters.

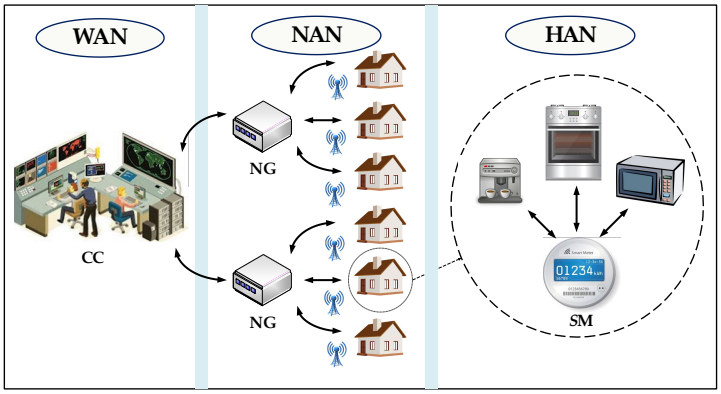


Figure 1: The network model of SG

- Implementation of different cryptographic operations on both 8-bit *AVR* and 32-bit *ARM*.

The remainder of this paper is structured as follows. The network model, attack model, and design objectives are explained in Section 2. In Section 3, the proposed anonymous communication scheme is delineated. In Sections 4 and 5, the descriptive security analysis and automatic formal verification of security using ProVerif are presented, respectively. The efficiency analysis and experimental study are presented in Section 6. Finally, Section 7 concludes this paper.

2 Models and Design Objectives

2.1 Network Model

The communication model of *SG* has been presented in several papers [4,7,20]. As shown in Figure 1, a smart metering communication system consists of *SM*s, which are responsible for energy usage measurement and collection; home area network (HAN), which is a network formed by an *SM* and its controlling smart appliances, neighborhood area network (NAN), which is composed of hundreds of *SM*s and collects information from multiple *HAN*s; wide area network (WAN), which is referred to as the backhaul and carries metering data to the control center (CC); and gateways, which gather energy consumption reports from several *HAN*s. The focus of this paper is on the secure bidirectional communications of *SM*s and *NG*s.

2.2 Attack Model

In this paper, we assume three kinds of adversary as follows.

1) **The external adversary.** The external adversary ε can eavesdrop or alter the exchanging messages between *SM*s and *NG*s and can perform replay, *DoS*, message injection, message modification, and message analysis attacks.

2) **The internal adversary.** The internal adversary $\jmath$ not only has the power of the external adversary, but

also can gain access to the stored records of the *NG* database.

3) **The global adversary.** The global adversary φ not only has the power of the internal adversary, but also can read the content of *SM*s flash storage.

2.3 Design Objectives

In this paper, we aim to propose a two-way anonymous communication scheme which can fulfil the following goals.

1) **Anonymity.** The proposed communication protocol should be designed such that an adversary cannot distinguish the real identifier of *SM*s.

2) **Near real-time authentication.** Each *SM* must be able to check that the received message has been sent from the authorized *NG* and nobody has impersonated *NG*. Similarly, *NG* must be able to check that the received message has been sent from the intended *SM*. Both of these actions should be done in a short amount of time.

3) **Confidentiality.** The exchanging messages between *SM*s and *NG*s must only be accessed by the intended party. That is to say, except authorized *SM*s and *NG*s, nobody else must be able to gain access to confidential messages.

4) **Message modification attack resistance.** Both *SM*s and *NG*s must be able to check whether or not a received message has been altered by an adversary during the transfer.

5) **Message injection attack resistance.** Both *SM*s and *NG*s must be able to filter the fabricated messages that may be sent by an attacker.

6) ***DoS* attack resistance.** Both *SM*s and *NG*s must be able to detect the modifications that make services unavailable.

7) **Message analysis attack resistance.** The adversary must not be able to recover the consumption reports or control messages by just eavesdropping the exchanging messages.

8) **Replay attack resistance.** Both *SM*s and *NG*s must be able to verify that a valid message is not a repeated one.

9) **Insider attack resistance.** In this paper, we assume the insider as a person who can easily gain access to the *NG* database and an insider attacker as an adversary who is an insider. The proposed scheme must be able to resist the attacks that may be performed by an insider attacker.

10) ***SM* memory modification attack resistance.** The protocol must be designed such that the stored data in the flash memory of *SM*s be kept confidential and even if an adversary alters them, the tampering could be revealed so soon.

11) **Low storage and computational costs.** Due to the constrained resources of *SM*s, the proposed protocol must be as lightweight as possible.

3 Proposed Lightweight Scheme

In this section, a complete description of the proposed anonymous lightweight communication scheme is given. Our scheme can be employed effectively for secure bidirectional communications of *SM*s and *NG*s in *SG*. In our scheme, each day, every 15 minutes, the i^{th} *SM*, SM_i, measures consumption report D_j^i, where $j = 1, 2, \ldots, 96$, and sends it to *NG*. Meanwhile, *NG* may send four control messages CM_k^i, where $k = 1, 2, 3, 4$, in order to be performed by SM_i. Here, same as the other related schemes, we consider 15 minutes time intervals for consumption reports collection. However, in the "efficiency analysis and experimental study" section of this paper, we will evaluate the schemes according to different time intervals.

Our scheme is composed of three phases, namely "initialization," "shared key generation and storage," and "secure message transmission." In the following subsections, we elaborate each phase. The notations used in our scheme together with their definitions have been listed in Table 1.

Table 1: Notations and their definitions

Notation	Definition
SM_i	i^{th} smart meter
NG	neighborhood gateway
ID_i	identifier of SM_i
ID_{NG}	identifier of NG
m_i	secret key of SM_i
s	secret key of NG
K_i^{NG},K_i^{SM}	shared key between NG and SM_i
D_j^i	j^{th} usage report of SM_i
CM_k^i	k^{th} control message for SM_i
T_i	j^{th} timestamp of data collection
CT_i	current time of SM_i
CT_{NG}	current time of NG
Enc	symmetric encryption
Dec	symmetric decryption
Δt	predefined maximum transmission delay

3.1 Initialization

In this phase, for each *SM*, *NG* first generates a random number r_i, then, computes encrypted identifier EID_i as

Equation (1), where Enc_s is the symmetric encryption using the key s, s is the private key of NG, and ID_i is the identifier of SM_i. Finally, NG sends the generated EID_i to SMs through a reliable medium.

$$EID_i = Enc_s(ID_i \,\|\, r_i). \tag{1}$$

3.2 Shared Key Generation and Storage

In this phase, SM_i and NG share a key K_i by running the proposed protocol in [2]. The following steps are done after the shared key generation.

Step 1. NG first calculates H_i^{NG} and E_i^{NG} as Equations (2) and (3), next, it adds (ID_i, E_i^{NG}, H_i^{NG}) record to its database. Here, K_i^{NG} is the NG side K_i.

$$H_i^{NG} = h(K_i^{NG} \,\|\, ID_i), \tag{2}$$
$$E_i^{NG} = K_i^{NG} \oplus h(s). \tag{3}$$

Step 2. SM_i first computes E_i^{SM} as Equation (4), where K_i^{SM} and m_i are the SM_i side K_i and SM_i's private key, respectively. Afterwards, it stores E_i^{SM} and EID_i in its flash storage. It should be noted that K_i^{NG} and K_i^{SM} are identical and we have named them differently to make them distinguishable.

$$E_i^{SM} = K_i^{SM} \oplus h(m_i). \tag{4}$$

Step 3. According to the security policies of system, the shared key K_i can be updated by rerunning the presented protocol in [2].

3.3 Secure Message Transmission

In this phase, every 15 minutes, SM_i sends usage report D_j^i to NG and meanwhile, NG may send 4 control messages CM_k^i to SM_i. The following steps are done in this phase. An illustration of this phase is depicted in Figure 2.

1) SM_i **to** ***NG*** **message transmission**

Step 1. SM_i first retrieves E_i^{SM} from its flash memory, then, computes K_i^{SM} as Equation (5).

$$K_i^{SM} = E_i^{SM} \oplus h(m_i). \tag{5}$$

Step 2. SM_i computes verifier V_j^i as Equation (6). This verifier will be used by NG for checking the message integrity, the SM_i authentication, and SM_i memory modification attack check. Here, T_j is the j^{th} timestamp of data collection.

$$V_j^i = h(D_j^i \oplus T_j \oplus ID_i). \tag{6}$$

Step 3. SM_i computes M_j^i as Equation (7).

$$M_j^i = Enc_{K_i^{SM}}(D_j^i). \tag{7}$$

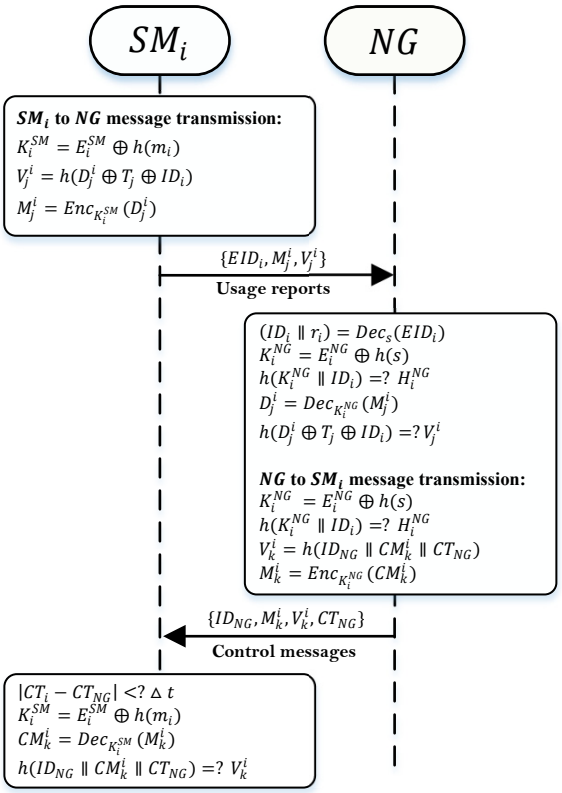


Figure 2: Secure message transmission phase of the proposed scheme

Step 4. SM_i sends $\{EID_i, M_j^i, V_j^i\}$ to NG. Here, there is no need to send timestamp T_j since NG can obtain an identical one. We refer readers to [4] for the detailed discussion.

Step 5. Upon receiving the message $\{EID_i, M_j^i, V_j^i\}$, NG first decrypts EID_i to achieve ID_i. Then, according to the obtained ID_i, retrieves H_i^{NG} and E_i^{NG} from its database. Finally, achieves K_i^{NG} as Equation (8) and checks whether $h(K_i^{NG} \,\|\, ID_i) = H_i^{NG}$ holds to ensure the integrity of ID_i, E_i^{NG}, and H_i^{NG}.

$$K_i^{NG} = E_i^{NG} \oplus h(s). \tag{8}$$

Step 6. NG obtains usage report D_j^i as Equation (9).

$$D_j^i = Dec_{K_i^{NG}}(M_j^i). \tag{9}$$

Step 7. Using its current time, NG first achieves timestamp T_j, then checks whether $h(D_j^i \oplus T_j \oplus ID_i) = V_j^i$ holds or not in order to ensure the message has been sent from the intended SM_i, the message has not been altered during its transfer, the message is not a repeated one, and the SM_i memory has not been changed.

Step 8. NG compares D_j^i with predefined format and if it conforms, accepts D_j^i from SM_i.

2) ***NG*** **to** SM_i **message transmission**

Step 1. NG picks intended control message CM_k^i that is needed to be performed by SM_i, where k = 1,2,3,4.

Step 2. NG retrieves H_i^{NG} and E_i^{NG} corresponding to ID_i from its database and achieves K_i^{NG} as Equation (8).

Step 3. NG checks the equality of $h(K_i^{NG} \| ID_i) = H_i^{NG}$ to ensure the integrity of ID_i, E_i^{NG}, and H_i^{NG}.

Step 4. NG computes verifier V_k^i as Equation (10) that will be used by SM_i for the message integrity check, the NG authentication, and SM_i memory modification attack check. Here, CT_{NG} is the current time of NG.

$$V_k^i = h(ID_{NG} \| CM_k^i \| CT_{NG}) \quad (10)$$

Step 5. NG computes M_k^i as Equation (11) in order to ensure that the control message CM_k^i can only be accessed by SM_i.

$$M_k^i = Enc_{K_i^{NG}}(CM_k^i). \quad (11)$$

Step 6. NG sends $\{ID_{NG}, M_k^i, V_k^i, CT_{NG}\}$ to SM_i.

Step 7. Upon receipt of the message, SM_i checks whether $|CT_i - CT_{NG}| < \Delta t$ holds or not to ensure that the received message is not a repeated one. CT_i is the current timestamp of SM_i and Δt is a predefined maximum transmission delay.

Step 8. SM_i first retrieves E_i^{SM} from its flash memory, then, computes K_i^{SM} as Equation (5).

Step 9. SM_i achieves control message CM_k^i as Equation (12).

$$CM_k^i = Dec_{K_i^{SM}}(M_k^i). \quad (12)$$

Step 10. SM_i checks the equality of Equation (10) to ensure that the message has been sent from the authentic NG, it has not been changed during the transfer, and its memory has not been altered.

Step 11. SM_i checks the CM_k^i format, then executes it.

It is worth noting that to guarantee the strong anonymity of SMs, NG needs to generate a new random number and update the EID_i. The new generated EID_i can be sent to SM via a control message.

A feature-based comparison with similar recently-published schemes is presented in Table 2.

4 Security Analysis

According to our objectives and attack model, in this section, we present the security analysis of the proposed scheme. We indicate that our scheme not only can provide confidentiality and anonymity, but also is secure against the

a. Message analysis;
b. Impersonation;
c. Modification;
d. Injection;
e. Replay;
f. DoS;
g. Insider;
h. SM memory modification attacks.

The details are as follows.

4.1 Providing Confidentiality, Preserving Anonymity, and Message Analysis Attack Resistance

In our scheme, an attacker $\mathscr{A}$ (either external, internal, or global), who is eavesdropping the communication channels, can get access to $\{EID_i, M_j^i, V_j^i\}$ and $\{ID_{NG}, M_k^i, V_k^i, CT_{NG}\}$ messages. In these messages, ID_{NG} and CT_{NG} are public parameters, V_j^i and V_k^i are two hash outputs, and M_j^i and M_k^i are two encrypted values using the shared key of SM_i and NG. Therefore, because of the one-way property of hash function, $\mathscr{A}$ cannot achieve D_j^i and CM_k^i from V_j^i and V_k^i. Moreover, having access to M_j^i or M_k^i, he/she cannot extract or recover the consumption reports D_j^i and control messages CM_k^i without knowing the K_i^{SM} or K_i^{NG}. The K_i^{SM} and K_i^{NG} are also kept secure using the secret keys of SM_i and NG. Hence, the proposed scheme provides confidentiality and is secure against the message analysis attack. In addition, since the EID_i is the encrypted value of identifier, $\mathscr{A}$ cannot identify the identity of SMs without knowing the private key of NG.

4.2 Impersonation, Modification, and Injection Attacks Resistance

In the proposed scheme, when the SM_i wants to send its usage report D_j^i, it first computes the $h(D_j^i \oplus T_j \oplus ID_i)$, and then sends $\{EID_i, M_j^i, V_j^i\}$ to NG. If the adversary $\mathscr{A}$, either external ε, internal $\jmath$, or global φ, tries to impersonate SM_i and send a forgery message by altering M_j^i, he/she will not be able to compute the proper V_j^i. Therefore, when NG checks the equality of $h(D_j^i \oplus T_j \oplus ID_i) = V_j^i$, it can detect any tampering. By checking this equation, NG becomes certain that the received message is from the real intended SM_i and nobody has modified the D_j^i. Same strategy is done when the NG sends a control message to SM_i. As soon as SM_i checks the equivalence of $h(ID_{NG} \| CM_k^i \| CT_{NG}) = V_k^i$, it not only ensures that the message has not been altered during the transfer, but also becomes sure that the message has

Table 2: Features comparison

Scheme	F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8	F_9	F_{10}	F_{11}	F_{12}	F_{13}	F_{14}	F_{15}	F_{16}	F_{17}
[20]	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No	Yes	Yes	No	No	No	Yes
[21]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes	No	Yes	Yes	No	Yes	No	No
[10]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes	No	Yes	Yes	No	Yes	No	No
[28]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No	Yes	No	No	Yes	No	Yes
[19]	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	Yes	No	No	Yes	No	No	No	Yes
Proposed	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes

F_1. Providing confidentiality F_2. Near real-time authentication F_3. Injection attack resistance F_4. Analysis attack resistance F_5. Modification attack resistance F_6. Replay attack resistance F_7. DoS attack detection capability F_8. Providing anonymity F_9. Considering *SM* memory modification attack resistance F_{10}. Formal verification/proof F_{11}. Extensive hardware implementation on different suitable testbeds F_{12}. Low communication overhead F_{13}. Low computational overhead F_{14}. Being extremely lightweight F_{15}. Presenting key agreement details F_{16}. Presenting the details of shared key storage and retrieval F_{17}. Presenting details of usage reports transmission.

been sent from the authorized *NG*. With the same reason, the scheme is secure against the message injection/pollution attack.

4.3 Replay Attack Resistance

By the employment of T_j in the computation of V_j^i, the messages from *SM*s to *NG* will be kept secure against the replay attack. Further, for the messages from *NG* to *SM*s, when a message is received by *SM*, it first generates a fresh timestamp, then, compares its generated timestamp with the received one. If the elapsed time is shorter than predefined maximum transmission delay, it will accept the message as a non-repeated one. Hence, the proposed scheme can properly withstand the replay attack.

4.4 DoS Attack Resistance

Since in the proposed scheme any tampering on (a) exchanging messages, (b) *SM* flash memory, and (c) *NG* database can be detected very soon, $\mathscr{A}$ cannot perform *DoS* attack. For the exchanging messages or *SM* flash memory, the tampering is revealed when equations $h(D_j^i \oplus T_j \oplus ID_i) = V_j^i$ and $h(ID_{NG} \| CM_k^i \| CT_{NG}) = V_k^i$ are checked. Additionally, as will be stated in the next part, if the internal adversary $\jmath$, who has access to the *NG* database, changes even one field of a record, *NG* will be informed very soon.

4.5 Insider Attack Resistance

In our proposed scheme, an insider adversary $\jmath$ cannot access the confidential data nor can perform the *DoS* attack. Since in the proposed scheme, the confidential data are saved as encrypted, the insider attacker $\jmath$ is not a privileged attacker. As a result, he/she cannot perform any special attack by having access to the *NG* database. Since in the *NG* database only the encrypted form of K_i is saved, $\jmath$ cannot get access to shared keys and if he/she tries to alter a field of a record, the tampering will be detected as soon as the equivalence of $h(K_i^{NG} \| ID_i) = H_i^{NG}$ is checked.

4.6 Memory Modification Attack Resistance

In the proposed scheme, since K_i is obfuscated using the Exclusive-OR operation, it cannot be meaningfully modified without having the secret key of the *SM*. Therefore, this scheme can withstand the *SM* memory modification attack. As stated in part 4.4, at the worst case scenario, the adversary φ who has gained access to the *SM* memory, cannot even perform *DoS* attack.

5 Automatic Formal Verification

In order to ensure that none of the usage reports or control messages can be accessed by an adversary and the impersonation or replay attack cannot take place, we have used a well-known and popular automatic protocol verifier called ProVerif [8]. Figure 3 shows the obtained output from this tool.

The first two results are the results of two injective correspondence that assures SM_i has really executed the protocol with *NG* and vice versa and also the received messages by each of these two entities are fresh. Therefore, these two results prove the resistance of the protocol against impersonation and replay attacks. In the ProVerif, proving the reachability properties is among the most basic capabili-

```
Output: root@ubuntu:~/Proverif/proverif1.94pl1# ./proverif SMandNG.pv
-- Query inj-event(endNG) ==> inj-event(startNG)
Completing...
Starting query inj-event(endNG) ==> inj-event(startNG)
RESULT inj-event(endNG) ==> inj-event(startNG) is true.
-- Query inj-event(endSMi) ==> inj-event(startSMi)
Completing...
Starting query inj-event(endSMi) ==> inj-event(startSMi)
RESULT inj-event(endSMi) ==> inj-event(startSMi) is true.
-- Query not attacker(Ki[])
Completing...
Starting query not attacker(Ki[])
RESULT not attacker(Ki[]) is true.
-- Query not attacker(CMik[])
Completing...
Starting query not attacker(CMik[])
RESULT not attacker(CMik[]) is true.
-- Query not attacker(Dij[])
Completing...
Starting query not attacker(Dij[])
RESULT not attacker(Dij[]) is true
```

Figure 3: The results of analysing the proposed protocol using Proverif

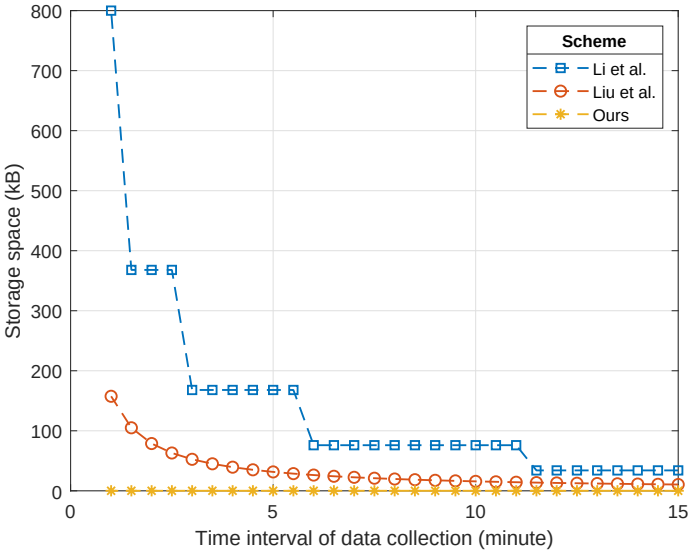


Figure 4: Storage space comparison for different time intervals

ties that lets us to check whether a term can be accessed by an attacker or not. The last three results are the results of such queries that indicate the attacker cannot obtain K_i, CM_k^i, and D_j^i. Therefore, the achieved results prove the secrecy of shared keys, control messages, and usage reports.

6 Efficiency Analysis and Experimental Study

In this section, we compare our proposed anonymous lightweight communication protocol with the related ones. Our comparative analysis shows better performance in terms of

a. Storage;
b. Communication;
c. Computational costs.

In the following, we present the detailed discussion.

6.1 Storage Space

Since only Liu *et al.* [20] and Li *et al.* [19] have discussed the storage space, in this section, we compare the proposed scheme with these two.

In Liu *et al.*'s scheme [20], *SM* needs to store r_j, R_j, and C_j, where $j = 1, 2, \ldots, 96$. The required storage space for r_j is 128×96 bits, the needed storage space for R_j is 256×96 bits, and the required storage space for C_j is 512×96 bits. Hence, the total required storage space is 10.5 kB.

In Li *et al.*'s scheme [19], the *SM* needs to store r_j, C_j, and API_j, where $j = 1, 2, \ldots, 128$. Considering the length of each random number to be 128 bits, the storage space of r_j is 128×128 bits, the storage space of $C_j = Enc_{k_i}(r_j \| TS_j)$ is $2 \times 128 \times 128$ bits, and the storage space of API_j, where each API_j contains seven hash values, is $7 \times 256 \times 128$ bits. Thus, the total required storage space is 34 kB.

In comparison to the previous schemes, in our scheme, the *SM* only needs to store E_i^{SM} and EID_i. Therefore, the required storage space is significantly reduced to only $256 + 256 = 512$ bits. Table 3 shows the comparison. Further, the required storage space for different time intervals of data collection, from 1 to 15 minutes, is illustrated in Figure 4.

Table 3: Storage space

[20]	[19]	Proposed
10.5 kB	34 kB	512 bits

6.2 Communication Cost

For Liu *et al.*'s scheme [20], the communication cost includes the encrypted r_{NG} which is 512 bits, 96 encrypted coefficients of the "f" function that are 128×96 bits, the reports $\{ID_i \| C_j \| S_j\}$ that are $(128 \times 96) + (512 \times 96) + (256 \times 96)$ bits, and the control messages $\{ID_{NG} \| M_k^1 \| M_k^2\}$ that are $(128 \times 4) + (256 \times 4) + (256 \times 4)$ bits. Therefore, the total communication cost is 12.375 kB.

For Li *et al.*'s scheme [19], the communication cost includes the encrypted root node value which takes 256 bits and $\{U_i \| C_j \| S_j \| API_j\}$ reports which

Table 4: Daily Communication Cost

[20]	[21]	[10]	[28]	[19]	Proposed
12.37 kB	7.81 kB	7.81 kB	8.12 kB	27.03 kB	7.81 kB

are sent from *SM* to *NG*. As a result, the total communication cost is $(256) + (128 \times 96) + (2 \times 128 \times 96) + (128 \times 96) + (7 \times 256 \times 96)$ bits = 27.03125 kB.

For Uludag *et al.*'s scheme [28], the communication cost includes $\{ID_j, SKE(K_{MD_j}^{DC_i}, T \| PRODATA \| HASH(DK, PRODATA))\}$ and $\{ID_{DC_i}, SKE(GK_i, SIGN(PO, COMD) \| COMD)\}$ messages. Therefore, the total communication cost is $(128 \times 96) + (128 \times 96) + (128 \times 96) + (256 \times 96) + (128 \times 4) + (1024 \times 4) + (128 \times 4)$ bits = 8.125 kB.

For Mahmood *et al.*'s [21] and Fouda *et al.*'s [10] schemes, the communication cost includes $\{ID_i, E_{K_i}(M_i \| T_i \| HMAC_{K_i})\}$ and $\{ID_i, E_{K_{ij}}(M_i \| t_{if} \| HMAC_{K_{ij}}(M_i))\}$ messages (and same command messages), respectively. As a result, the total communication cost of both is $(128 \times 96) + (128 \times 96) + (128 \times 96) + (256 \times 96) + (128 \times 4) + (128 \times 4) + (128 \times 4) + (256 \times 4)$ bits = 7.8125 kB.

The communication cost of our scheme includes the reports $\{EID_i, M_j^i, V_j^i\}$ that are $(256 \times 96) + (128 \times 96) + (256 \times 96)$ bits and the control messages $\{ID_{NG}, M_k^i, V_k^i, CT_{NG}\}$ which are $(128 \times 4) + (128 \times 4) + (256 \times 4) + (128 \times 4)$ bits. Thus, the total communication cost of our scheme is 7.8125 kB. Note that if M_k^i contains the updated encrypted identifier, its size will be increased to 256 bits. Table 4 shows the communication cost comparison. In addition, the communication cost for different time intervals of data collection is illustrated in Figure 5.

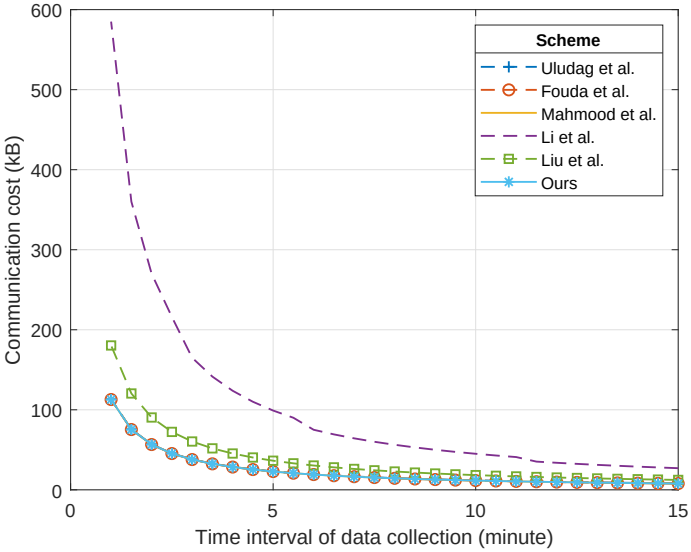


Figure 5: Communication cost comparison for different time intervals

6.3 Computational Cost

In this paper, in order to obtain the *SM* side computational cost of our scheme and the related ones, we have implemented different cryptographic operations on two testbeds. First is an *AVR* microcontroller called ATmega2560 which has 256 kB flash memory, 8 kB SRAM, 4 kB EEPROM, and clock speed of 16 MHz. Second is an *ARM* Cortex-M3 microcontroller called AT91SAM3X8E that has 512 kB flash memory, 96 kB SRAM, and clock speed of 84 MHz. The results have been achieved by utilization of the cryptographic library of ArduinoLibs [1]. On *AVR*, The *RSA* signature verification takes 670 ms; for the *AES*-256 encryption/decryption, 228.96 μs are spent for setting the key, 46.88 μs for encrypting each byte, and 90.05 μs for decrypting each byte; and for the hash operation, 43.89 μs are spent per byte. On *ARM*, The *RSA* signature verification takes 34 ms; for the *AES*-256 encryption/decryption, 46.97 μs are spent for setting the key, 8.04 μs for encrypting each byte, and 14.73 μs for decrypting each byte; and for the hash operation, 1.2 μs are spent per byte. Table 5 shows the execution time of cryptographic operations and Table 6 indicates the comparative computational cost. In Table 6, T_h, T_H, T_{Enc}, T_{Dec}, T_{Rnd}, T_{Pol}, and T_{Ver} are the execution time of one-way hash operation, *HMAC* operation, symmetric encryption, symmetric decryption, random generation, polynomial generation, and *RSA* signature verification, respectively. Moreover, Figures 6 and 7 depict the computational cost comparison for different time intervals of data collection, from 1 to 15 minutes, on *AVR* and *ARM*, respectively.

7 Conclusion

Recently, a number of lightweight communication schemes have been proposed to be employed in the context of smart grid. Nevertheless, most of them are not anonymous and some cannot resist the well-known attacks like the pollution attack. Therefore, in this paper, to remedy the existing challenges, we have proposed an efficient anonymous communication protocol that can properly withstand the common attacks. Moreover, we have implemented the

Table 5: Execution Time Cryptographic Operations on AVR and ARM

Operation	ATmega2560	AT91SAM3X8E
AES-256 ECB Setting the Key	228.96 μs	46.97 μs
AES-256 ECB Encryption (16 Bytes)	750.08 μs	128.64 μs
AES-256 ECB Decryption (16 Bytes)	1440.8 μs	235.68 μs
SHA256 (16 Bytes)	702.24 μs	19.2 μs
HMAC KEY Setup	2836 μs	81 μs
Polynomial Generation	160 ms	10 ms
Random Generation	12 ms	80 μs
RSA Signature Verification	670 ms	34 ms

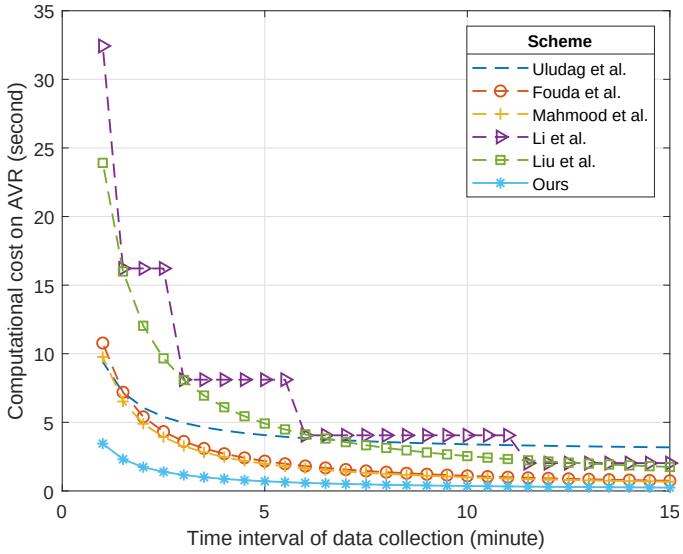


Figure 6: Computational cost comparison for different time intervals on AVR

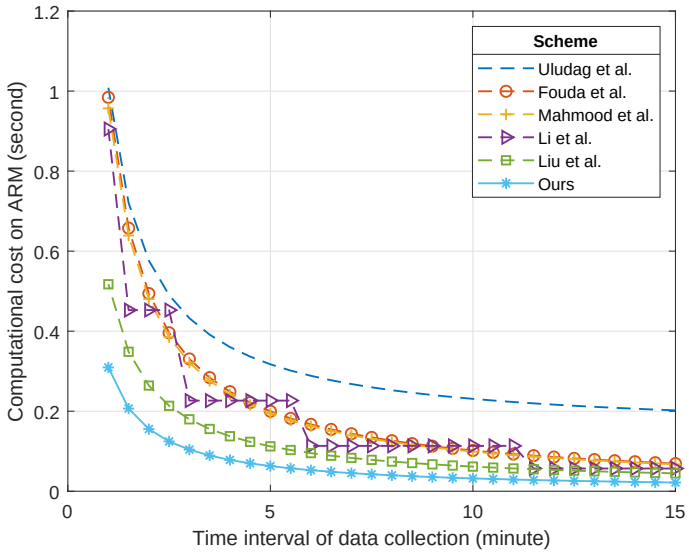


Figure 7: Computational cost comparison for different time intervals on ARM

cryptographic operations on both *AVR* and *ARM* and have compared our scheme with the related ones based on the obtained results on these two hardware. The achieved results indicate the superiority of the proposed scheme in terms of storage, communication, and computational costs. We hope that the presented results of this paper be useful for future researches in this field.

References

[1] "ArduinoLibs: Cryptographic library," 2018. (`http://rweather.github.io/arduinolibs/crypto.html`).

[2] D. Abbasinezhad-Mood, M. Nikooghadam, "An anonymous ECC-based self-certified key distribution scheme for the smart grid," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 10, pp. 7996–8004, 2018.

[3] D. Abbasinezhad-Mood, M. Nikooghadam, "Design and extensive hardware performance analysis of an efficient pairwise key generation scheme for smart grid," *International Journal of Communication Systems*, 2018. (`https://onlinelibrary.wiley.com/doi/abs/10.1002/dac.3507`)

[4] D. Abbasinezhad-Mood, M. Nikooghadam, "An ultra-lightweight and secure scheme for communications of smart meters and neighborhood gateways by utilization of an ARM Cortex-M microcontroller," *IEEE Transactions on Smart Grid*, vol. 9, no. 6, pp. 6194-6205, 2017.

[5] D. Abbasinezhad-Mood, M. and Nikooghadam, "Efficient anonymous password-authenticated key exchange protocol to read isolated smart meters by utilization of extended chebyshev chaotic maps," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 11, pp. 4815-4828, 2018.

[6] D. Abbasinezhad-Mood, and M. Nikooghadam, "Design and hardware implementation of a security-enhanced elliptic curve cryptography based lightweight authentication scheme for smart grid communications," *Future Generation Computer Systems*, vol. 84, pp. 47–57, 2018.

Table 6: Daily computational cost

Scheme	[20]	[21]	[10]	[28]	[19]	Proposed
Operations	$96\,T_h$ $96\,T_{Enc}$ $1\,T_{Dec}$ $96\,T_{Rnd}$ $1\,T_{Pol}$	$100\,T_H$ $96\,T_{Enc}$ $4\,T_{Dec}$	$100\,T_H$ $96\,T_{Enc}$ $4\,T_{Dec}$	$96\,T_h$ $96\,T_{Enc}$ $4\,T_{Dec}$ $4\,T_{Ver}$	$255\,T_h$ $129\,T_{Enc}$ $128\,T_{Rnd}$	$200\,T_h$ $96\,T_{Enc}$ $4\,T_{Dec}$
Execution Time on ARM	$\approx$ 45 ms	$\approx$ 67 ms	$\approx$ 69 ms	$\approx$ 202 ms	$\approx$ 57 ms	$\approx$ 22 ms
Execution Time on AVR	$\approx$ 1.75 s	$\approx$ 673 ms	$\approx$ 740 ms	$\approx$ 3.18 s	$\approx$ 2.03 s	$\approx$ 246 ms

[7] D. Abbasinezhad-Mood, and M. Nikooghadam, "Efficient design and extensive hardware evaluation of an anonymous data aggregation scheme for smart grid," *Security and Privacy*, 2018. (`https://onlinelibrary.wiley.com/doi/full/10.1002/spy2.24`)

[8] B. Blanchet, V. Cheval, X. Allamigeon, and B. Smyth, *ProVerif: Cryptographic Protocol Verifier in the Formal Model*, 2010. (`http://prosecco.gforge.inria.fr/personal/bblanche/proverif`)

[9] K. Chatterjee, and L. Priya, "HKDS: A hierarchical key distribution scheme for wireless ad hoc network," *International Journal of Network Security*, vol. 20, no. 2, pp. 243–255, 2018.

[10] M. M. Fouda, Z. M. Fadlullah, N. Kato, R. Lu, and X. S. Shen, "A lightweight message authentication scheme for smart grid communications," *IEEE Transactions on Smart Grid*, vol. 2, no. 4, pp. 675–685, 2011.

[11] M. S. Hwang and C. C. Lee, "Research issues and challenges for multiple digital signatures," *International Journal Network Security*, vol. 1, no. 1, pp. 1–7, 2005.

[12] M. S. Hwang and T. Y. Chang, T. Yi, "Threshold signatures: Current status and key issues," *International Journal Network Security*, vol. 1, no. 3, pp. 123–137, 2005.

[13] M. S. Hwang and C. Y. Liu, "Authenticated encryption schemes: Current status and key issues," *International Journal Network Security*, vol. 1, no. 2, pp. 61–73, 2005.

[14] M. S. Hwang, C. C. Lee, and W. P. Yang, "An improvement of mobile users authentication in the integration enviromnents," *AEU-International Journal of Electronics and Communications*, vol. 56, no. 5, pp. 293–297, 2002.

[15] Y. Kabalci, "A survey on smart metering and smart grid communication," *Renewable and Sustainable Energy Reviews*, vol. 57, pp. 302–318, 2016.

[16] A. V. N. Krishna, A. H. Narayana, and 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.

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

[18] C. Li, H. Cheung, C. Yang, "Secure and efficient authentication protocol for power system computer networks," *International Journal of Network Security*, vol. 20, no. 2, pp. 337–344, 2018.

[19] H. Li, R. Lu, L. Zhou, B. Yang, and X. Shen, "An efficient Merkle-tree-based authentication scheme for smart grid," *IEEE Systems Journal*, vol. 8, no. 2, pp. 655–663, 2014.

[20] Y. Liu, C. Cheng, T. Gu, T. Jiang, X. Li, "A lightweight authenticated communication scheme for smart grid," *IEEE Sensors Journal*, vol. 16, no. 3, pp. 836–842, 2016.

[21] K. Mahmood, S. A. Chaudhry, H. Naqvi, T. Shon, and H. F. Ahmad, "A lightweight message authentication scheme for smart grid communications in power sector," *Computers & Electrical Engineering*, vol. 52, pp. 114–124, 2016.

[22] K. Mahmood, S. A. Chaudhry, H. Naqvi, S. Kumari, X. Li, and A. K. Sangaiah, "An elliptic curve cryptography based lightweight authentication scheme for smart grid communication," *Future Generation Computer Systems*, vol. 81, pp. 557–565, 2018.

[23] N. B. M. Mohan, A. S. N. Chakravarthy, and C. Ravindranath, "Cryptanalysis of design and analysis of a provably secure multi-server authentication scheme," *International Journal of Network Security*, vol.20, no. 2, pp. 217–224, 2018.

[24] N. T. Nguyen, H. D. Le, and C. C. Chang. "Provably secure and efficient three-factor authenticated key agreement scheme with untraceability," *International Journal of Network Security*, vol. 18, no. 2, pp. 335–344, 2016.

[25] R. Paspula, K. Chiranjeevi, and S. L. Kumar, "Hidden data transmission with variable DNA technology," *International Journal of Electronics and Information Engineering*, vol. 7, no. 2, pp. 96–106, 2017.

[26] S. K. Ravva, "Common private exponent attack on multi prime RSA," *International Journal of Electronics and Information Engineering*, vol. 7, no. 2, pp. 79–87, 2017.

[27] C. S. Tsai, C. C. Lee, and M. S. Hwang, "Password authentication schemes: Current status

and key issues," *International Journal Network Security*, vol. 3, no. 2, pp. 101–115, 2006.

[28] S. Uludag, K. S. Lui, W. Ren, and K. Nahrstedt, "Secure and scalable data collection with time minimization in the smart grid," *IEEE Transactions on Smart Grid*, vol. 7, no. 1, pp. 43–54, 2016.

[29] M. Wazid, A. K. Das, N. Kumar, and J. J. Rodrigues, "Secure three-factor user authentication scheme for renewable-energy-based smart grid environment," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 6, pp. 3144–3153, 2017.

Biography

Dariush Abbasinezhad-Mood received the B.Sc. degree in Computer Science from Payame Noor University, Mashhad, Iran, in 2009 and the M.Sc. degree in Secure Communications from Imam Reza International University, Mashhad, Iran, in 2016 with the first rank. Further, his M.Sc. thesis was elected as the top thesis. His research interests include Cryptography, Trust and Reputation based Systems, Authentication Protocols, Smart Grid Security, Wireless Sensor Networks, Internet of Things, and Embedded Systems. He is a reviewer for several well-known journals, such as IEEE Transactions on Smart Grid, IEEE Transactions on Industrial Informatics, IEEE Systems Journal, IEEE Internet of Things Journal, and IEEE Communications Letters.

Arezou Ostad-Sharif received the B.Sc. degree in Information Technology from Safahan University, Esfahan, Iran, in 2015 and M.Sc. degree in Information Security from Imam Reza International University, Mashhad, Iran, in 2017. Her M.Sc. thesis was elected as the top thesis in 2018. She hopes to advance her education in network security. Her research interest focuses on the security protocols for Wireless Sensor Networks, Internet of Things, Smart Grid, and Telecare Medical Information Systems.

Morteza Nikooghadam received the B.Sc. degree from university of Sadjad, Iran, in 2006, M.Sc. from the Shahid Beheshti University, Iran, in 2008, and Ph.D. from Shahid Beheshti University, Iran, in 2012. He is currently an assistant professor in the Department of Computer Engineering and Information Technology at Imam Reza International University, Mashhad, Iran. His research focuses on Data Security, Cryptography, and Sensor Network Security. His current research interests are Reconfigurable Architectures for multipliers under Galois Field $GF(2^m)$.

A PUF-based Group Key Transfer Protocol for Bluetooth

Sensen Li, Bin Yu, and Yicai Huang
(Corresponding author: Sensen Li)

Department of Computer Engineering, Zhengzhou Information Science and Technology Institute
Zhengzhou 450001, China
(Email: lss589@163.com)

(Received Dec. 5, 2017; Revised and Accepted Mar. 19, 2018; First Online Mar. 2, 2019)

Abstract

Group key is the basis for ensuring the security of bluetooth broadcast messages. Recently, aiming at establishing group key among resource-constrained devices, Hsu *et al.* and Piao *et al.* respectively proposed a group key transfer protocol, but our analysis demonstrates that the two protocols can neither satisfy the security requirements of bluetooth. Physical Unclonable Function (PUF) is a mapping relationship based on the physical characteristics of a given device, which has a broad application prospect in the field of information security. In this paper, we propose a PUF-based group key transfer protocol for bluetooth, which can establish the shared group key between bluetooth master device and slave devices on the condition of slave devices not storing any secret parameters. The proposed protocol not only resists the traditional attacks such as eavesdropping, tampering and replaying, but also protects the bluetooth devices from replication attack. Compared with related protocols, this protocol has a higher security and obviously decreases the computation, storage and communication overhead.

Keywords: Bluetooth; Group Key Transfer; Physical Unclonable Function; Replication attack; Traditional Attacks

1 Introduction

Bluetooth has become one of the main communication ways of wireless sensors for the features of low cost, low complexity and high reliability. Nowadays, bluetooth is being used in smart home, medical care, indoor positioning and many other areas [6, 20, 26]. With the popularization of its application, the security issues of bluetooth have attracted extensive attention and gradually become the key factor constraining its development in the high security requirements fields, such as finance and military [14].

In the application of sensor networks, the basic communication topology of bluetooth is the piconet consisting of a master and several slavers. Key agreement is a pivotal step for secure information interaction among bluetooth devices. The bluetooth specification [2] implements the establishment of a point-to-point link key between the master and slaver by defining the processes called pairing and binding. However, the specification does not provide a mechanism for establishing the group key. Adversaries can attack the bluetooth broadcast channel by eavesdropping, tampering, replaying, etc. Besides, they are able to capture the slaver, whose self-protection ability is poor, then extract secret parameters from the slaver's storage medium and replicate malicious devices. Consequently, for the purpose of ensuring the security of the data in the broadcast channel, a lightweight cryptographic protocol should be used to establish the group key shared by the master with multiple slavers. And the protocol should meet the link security requirements and protect the device from replication attack.

The traditional group key management protocols can be divided into two categories: Centralized protocols and distributed protocols. The centralized ones require a device with strong computing and storage capabilities, called KGC, to generate and distribute the group key, while the distributed protocols don't have an explicit KGC and the group key is obtained by all group members through negotiation. In the bluetooth piconet, the master usually has abundant resources, while the slavers are usually nodes with simple structure and limited resources. Therefore, the centralized protocols are more suitable for bluetooth and the master can act as KGC.

In recent years, many researchers have studied the group key management in wireless sensor networks. Harn *et al.* [9] pointed out that the traditional centralized protocols [7, 10, 18, 23] and distributed protocols [3, 5, 13, 15, 28] have the problems like high computational complexity and the prolonged time delay of setting up a group key, so they proposed a group key transfer protocol based on (t, n) secret sharing scheme. In Harn's protocol, KGC generates the group key and broadcasts related secret information to group members. When receiving the secret information, each authorized member needs

to calculate a t-degree interpolation polynomial to recover the group key. But Nam *et al.* [21] proved that Harn's protocol was unsecured. To improve Harn's protocol, Liu *et al.* [17] proposed a new protocol, which achieved the security at the cost of higher computation overhead. Based on the secret sharing scheme and factoring assumption, Hsu *et al.* [11] proposed an efficient group key transfer protocol for WSNs, whose communication and computation overheads are less than those of Harn's protocol, but unfortunately, this protocol can't achieve the claimed security for the reason that inside members can obtain the secret key shared by another group member with KGC. To fill the gap, Hsu *et al.* [12] improved their former protocol [11] by using the hash function to ensure the confidentiality of the key shared by group member and KGC. However, in this protocol, the KGC needs to perform $(t+1)$ times hash function, where t is the number of group members, so the computation overhead is too high to be suitable for the resource-constrained bluetooth devices. Piao *et al.* [24] employed a polynomial to implement group key transfer. The protocol is lightweight and simple, but it can be proved that this protocol can't guarantee the forward security and backward security of the group key. In addition, the protocols above are all establishing the group key on the basis of the secret key shared by each group member with KGC, so each device needs to store such secret information in its memory. However, [19, 25] pointed out this storage way can't resist the replication attack on the devices, especially for the resource-constrained devices. By capturing one authorized group member and extracting the point-to-point key from its device memory, attackers can easily recover all the group keys for communications the captured member has participated in.

PUF [22] is a special mapping relationship between the input challenges and the output responses. Similar with using the unique features of human body like fingerprint, iris and so on, the mapping is based on the intrinsic physical characteristics of the device, which can be expressed as hardware fingerprint. With the features of uniqueness, unclonability, unpredictability and lightweight [27], PUF can be applied in authentication, key generation and many other fields. Many researches [1, 4, 8, 16] have probed the application of PUF in resource-constrained devices and significantly improved the security of these devices. However, the existing researches mainly focus on key generation, device authentication and point-to-point key agreement. Up to now, PUF has not been used for establishing the group key.

In this paper, we propose a PUF-based bluetooth group key transfer protocol that obviously decreases the resource overhead and improves the security. In our protocol, PUF is adopted to hide the information related to group key and the resource-constrained slave devices don't need to store any secret parameters, which effectively prevents the replication attack as well as traditional link attacks. The rest of this paper is organized as follows: In the next section, we briefly review the related protocols and then analyze their security weaknesses respectively. In Section 3, we propose our PUF-based group key transfer protocol for bluetooth. In Section 4, we analyze the correctness and security of our protocol. Section 5 provides the performance evaluation of the proposed protocol. Concluding remarks are given in Section 6.

2 Related Protocols and Their Security Analysis

2.1 Analysis of Hsu's Protocol

Hsu *et al.* [11] proposed a group key transfer protocol with low resource overhead, but we found that the protocol couldn't resist the insider attack. This section firstly reviews this protocol briefly, then gives the specific attack method of the malicious inside members.

1) Protocol review: Hsu's protocol consists of three phases: KGC initialization, user registration, group key generation and distribution. We only briefly introduce the most important phase: Group key generation and distribution, the process is as follows.

   Step1. The initiator sends a list of target group members $\{1, \cdots, t\}$to KGC, as a group key transfer request.

   Step2. When receiving the request, KGC broadcasts the group list $\{1, \cdots, t\}$.

   Step3. Each participating member sends a random number $R_i (i = 1, \cdots, t)$ to KGC.

   Step4. KGC randomly selects a group key K_G and a random number R_0. KGC also computes $U_i = (K_G - K_i) \bmod m \ (i = 1, \cdots, t)$ and authentication code $Auth$, where $K_i = (v(x_i), \overrightarrow{r}) = R_0 + R_1 x_i + R_2 {x_i}^2 + \cdots + R_t {x_i}^t$ and x_i is the secret parameter shared by the member i and KGC. Then, KGC broadcasts $\{Auth, R_0, U_i\}$ $(i = 1, \cdots, t)$.

   Step5. Each participating member i, knowing the x_i, is able to compute $K_i = (v(x_i), \overrightarrow{r})$ and recover the group key $K_G = (U_i + K_i) \bmod m$. Then the member i uses the authentication code $Auth$ to check the correctness of K_G.

2) Feasible attack method: The malicious group member *eve* can obtain the secret parameter shared by the member *target* with KGC. The detailed attack process is as follows.

   As the initiator, *eve* firstly sends the group key transfer request$\{eve, target\}$ to KGC, then these three parties, KGC, *eve* and *target*, execute the group key transfer process above. During this process, *eve* can get the parameter U_{target} and have the ability to compute $K_{target} = K_{G1} - U_{target}$. Therefore, *eve* obtains the Equation (1).

$$K_{target} = R_0 + R_{eve} x_{target} + R_{target} x_{target}^2. \quad (1)$$

By repeating the above procedure, *eve* gets the following Equation (2).

$$K'_{target} = R'_0 + R'_{eve}x_{target} + R'_{target}x^2_{target}. \quad (2)$$

Using the public parameters $\{R_0, R'_0, R_{eve}, R'_{eve}, R_{target}, R'_{target}\}$, *eve* can recover the secret x_{target} by executing $(1) \times R'_{target} - (2) \times R_{target}$, as shown in Equation (3).

$$\begin{aligned} x_{target} &= [(K_{target} - R_0) \times R'_{target} - \\ &\quad (K'_{target} - R'_0) \times R_{target}]/ \\ &\quad (R_{eve} \times R'_{target} - R'_{eve} \times R_{target}). \end{aligned} \quad (3)$$

2.2 Analysis of Piao's Protocol

Piao *et al.* [24] used the polynomial to implement the secret distribution of group keys. The implementation process is simple, but our analysis shows that the protocol can't guarantee the forward security and backward security of the group key.

1) Protocol review: Each group member $i(i = 1, \cdots, t)$ firstly establishes the secret KEY_i shared with KGC by registering, and then performs the following steps.

   Step 1. KGC randomly selects a group key K_G and generates the related polynomial P, as shown in Equation (4).

$$\begin{aligned} P &= (x - KEY_1)(x - KEY_2)... \\ &\quad (x - KEY_t) + K_G \end{aligned} \quad (4)$$

   Then, KGC broadcasts the polynomial P to group members $\{1, \cdots, t\}$.

   Step 2. When receiving the polynomial, group member$i(i = 1, \cdots, t)$recovers the group key K_G by using the method shown in Equation (5).

$$\begin{aligned} K_G &= (x - KEY_1)(x - KEY_2) \cdots \\ &\quad (x - KEY_t) + K_G, \text{ where } x = KEY_i. \end{aligned} \quad (5)$$

2) Forward security analysis: Before joining the group, the node w can obtain the polynomialP_1, as shown in Equation (6), by monitoring the public channel. The constant term of this polynomial is $c_1 = (-1)^t \times KEY_1 \times ... \times KEY_t + K_{G1}$.

$$\begin{aligned} P_1 &= (x - KEY_1)(x - KEY_2)...(x - KEY_t) \\ &\quad + K_{G1}, w \notin \{1, \cdots, t\}. \end{aligned} \quad (6)$$

After the node w becomes a group member, the polynomial P_2, as shown in Equation (7), can be obtained during the group key distribution.

$$\begin{aligned} P_2 &= (x - KEY_1)(x - KEY_2)...(x - KEY_t) \\ &\quad \times (x - KEY_w) + K_{G2}. \end{aligned} \quad (7)$$

The constant term of P_2 is $c_2 = (-1)^{t+1} \times KEY_1 \times ... \times KEY_t \times KEY_w + K_{G2}$. Using the secret KEY_w, group member w is able to recover the group key K_{G2}. And Equation (8) shows the method to get the forward group keyK_{G1}, which shouldn't have been known by node w.

$$KG_1 = c_1 + \frac{c_2 - KG_2}{KEY_w} \quad (8)$$

3) Backward security analysis: In the process of group key distribution, group member i obtains the polynomial P_3, as shown in Equation (9).

$$\begin{aligned} P_3 &= (x - KEY_1)(x - KEY_2)...(x - KEY_t) \\ &\quad + K_{G3}, i \in \{1, \cdots, t\}. \end{aligned} \quad (9)$$

The constant term of P_3 is $c_3 = (-1)^t \times KEY_1 \times ... \times KEY_t + K_{G3}$. And group member i uses KEY_i to recover the group key K_{G3}.

After leaving the group, node i can get the polynomial P_4, as shown in Equation (10), by monitoring the channel. And its constant term is $c_4 = (-1)^{t-1} \times KEY_1 \times ... \times KEY_{i-1} \times KEY_{i+1} \times ... \times KEY_t + K_{G4}$.

$$\begin{aligned} P_4 &= (x - KEY_1)...(x - KEY_{i-1}) \\ &\quad \times (x - KEY_{i+1})...(x - KEY_t) \\ &\quad + K_{G4}. \end{aligned} \quad (10)$$

Utilizing the method shown in Equation (11), node i can recover the backward group key KG_4.

$$KG_4 = c_4 + \frac{c_3 - KG_3}{KEY_i} \quad (11)$$

From the above analysis, we can see that Piao's protocol can't guarantee the forward and backward security of the group key, so its security needs to be improved.

2.3 Analysis of Other Protocols

The existing group key transfer protocols all need to store sensitive parameters in device memory: the protocols based on symmetric cryptography need to save symmetric key information, while the ones based on public key cryptography need to save the device's private key.

With a simple structure and limited resources, the bluetooth slave device is usually difficult to achieve self-protection. If attackers capture a slave device, they may extract sensitive information, such as keys or algorithm parameters, using the method shown in Figure 1. Utilizing these information, attackers can not only figure out the group key to decrypt broadcast messages, but also replicate a similar node to forge the identity of the legal device and deliver false information, which poses a tremendous threat to the security of the bluetooth.

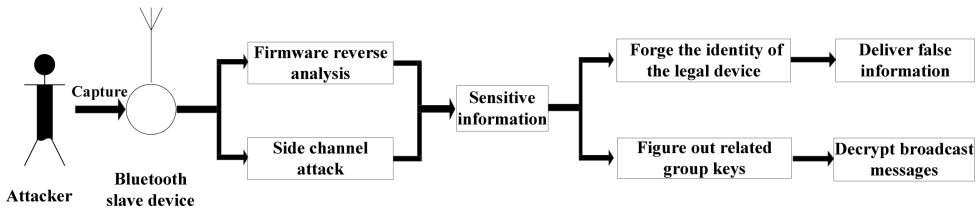


Figure 1: Group key establishment process

3 The Proposed Protocol

This section firstly gives the basic model of the bluetooth network, according to which we propose the attack assumptions. Then, the detailed process of the PUF-based group key transfer protocol is introduced.

3.1 Network Model and Attack Assumptions

The participants of the proposed protocol include a bluetooth master device (*Master*) and multiple bluetooth slave devices (*Slavers*). Playing the role of KGC, *Master*, with strong computing, storage and self-protection ability, is responsible for the generation and distribution of group keys; while *Slaver* is usually an information gathering node with a simple structure and limited resources, which is vulnerable to replication attack for its poor self-protection ability.

According to the characteristics of different devices, we put forward the following attack assumptions.

1) Adversaries can not only attack the network by traditional manners, such as eavesdropping, tampering and replaying, but also capture the bluetooth devices in the open environment.

2) Once the *Slaver* is captured, adversaries can obtain all the secret information in the device's storage medium by replication attack.

3) The *Master* has sufficient self-protection capability to resist replication attack.

3.2 Protocol Details

The proposed group key transfer protocol consists of two processes: initialization and group key establishment. The initialization process accomplishes the selection of basic parameters and the registration of *Slavers*. In group key establishment process, *Master* generates the group key and distributes it to *Slavers* based on the PUF.

1) Initialization: The *Master* randomly chooses two secure primes, p and q, and computes $n = pq$. Here, n is made publicly known. Then, the *Master* randomly selects $C_1, C_2 \in \mathbb{Z}_n^*$ as the challenges of PUF. All computations of the proposed protocol are performed in $\mathbb{Z}_n^*$.

Each device is required to register to the *Master* when joining the network. And this process is performed in a secure manner, for example, it can be achieved with the help of the network administrator. For the device $Slaver_i$, *Master* uses the PUF of $Slaver_i$ to obtain the unique responses$R_{i,1}$ and $R_{i,2}$, where $R_{i,1} = \mathrm{PUF}_i(C_1)$ and $R_{i,2} = \mathrm{PUF}_i(C_2)$. Then, *Master* stores the tuple $(i, R_{i,1}, R_{i,2})$.

After the initialization process is completed, $Slaver_i$ destroys its one-time PUF external interface for the purpose that adversaries outside the device chip have no access to challenge-response pairs of the PUF.

2) Group key establishment: The group key establishment process contains five steps and the detailed description is as follows.

Step1. The initiator, *init*, sends the list of target group members $\{1, 2, \cdots, t\}$, $init \in [1,\ t]$ to *Master*, as a group key establishment request.

Step2. For each group member $Slaver_i$, $i = 1, \cdots, t$, *Master* randomly selects a secret parameter $x_i \in \mathbb{Z}_n^*$. Then, *Master* broadcasts the message$\{C_1, C_2, i, R_{i,1} \oplus x_i\}$.

Step3. When receiving the message, $Slaver_i$ uses its PUF to acquire the responses $R_{i,1} = \mathrm{PUF}_i(C_1)$ and $R_{i,2} = \mathrm{PUF}_i(C_2)$, then figures out the secret parameter x_i. After that, $Slaver_i$ generates a random number $y_i \in \mathbb{Z}_n^*$ and sends the message $\{R_{i,2} \oplus y_i\}$ to Master.

Step4. Knowing the response$R_{i,2}$, *Master* is able to recover the secret y_i generated by each group member. *Master* randomly selects the group key $K_G \in \mathbb{Z}_n^*$ and constructs the t-degree equation $a_1 x + a_2 x^2 + \cdots + a_t x^t = K_G$, whose roots are $\{k_1, k_2, \cdots, k_t\}$ and $k_i = x_i \oplus rev(y_i)$. (Here, $rev(y_i)$ represents the reverse order of the binary sequence y_i and the method to get the equation coefficients $\{a_1, a_2, \cdots, a_t\}$ is introduced in Section 4.1.) Then, *Master* computes the group key authentication code $Auth = \mathrm{H}(a_1||a_2||\cdots||a_t||K_G)$ and broadcasts the message $\{a_1, a_2, \cdots, a_t, Auth\}$.

Step5. Knowing the secret parameters x_i and y_i, $Slaver_i$ is able to figure out a equation root $k_i = x_i \oplus rev(y_i)$. And the group key

K_G can be recovered by computing $K_G = f(k_i)$, where $f(x) = a_1x + a_2x^2 + \cdots + a_tx^t$. Then, $Slaver_i$ checks the validity of K_G by computing $\mathrm{H}(a_1||a_2||\cdots||a_t||K_G)$ and comparing whether the hash value is equal to $Auth$. Figure 2 shows the process of group key establishment for a group whose members are $\{Slaver_A, Slaver_B, Slaver_C\}$.

After the communication is completed, each group member deletes the group key and other related parameters in the device. In other words, *Slavers* in the open environment don't store any secret parameters.

4 Protocol Analysis

In this section, we analyze the correctness and security of the proposed protocol respectively. Then, the security comparison between this protocol and other related protocols is listed.

4.1 Correctness Analysis

In our protocol, *Master* is required to figure out the proper coefficients $\{a_1, a_2, \cdots, a_t\}$ in order to make the roots of the equation $a_1x + a_2x^2 + \cdots + a_tx^t = K_G$ be $\{k_1, k_2, \cdots, k_t\}$, where $k_i = x_i \oplus rev(y_i)$. It's a key issue to prove that no matter what the values of $\{k_1, k_2, \cdots, k_t\}$ are, the coefficients are surely existed, which is vital to the correctness of the proposed protocol. The detailed proof process is as follows.

By substituting $\{k_1, k_2, \cdots, k_t\}$ into the equation $a_1x + a_2x^2 + \cdots + a_tx^t = K_G$, we can obtain the linear equations, whose unknowns are $\{a_1, a_2, \cdots, a_t\}$, as shown in Equation (12).

$$\begin{cases} k_1a_1 + k_1^2a_2 + \cdots + k_1^ta_t = K_G \\ k_2a_1 + k_2^2a_2 + \cdots + k_2^ta_t = K_G \\ \cdots \\ k_ta_1 + k_t^2a_2 + \cdots + k_t^ta_t = K_G \end{cases} \tag{12}$$

The coefficient matrix of the linear equations is A, as shown in Equation (13).

$$A = \begin{bmatrix} k_1 & k_1^2 & \cdots & k_1^t \\ k_2 & k_2^2 & \cdots & k_2^t \\ \cdots & \cdots & \cdots & \cdots \\ k_t & k_t^2 & \cdots & k_t^t \end{bmatrix} \tag{13}$$

Through calculation, we can see that the determinant of matrix A is $|A| = (k_1k_2\cdots k_t) \prod_{1\le j<i\le n} (k_i - k_j)$. On account that x_i and y_i are randomly generated and $k_i = x_i \oplus rev(y_i)$, it is reasonable to think that when $i \neq j$, $k_i \neq k_j$ and then $|A| \neq 0$. According to the *Cramer Rule*, when $|A| \neq 0$, the linear Equation (12) has the unique solution $\{a_1, a_2, \cdots, a_t\}$.

For the reason that $\{k_1, k_2, \cdots, k_t\}$ are the roots of the equation $a_1x + a_2x^2 + \cdots + a_tx^t = K_G$, we can conclude this equation is equivalent to Equation (14). And the parameters $\{a_1, a_2, \cdots, a_t\}$ can be obtained by expanding Equation (14).

$$a_t(x - k_1)(x - k_2)\cdots(x - k_t) = 0. \tag{14}$$

When receiving the $\{a_1, a_2, \cdots, a_t\}$, each group member $Slaverr_i$ is able to recover the group key $K_G = f(k_i)$ by substituting $k_i = x_i \oplus rev(y_i)$ into the function $f(x) = a_1x + a_2x^2 + \cdots + a_tx^t$.

4.2 Security Analysis

The security of the proposed protocol depends on the confidentiality of PUF's challenge-response pairs (CRPs), which can be guaranteed by the unclonability and unpredictability of PUF. That is to say, the following security conditions are true.

1) The unclonability of PUF: For a given PUF, it's infeasible to construct a PUF' by physical manners enabling $PUF'(c) = PUF(c)$ for any challenge signal c.

2) The unpredictability of PUF: Given a CRPs set $L = \{(c_i, PUF(c_i))|i = 1, 2, \cdots l\}$, the probability to predict the response $PUF(c_x)$ is negligible, where c_x is a random challenge signal and $(c_x, PUF(c_x)) \notin L$.

Based on the above security conditions, we prove the security of the proposed protocol by the following two theorems.

Theorem 1. *The proposed protocol can guarantee the freshness, confidentiality, authentication, forward security and backward security of the group key.*

1) Key freshness. When receiving the group key establishment request from the initiator, *Master* randomly selects the group key K_G and secretly distributes K_G to group members by constructing the equation $a_1x + a_2x^2 + \cdots + a_tx^t = K_G$. The roots of the equation are $\{k_1, k_2, \cdots, k_t\}$ and $k_i = x_i \oplus rev(y_i)$, where x_i and y_i are the random parameters selected by *Master* and $Slaver_i$ respectively. Therefore, using random numbers as the fresh factor, the proposed protocol can ensure the freshness of the group key.

2) Key confidentiality. In the process of group key distribution, the parameters transmitted in the public channel include $\{C_1, C_2, R_{i,1} \oplus x_i, R_{i,2} \oplus y_i, a_1, a_2, \cdots, a_t, Auth\}$. Because of the unclonability and unpredictability of PUF, the attacker can't figure out the corresponding response signals R_1 and R_2 as well as the equation's root $k_i = x_i \oplus rev(y_i)$. In addition, due to the one-way nature of the hash function, the attacker can't obtain any secret information from the authentication code $Auth$. For the function $f(x) = a_1x + a_2x^2 + \cdots + a_tx^t$, it is impracticable to calculate $K_G = f(k_i)$ only by using the coefficients $\{a_1, a_2, \cdots, a_t\}$. In conclusion, the group key in the protocol is confidential.

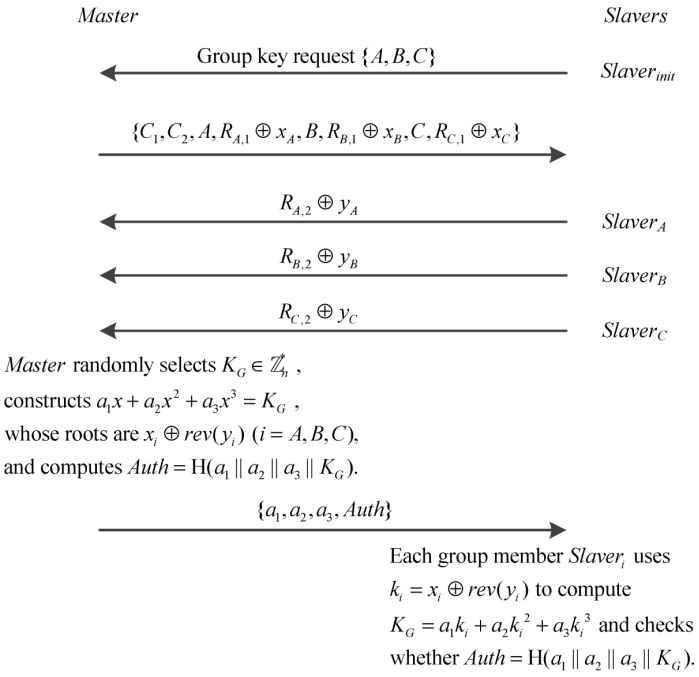


Figure 2: Group key establishment process

3) Key authentication. Group members use the authentication code $Auth$, which is the hash value of equation coefficients $\{a_1, a_2, \cdots, a_t\}$ and the group key K_G, to judge the validity of the group key. The attackers outside the group can't correctly forge the authentication code for the freshness and confidentiality of K_G. Any group member, knowing the group key, also can't forge the authentication code without being detected, for the reason that the coefficients $\{a_1, a_2, \cdots, a_t\}$ are related to the secrets shared between each group member and *Master*. Consequently, as long as the calculated hash value $\mathrm{H}(a_1||a_2||\cdots||a_t||K_G)$ is equal to the received $Auth$, group members can believe K_G is generated by *Master*.

4) Forward security and backward security. Since the group key is randomly generated by *Master* and the parameters of the equation are fresh, group keys generated at different times have nothing to do with each other. Therefore, the unauthorized device can't infer other group keys through a group key. In other words, the protocol can guarantee the forward security and backward security of the group key.

Theorem 2. *While resisting the attackers outside the group, the proposed protocol provides protection against the insider attack, which is initiated by the malicious inside group member.*

Proof. According to *Theorems 1*, the attackers outside the group can't obtain the group key by eavesdropping the bluetooth channel, so the protocol can resist the attacks from malicious outside devices. Different from the attackers outside the group, the malicious inside attackers are authorized to know the group keys and their attack attempt is to recover the CRPs of another member' s PUF. □

In order to facilitate the representation, we assume that the malicious device in the group is $Slaver_{eve}$ and its attack target is $Slaver_{target}$. $Slaver_{eve}$ can obtain the polynomial $a_1 k_{target} + a_2 k_{target}^2 = KG \pmod n$, where $k_{target} = x_{target} \oplus rev(y_{target})$, by sending the group key establishment request $\{eve, target\}$ to *Master*. And sending the same request again, $Slaver_{eve}$ will get a similar but different polynomial $a_1{}' k'_{target} + a_2{}' (k'_{target})^2 = KG' \pmod n$. Since all the parameters in the polynomial are fresh, there is no correlation between the polynomials. In other words, $Slaver_{eve}$ can only get the secret parameter k_{target} by solving the polynomial $a_1 k_{target} + a_2 k_{target}^2 = KG \pmod n$. But Harn *et al.* [9] pointed out that this is an intractable problem due to the *Factoring Assumption*. For our protocol, even if $Slaver_{eve}$ has the ability to solve the factorization problem and figure out k_{target}, he still obtains nothing about the CRPs of $Slaver_{target}$, $(C_1, R_{target,1})$ and $(C_2, R_{target,2})$, since $k_{target} = x_{target} \oplus rev(y_{target})$. Therefore, the proposed protocol provides protection against the insider attack while resisting the attackers outside the group.

For the reason that the bluetooth slave devices don't need to store any secret parameters and the PUF has unclonability and unpredictability, the proposed protocol can not only resist the traditional attacks such as eavesdropping, tampering and replaying, but also effectively prevent the possible replication attack on the slave devices. The security comparison between the proposed

protocol and other related protocols is shown in Table 1.

5 Performance Evaluation

This section firstly analyzes the performance of the proposed group key transfer protocol from three aspects, computation, communication and storage overhead. Then, we compare our protocol with Liu's protocol [17], which is more secure than other existing protocols. The proposed protocol consists of two processes: initialization and group key establishment. This section mainly analyzes the resource overhead of group key establishment process, since the former process only needs to perform one time while the latter process performs as long as *Master* has received the "group key establishment request".

In the bluetooth network, the master device, *Master*, usually has strong computation, communication and storage capabilities, while the slave devices, *Slavers*, only possess limited resources. Therefore, it is more important to consider reducing the resource overhead of the slave devices when designing the protocol.

For the convenience of description, we assume that the slave devices in the network are $\{1, 2, \cdots, m\}$, the group members are $\{1, 2, \cdots, t\}(t \leq m)$, the length of each parameter in $\mathbb{Z}_n^*$ is $|n|$ and the length of the hash is $|H|$.

5.1 Computation Overhead

We use T_M, T_I and T_H, respectively, to represent the time required to perform modular multiplication, modular inversion and hash. Compared to T_M, T_I and T_H, the time required for other operations, such as modular addition and subtraction, can be ignored [11].

In the proposed protocol, *Master* needs to perform $\frac{1}{2} \times t \times (t+1)$ times modular multiplication and one hash operation, so its computation overhead is $\frac{1}{2} \times t \times (t+1) \times T_M + t \times T_H$, while the computation overhead of $Slave_i (i = 1, 2, \cdots, t)$ is $(2t-1) \times T_M + T_H$ for performing $(2t-1)$ times modular multiplication and one hash operation. In the same way, we can get that, in Liu's protocol, the computation overhead of *Master* is $t \times (t+1) \times t \times (T_M + T_I) + (t+1) \times T_H$ and the computation overhead of $Slave_i$ is $(t+1) \times t \times (T_M + T_I) + 2 \times T_H$.

Table 2 shows the comparison of computation overhead between the proposed protocol and Liu's protocol. It can be seen that the proposed protocol obviously reduces the computation overhead of devices, include *Master* and *Slavers*.

5.2 Communication Overhead

The communication overhead is measured using the length of the messages sent by the device in group key establishment process. In the proposed protocol, the communication overheads of *Master* and $Slave_i (i = 1, 2, ..., t)$ are approximately $(2t+2)|n| + |H|$ and $|n|$, respectively. In Liu's protocol, *Master*'s communication overhead is about $2t|n| + |H|$ and $Slaver_i$'s is about $|n|$.

The communication overhead of each protocol is shown in Table 3. The overall communication overhead of the proposed protocol is almost equal to Liu's protocol. And in the two protocols, the overheads of resource-constrained devices, *Slavers*, are identical.

5.3 Storage Overhead

In the proposed protocol, *Master* needs to store the challenge signals of PUF, $\{C_1, C_2\}$, and the response signals of $Slave_j (j = 1, 2, \cdots, m)$, $\{R_{j,1}, R_{j,2}\}$, while *Slavers* don't need to store any parameter. In other words, the storage overhead of *Master* is about $2|n| + 2m|n|$ and $Slaver_j$'s overhead is 0. In Liu's protocol, *Master*'s storage overhead is $2m|n|$ and $Slaver_j$'s is $2|n|$.

Compared with Liu's protocol, the proposed protocol significantly reduces the total storage overhead of the network and the overheads of resource-constrained devices are lower.

6 Conclusions

In this paper, we analyze the security of the existing group key transfer protocols when applied to the bluetooth network and put forward several feasible attack methods respectively. As a remedy, we have proposed a PUF-based group key transfer protocol for bluetooth. The security of the proposed protocol is based on the unclonability and unpredictability of PUF. Compared with related protocols, this protocol significantly reduces the resource overhead of the device and its security is higher.

References

[1] M. Aman, K. C. Chua, and B. Sikdar, "Mutual authentication in iot systems using physical unclonable functions," *IEEE Internet of Things Journal*, vol. 4, no. 5, pp. 1327–1340, 2017.

[2] Bluetooth SIG, *Specification of the Bluetooth System: Core Package Version 4.0*, Technical Report Bluetooth Core v4.0, Dec. 2009.

[3] E. Bresson, O. Chevassut, and D. Pointcheval, "Provably secure authenticated group diffie-hellman key exchange," *Acm Transactions on Information and System Security Journal*, vol. 10, no. 3, pp. 255–264, 2007.

[4] U. Chatterjee, R. S. Chakraborty, and D. Mukhopadhyay, "A puf-based secure communication protocol for IOT," *Acm Transactions on Embedded Computing Systems*, vol. 16, no. 67, pp. 1–25, 2016.

[5] J. C. Cheng and C. S. Laih, "Conference key agreement protocol with non-interactive fault-tolerance over broadcast network," *International Journal of Information Security*, vol. 8, no. 1, pp. 37–48, 2009.

[6] J. J. V. Diaz, A. B. R. Gonzalez, and M. R. Wilby, "Bluetooth traffic monitoring systems for travel time estimation on freeways," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 1, pp. 123–132, 2016.

[7] M. Eltoweissy, M. H. Heydari, L. Morales, and I. H. Sudborough, "Combinatorial optimization of group key management," *Journal of Network and Systems Management*, vol. 12, no. 1, pp. 33–50, 2004.

Table 1: Security comparison

Scheme	Resisting Outsider Attack	Resisting Insider Attack	Resisting Replication Attack	Forward Security	Backward Security
Harn's protocol [9]	√	×	×	√	√
Liu's protocol [17]	√	√	×	√	√
Hsu's protocol [11]	√	×	×	√	√
Piao's protocol [24]	√	√	×	×	×
Our protocol	√	√	√	√	√

Table 2: Comparison of computation overhead

Protocol	***Master***	***Slaver***
Our protocol	$\frac{1}{2} \times t \times (t+1) \times T_M + T_H$	$(2t-1) \times T_M + T_H$
Liu's protocol	$t \times (t+1) \times t \times (T_M + T_I) + (t+1) \times T_H$	$(t+1) \times t \times (T_M + T_I) + 2 \times T_H$

Table 3: Comparison of communication overhead

Protocol	***Master***	***Slaver***
Our protocol	$(2t+2)\|n\| + \|H\|$	$\|n\|$
Liu's protocol	$2t\|n\| + \|H\|$	$\|n\|$

Table 4: Comparison of communication overhead

Protocol	***Master***	***Slaver***	Totally
Our protocol	$2\|n\|+2m\|n\|$	0	$2\|n\|+2m\|n\|$
Liu's protocol	$2m\|n\|$	$2\|n\|$	$4m\|n\|$

[8] Y. B. Guo, Z. N. Zhang, and K. W. Yang, "Authenticated key exchange protocol based on physical unclonable function system in wireless sensor networks," *International Journal of Advancements in Computing Technology*, vol. 4, no. 23, pp. 300–308, 2012.

[9] L. Harn and C. Lin, "Authenticated group key transfer protocol based on secret sharing," *IEEE Transactions on Computers*, vol. 59, no. 6, pp. 842–846, 2010.

[10] H. Harney, C. Muckenhirn, and T. Rivers, *Group Key Management Protocol (GKMP) Architecture*, RFC 2094, July 1997.

[11] C. F. Hsu, L. Harn, T. He, and M. Zhang, "Efficient group key transfer protocol for WSNs," *IEEE Sensors Journal*, vol. 16, no. 11, pp. 4515–4520, 2016.

[12] C. F. Hsu, L. Harn, Y. Mu, M. Zhang, and X. Zhu, "Computation-efficient key establishment in wireless group communications," *Wireless Networks*, vol. 23, no. 1, pp. 1–9, 2016.

[13] K. H. Huang, Y. F. Chung, H. H. Lee, F. Lai, and T. S. Chen, "A conference key agreement protocol with fault-tolerant capability," *Computer Standards and Interfaces*, vol. 31, no. 2, pp. 401–405, 2009.

[14] M. S. Hwang, C. C. Lee, J. Z. Lee, and C. C. Yang, "A secure protocol for bluetooth piconets using elliptic curve cryptography", *Telecommunication Systems*, vol. 29, no. 3, pp. 165-180, 2005.

[15] J. Katz and M. Yung, "Scalable protocols for authenticated group key exchange," *Journal of Cryptology*, vol. 20, no. 1, pp. 85–113, 2003.

[16] Y. S. Lee, H. J. Lee, and E. Alasaarela, "Mutual authentication in wireless body sensor networks (WBSN) based on physical unclonable function (PUF)," in *International Wireless Communications and Mobile Computing Conference (IWCMC'13)*, pp. 1314–1318, July 2013.

[17] Y. Liu, C. Cheng, J. Cao, and T. Jiang, "An improved authenticated group key transfer protocol based on secret sharing," *IEEE Transactions on Computers*, vol. 62, no. 11, pp. 2335–2336, 2013.

[18] J. W. Lo, S. C. Lin, and M. S. Hwang, "A parallel password-authenticated key exchange protocol for wireless environments," *Information Technology and Control*, vol. 39, no. 2, pp. 146–151, 2010.

[19] C. Marchand, L. Bossuet, U. Mureddu, N. Bochard, and A. Cherkaoui, "Implementation and characterization of a physical unclonable function for IOT: A case study with the TERO-PUF," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 37, no. 1, pp. 97–109, 2018.

[20] G. Mokhtari, Q. Zhang, G. Nourbakhsh, S. Ball, and M. Karunanithi, "Bluesound: A new resident identification sensor Xusing ultrasound array and ble technology for smart home platform," *IEEE Sensors Journal*, vol. 17, no. 5, pp. 1503–1512, 2017.

[21] J. Nam, M. Kim, J. Paik, W. Jeon, and B. Lee, "Cryptanalysis of a group key transfer protocol based on secret sharing," in *Future Generation Information Technology - Third International Conference*, pp. 309–315, Dec. 2011.

[22] R. Pappu, B. Recht, J. Taylor, and N. Gershenfeld, "Physical one-way functions," *Science*, vol. 297, no. 5589, pp. 2026–2030, 2002.

[23] A. Perrig, D. Song, and J. D. Tygar, "Elk, a new protocol for efficient large-group key distribution,"

in *Proceedings of IEEE Symposium on Security and Privacy*, pp. 247–262, May 2001.

[24] Y. Piao, J. U. Kim, U. Tariq, and M. Hong, "Polynomial-based key management for secure intra-group and inter-group communication," *Computers and Mathematics with Applications*, vol. 65, no. 9, pp. 1300–1309, 2013.

[25] S. Skorobogatov, "Flash memory 'bumping' attacks," in *International Conference on Cryptographic Hardware and Embedded Systems (CHES'10)*, pp. 158–172, Aug. 2010.

[26] M. Singh and N. Jain, "Performance and evaluation of smartphone based wireless blood pressure monitoring system using bluetooth," *IEEE Sensors Journal*, vol. 16, no. 23, pp. 8322–8328, 2016.

[27] G. E. Suh and S. Devadas, "Physical unclonable functions for device authentication and secret key generation," in *Design Automation Conference (DAC'07)*, pp. 9–14, June 2007.

[28] C. C. Yang, T. Y. Chang, and M. S. Hwang, "A new anonymous conference key distribution system based on the elliptic curve discrete logarithm problem," *Computer Standards and Interfaces*, vol. 25, no. 2, pp. 141–145, 2003.

Biography

Sensen Li is currently an Assistant in Zhengzhou Information Science and Technology Institute, China. His research interests include wireless sensor networks, information security and bluetooth.

Bin Yu is currently a Professor with Zhengzhou Information Science and Technology Institute, China. His research interests include information security, wireless sensor networks, embedded systems and visual cryptography.

Yicai Huang is currently a Lecturer with Zhengzhou Information Science and Technology Institute, China. His research interests include wireless sensor networks, information security and bluetooth.

SPA Resistant Scalar Multiplication Using Pell Lucas Type Chain

Shuang-Gen Liu and Hui Zhao
(Corresponding author: Shuang-Gen Liu)

Department of Information and Communication Engineering, Xi'an University of Posts and Telecommunications
Xi'an 710121, China
(Email: liusgxupt@163.com)

(Received Dec. 13, 2017; Revised and Accepted Mar. 26, 2018; First Online Mar. 4, 2019)

Abstract

A new fast and secure elliptic curve scalar multiplication algorithm is presented. The method is to utilize the front and back ratio coefficient of Pell Lucas sequences. The outcome is a new addition chain: Pell Lucas Type Chain(PLTC), and combines the mixed coordinates which shortens the previous ones. The energy curve of PLTC algorithm is unified, and can resist simple power attacks. Based on theoretical assumption and simulation experiments, it can be obtained that the new scalar multiplication by the PLTC method is 22.7 % faster than the golden ratio addition chain.

Keywords: Golden Ratio Addition Chain; Pell Lucas Type Chain; Scalar Multiplication; Simple Power Attacks

1 Introduction

Elliptic curve cryptography was proposed independently by Koblitz [15] and Miller [18] in 1985. Compared with RSA public key cryptography and EIGamal [14] public cryptography, elliptic curve cryptography provides higher security strength. For example, a 160-bit elliptic curve public key could provide comparable security to a 1024-bit RSA public key.

Hence, the elliptic curve cryptography suits the environment when the storage is limited [5,23]. The dominant operation in elliptic curve cryptography cryptographic schemes is the scalar multiplication, which is represented as kP=P+P+...+P, where P is a point given by the elliptic curve E and k is an integer, which plays the role of secret key [3]. Scalar multiplication of any one point on elliptic curves seems to be a simple addition, and yet, in the underlying field, it involves so many of multiplications. It is of great significance to find out a new method to make the chain shorter. The elliptic curve has different computational efficiency under different coordinate system. Select a suitable coordinate is critical for the scalar multiplication optimization.

There are three main operations in the underlying of the scalar multiplication: inverse, multiplication and square. The inverse is most time-consuming. Except for affine coordinate [21], coordinates which don't need inverse operation. To increase the efficiency of operation, the project coordinate [21] is often used. At the same time the Jacobian coordinate and the five element Jacobian coordinates are also used, which both proposed by Chudnovsky. It is difficult to improve the efficiency of operation by using only one coordinate [7]. But Cohen proposed that converting between the coordinates is easy, that is the characteristic of mixed coordinates [10, 19].

The core of the security chip is Cryptography algorithm. In the processing of information, there is a risk of information leakage, such as power, electromagnetic radiation, and running time. Attacker can collect and analyze the leak information then launch offensive attacks. In 1996, Kocher proposed the Side Channel Attacks (SCA) [20], it is divided into two categories: Simple Power Analysis(SPA) [6] and Differential Power Analysis (DPA) [4]. The simple power analysis is used to analyze the energy consumed by a single password operation. Because different operations have different energy consumption. For different energy consumption, an attacker can infer the order [10]. There are usually two ways to resist SPA attack. The first way is just using one kind algorithm, such as Golden Ratio Addition Chain(GRAC) [12] and the Montgomery Power Ladder [13]. The other way is to use the regular rules in algorithm, such as Double-and-add algorithm [17].

The paper presents a new $2P+Q$ algorithm using the best Mixed coordinate, which based on properties of the pell-lucas sequence and get the PLTC. The issue is mainly addressed in five parts. Part 1 gives an introduction to elliptic curve cryptography and the derivation of the pell-lucas sequence from the Lucas sequence. Part 2 introduced the new addition chain—-Pell Lucas Type Chain(PLTC). The application of PLTC in elliptic curve cryptosystems is introduced in Part 3. Part 4 makes a comparison between the PLTC and the previous algorithms under the same coordinate, at the same time,

analyze the resist of SPA attack.

2 Background

This part explain Elliptic Curve Cryptography and Pell Lucas sequence.

2.1 Elliptic Curve Cryptography

The elliptic curve E over the field K is defined by Weierstrass equation.

$$E : y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6. \quad (1)$$

Where $a_1, a_2, a_3, a_4, a_6 \in K$ and $\Delta \neq 0$, the Δ is discriminant of E. When the characteristic of the field K is greater than 3, the equation can be simplified to:

$$E : y^2 = x^3 + ax + b. \quad (2)$$

Where $a, b \in K$ and, $\Delta = 4a^3 + 27b^2 \neq 0$.
There are two infinite points on this curve:

$$\begin{aligned} P &= (x_1, y_1), \\ Q &= (x_2, y_2), \\ P+Q &= (x_3, y_3). \end{aligned}$$

• Point Addition ($P \neq Q$)

$$\begin{aligned} x_3 &= \lambda^2 - x_1 - x_2, \\ y_3 &= \lambda(x_1 - x_3) - y_1, \\ \lambda &= \frac{y_2 - y_1}{x_2 - x_1}. \end{aligned} \quad (3)$$

• Point Doubling ($P = Q$)

$$\begin{aligned} x_3 &= \lambda^2 - 2x_1, \\ y_3 &= \lambda(x_1 - x_3) - y_1, \\ \lambda &= \frac{3x_1^2 + a}{2y_1} \end{aligned} \quad (4)$$

The computation of $2P+Q$ often in two methods, compute $2P$ using the double point Equation (4), add Q using Equation (3).

Equations (3) and (4) consist of multiplication, inverse and square of large integer. The three methods are represented as *M*, *I*, *S*. Comparing with these three operations, the calculation of integer addition and large integer multiplication can be ignored. S/M is equal to 0.8. The I/M ratio is generally about 10 [10, 11]. The data show that the inverse operation is the most time-consuming. Under the affine coordinates, each cost time of $2P+Q$ is $1I+9M+2S$,*Ciet*. But in [10], the realization of point addition and double point operation in other coordinates does not need to compute the inverse operation. In this paper, discussion of the complexity of algorithm is based on the Mixed coordinate, the literature [1, 9, 10] state the operation method under the Mixed coordinate [22].

2.2 Pell Lucas Sequence

The Lucas sequence is an important result of the study by Lucas in the 19th century, now it has become an important integer sequence in the Theory of Numbers. There are some inseparable links between the Lucas sequence and the Fibonacci sequence.

Definition 1. *The Fibonacci sequence is defined as* $F_n = F_{n-1} + F_{n-2} (n \geq 2)$, *and* $F_0 = 0$, $F_1 = 1$.

Definition 2. *The Lucas sequence [2] is defined as* $L_{n+1} = L_n + L_{n-1} (n = 1, 2, \cdots)$ *and* $L_0 = 2$, $L_1 = 1$. *The general equation is*

$$\begin{aligned} L_n &= \alpha^n + \alpha^n (n \geq 0) \\ \alpha &= \frac{\sqrt{5}+1}{2} \\ \beta &= \frac{1-\sqrt{5}}{2} \end{aligned}$$

It can be seen that the Fibonacci sequence and Lucas sequence are different in beginning, but the relationship between the number is same.While the Lucas sequence is consists of two linear, so there is another way to define the Lucas sequence. Take the two integers P,Q to satisfy the equation: $\Delta = P^2 - 4Q > 0$.

So we can get the equation: $x^2 - Px + Q = 0$, the roots of equation are a, b, based on this, the Lucas sequence can also be defined as

$$\begin{aligned} U_n(P, Q) &= (a^n - b^n)/(a - b), \\ V_n(P, Q) &= (a^n + b^n). \end{aligned} \quad (5)$$

Where $n \geq 0$, so we can get

$$\begin{aligned} U_0(P, Q) &= 0 \\ U_1(P, Q) &= 1 \\ V_0(P, Q) &= 2 \\ V_1(P, Q) &= P. \end{aligned}$$

If take $(P, Q) = (1, -1)$ into U_n suquence, we can get the Fibonacci sequence:

$$\{0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377 \ldots\}.$$

If take $(P, Q) = (1, -1)$ into V_n sequence,we can get the Lucas sequence:

$$\{2, 1, 3, 4, 7, 11, 18, 29, 47, 76, 123, 199, 322, 521, 843 \ldots\}.$$

When $(P, Q) = (2, -1)$, the equation $V_n(2, -1)$ is Pell-Lucas sequence, which can be represented as follows:

$$\{2, 2, 6, 14, 34, 82, 198, 418, 1154, 2786, 6726, \ldots\}.$$

At the same time, $U_n(2, -1)$ is Pell sequence:

$$\{0, 1, 2, 5, 12, 29, 70, 169, 408, 985, 2378, 5741, \ldots\}.$$

The general term of Pell-Lucas and Pell sequence is

(Pell-Lucas)

$$V_n = (a^n + b^n)(a = 1 - \sqrt{2}, b = 1 + \sqrt{2}),$$
$$\lim_{n\to\infty} \frac{V_n}{V_{n+1}} = \lim_{n\to\infty} \frac{a^n - b^n}{a^{n+1} - b^{n+1}} \approx 0.414. \quad (6)$$

(Pell)

$$U_n = \frac{a^n - b^n}{a - b}(a = 1 - \sqrt{2}, b = 1 + \sqrt{2}),$$
$$\lim_{n\to\infty} \frac{U_n}{U_{n+1}} = \lim_{n\to\infty} \frac{(a^n - b^n)(a - b)}{(a^{n+1} - b^{n+1})(a - b)} \approx 0.414. \quad (7)$$

It can be seen that both of Pell-Lucas or Pell sequence satisfy the following properties:

$$L_{i+1} = L_{i-1} - 2L_i (i = 1, \ldots, n), \quad (8)$$
$$L_i = L_{i+1} \times 0.414 (i = 1, \ldots, n). \quad (9)$$

3 Pell Lucas Type Chain

Equations (8) and (9) can account for the Pell-Lucas and Pell sequence, and both of the sequence satisfy Equations (8) and (9). But it's easy to see that if one sequence is corresponds to formula $L_{i+1} = L_{i-1} - 2L_i (i = 1, \ldots, n)$, it is only going to fit the formula $L_i = L_{i+1} \times 0.414 (i = 1, \ldots, n)$ at the beginning. As the extended of sequence, the ratios of front and back are deviates from 0.414. The GRAC using GAP to determine the sequence of the gold addition chain. But select the number of GAP is a new major research problem. So we define a new sequence: Pell Lucas Type Chain(PLTC).

Definition 3. *The Pell Lucas Type Chain is a sequence satisfy the formula* $L_{i+2} = L_i - 2L_{i+1} (i = 1, 2, \ldots, n)$ *and* $L_{n+1} > L_n > 0$.

The PLTC can be applied to the scalar multiplication of Elliptic curve and can greatly shorten the length of the double-and-add chain.

PLTC is not a Standard Pell-Lucas sequence. PLTC is just a chain roughly satisfies the properties of the Pell-Lucas sequence. Applying this to the elliptic curve can get Algorithm 1.

For the facilitation of the calculation,three sets $e\{\}$, $s\{\}$ and $y\{\}$ must be used in Algorithm 1. The calculation begins with the integer number k. The first step is to obtain an integer number close to $k \times 0.414$. Then we can apply $u_{i+1} = u_{i-1} - u_i \times 2 (u_i > 1, i = 1, \ldots, l)$,base on this, there will be two situations.

A: $0 < u_{i+1} < u_i \to e_i = 1$,

B: $u_{i+1} \geq u_i or u_{i+1} \leq 0 \to e_i = 0$.

$u'_{i+1} = u_{i+1} \to u'_{i+1} = s_i \to u_{i+1} = \frac{1}{2} u_i$,
if $\text{Mod}(u_{i+1}, 2) = 1 \to \{e_{i+1} = 1, y = 1\}$;
if $\text{Mod}(u_{i+1}, 2) = 0 \to \{e_{i+1} = 1, y = 0\}$.

At the last step, the number is too small, so we have two cases for the end of the reference. One is end of e=1, another one is end of e=0. Each time we will get an s or a y. We call these two end of methods are the S type end mode and Y type end mode. Each situation is shown in case Example 1 and Example 2.

Algorithm 1 Pell Lucas-Type Addition Chain

Input: A positive integer **k**
Output: $e = \{e_1, e_2, \ldots, e_i\}$
$y = \{y_1, y_2, \ldots, y_j\}, s = \{s_1, s_2, \ldots, s_j\}$
$u_0 \leftarrow \mathbf{k}$
$e\{\}$
$u_1 \leftarrow u_0 \times 0.414$
$u_2 \leftarrow u_0 - 2u_1$
$e \leftarrow e \cup \{1\}$
$s\{\}$
$y\{\}$
while $u_i > 1$ **do**
  $u_{i+1} \leftarrow u_i \times 0.414$
  $e \leftarrow e \cup \{1\}$
  $u'_{i+2} \leftarrow u_i - 2 \times u_{i-1}$
  if $0 < u'_{i+2} < u_{i+1}$ **then**
    $e \leftarrow e \cup \{1\}$
    $u_{i+2} \leftarrow u'_{i+2}$
  end if
  if $u'_{i+2} \geq u_{i+1} or u'_{i+2} \leq 0$ **then**
    $e \leftarrow e \cup \{0\}$
    $s \leftarrow s \cup \{u'_{i+2}\}$
    $u_{i+2} \leftarrow \frac{u_{i+1}}{2}$
    if $u_{i+1} mod 2 = 1$ **then**
      $e \leftarrow e \cup \{1\}$
      $y \leftarrow y \cup \{1\}$
    end if
    if $u_{i+1} mod 2 = 0$ **then**
      $e \leftarrow e \cup \{1\}$
      $y \leftarrow y \cup \{0\}$
    end if
  end if
end while

From Example 1, we can get the three sets. But if it use this data to restore the k, three sets must be reversed and get the sets like :

$$e = \{1, 0, 1, 1, 0, 1, 1, 1, 1, 0, 1, 1, 1, 1\},$$
$$s = \{-1, -8, 2614\},$$
$$y = \{0, 1, 1\}.$$

Using the same method of Example 1, from Example 2, we can get the three sets:

$$e = \{0, 1, 1, 0, 1, 1\},$$
$$s = \{0, 11\},$$
$$y = \{0\}.$$

Example 1.k=131456, three sets:$e\{\}, y\{\}, s\{\}$

(**Y type end mode**)

$u_0 = \mathbf{k} = 131456$

$e = 1 \quad u_1 = u_0 \times 0.414 = 54423$

$e = 1 \quad u_2 = u_0 - 2u_1 = 22610$

$e = 1 \quad u_3 = u_1 - 2u_2 = 9203$

$e = 1 \quad u_4 = u_2 - 2u_3 = 4204$

$e = 0 \quad u_5 = u_3 - 2u_4 = 795$

$u'_6 = u_4 - 2u_5 = 2614, since 2614 > u_5,$

set $e = 0, s = s \cup \{u'_6 = 2614\}$

$since\ u_6 = \frac{u_5}{2}, and\ u_6 mod 2 = 1,$

set $e = 1, y = y \cup \{1\}$

$e = 1 \quad u_6 = \frac{u_5}{2} = 397 \cdots 1$

$e = 1 \quad u_7 = u_6 \times 0.414 = 164$

$e = 1 \quad u_8 = u_6 - 2u_7 = 69$

$e = 1 \quad u_9 = u_7 - 2u_8 = 26$

$e = 0 \quad u_{10} = u_8 - 2u_9 = 17$

$u'_{11} = u_9 - 2u_{10} = -8, since -8 < 0,$

set $e = 0, s = s \cup \{u'_{11} = -8\}$

$since\ u_{11} = \frac{u_{10}}{2}, and\ u_{11} mod 2 = 1,$

set $e = 1, y = y \cup \{1\}$

$e = 1 \quad u_{11} = \frac{u_{10}}{2} = 8 \cdots 1$

$e = 1 \quad u_{12} = u_{11} \times 0.414 = 3$

$e = 1 \quad u_{13} = u_{11} - 2u_{12} = 2$

$u'_{14} = u_{12} - 2u_{13} = -1, since -1 < 0$

set $e = 0, s = s \cup \{u'_{14} = -1\}$

$since\ u_{14} = \frac{u_{13}}{2}, and\ u_{14} mod 2 = 0,$

set $e = 1, y = y \cup \{0\}$

$e = 1 \quad u_{14} = \frac{u_{13}}{2} = 1 \cdots 0$

$since\ u_{13} - 2u_{14} = 0, and\ 0 < 1$

END

Example 2.k=175, three sets:$e\{\}, y\{\}, s\{\}$

(**S type end mode**)

$u_0 = \mathbf{k} = 175$

$e = 1 \quad u_1 = u_0 \times 0.414 = 72$

$e = 1 \quad u_2 = u_0 - 2u_1 = 31$

$e = 0 \quad u_3 = u_1 - 2u_2 = 10$

$u'_4 = u_2 - 2u_3 = 11, since 11 > u_3,$

set $e = 0, s = s \cup \{u'_4 = 11\}$

$since\ u_4 = \frac{u_3}{2}, and\ u_3 mod 2 = 0,$

set $e = 1, y = y \cup \{0\}$

$e = 1 \quad u_4 = \frac{u_3}{2} = 5 \cdots 0$

$e = 1 \quad u_5 = u_4 \times 0.414 = 2$

$e = 0 \quad u_6 = u_4 - 2u_5 = 1$

$u'_7 = u_5 - 2u_6 = 0 < 1$

set $e = 0, s = s \cup \{u'_7 = 0\}$

END

4 Application of PLTC to Elliptic Curve Cryptosystem

In Algorithm 2, there are two assignment required for each operation, T and T_0 are intermediate values in the algorithm, the cost time of assignment operation can be ignored. the last value is not remembered when the assignment end at each time, so it has no effect on memory space.

Algorithm 2 PLTC using to elliptic curve

Input: $e = \{e_1, e_2, \ldots, e_n\}, y = \{y_1, y_2, \ldots, y_i\}, s = \{s_1, s_2, \ldots, s_j\}$
Output:kP
Main loop
$i = 1$
$j = 1$
$n = 1$
if $e_n = 0$ **then**
  $T \leftarrow P$
  $P \leftarrow 2P + s_i P$
  $T_0 \leftarrow T$
  $i++$
  $n++$
end if
if $e_n = 1$ *and* $e_{n+1} = 0$ **then**
  $T \leftarrow P$
  $P \leftarrow 2P + y_j P$
  $T_0 \leftarrow T$
  $j++$
  $n++$
end if
if $e_n = 1$ *and* $e_{n+1} \neq 0$ **then**
  $T \leftarrow P$
  $P \leftarrow 2P + T_0 P$
  $T_0 \leftarrow T$
  $n++$
end if
$Q \leftarrow P$

Hence, the output is kP=Q. In Algorithm 2 operation, no mater the bit is 1 or 0, each scalar multiplication has one addition and one doubling. The two sets s and y does not affect the rate of calculation. Because all of their operations are contained in the operation of set e. Set s and set y are the fixed sequences of PLTC. These can be demonstrated in Example 3 and Example 4.

5 Discussion

5.1 Scalar Multiplication Analysis

Randomly selected 10000 of the large integers from 160 bits. Count the same chain length, According to the statistics, up to the most were 116, 117 and 118 bits. Choose the four times statistical results can obtain the Table 1. Count the length of chains from 111 to 120 and show in graph like Figure 1.

We can see from the Table 1 and Figure 1, 117 bit is always the most. The distribution of chain length is in accordance with the gaussian distribution. So the length of PLTC-160 can be seen as 117.

Example 3. e={1,0,1,1,0,1,1,1,1,0,1,1,1,1}
s={-1,-8,2614}
y={0,1,1}

$e_1 = 1, T = P, P = 2P + y_1P, T_0 = T$
$(P = 2P)$
$e_2 = 0, T = P, P = 2P + s_1P = 3P, T_0 = T$
$(P = 3P)$
$e_3 = 1, T = P, P = 2P + T_0P = 8P, T_0 = T$
$(P = 8P)$
$e_4 = 1, T = P, P = 2P + y_2P = 17P, T_0 = T$
$(P = 17P)$
$e_5 = 0, T = P, P = 2P + s_2P = 26P, T_0 = T$
$(P = 26P)$
$e_6 = 1, T = P, P = 2P + T_0P = 69P, T_0 = T$
$(P = 69P)$
$e_7 = 1, T = P, P = 2P + T_0P = 164P, T_0 = T$
$(P = 164P)$
$e_8 = 1, T = P, P = 2P + T_0P = 397P, T_0 = T$
$(P = 397P)$
$e_9 = 1, T = P, P = 2P + y_3P = 795P, T_0 = T$
$(P = 795P)$
$e_{10} = 0, T = P, P = 2P + s_3P = 4202P, T_0 = T$
$(P = 4204P)$
$e_{11} = 1, T = P, P = 2P + T_0P = 9203P, T_0 = T$
$(P = 9203P)$
$e_{12} = 1, T = P, P = 2P + T_0P = 22610P, T_0 = T$
$(P = 22610P)$
$e_{13} = 1, T = P, P = 2P + T_0P = 54423P, T_0 = T$
$(P = 54423P)$
$e_{14} = 1, T = P, P = 2P + T_0P = 131456P, T_0 = T$
$(P = 131456P)$
Q=131456P

Example 4. e={0,1,1,0,1,1}
s={0,11}
y={0}

$e_1 = 0, T = P, P = 2P + s_1P, T_0 = T$
$(P = 2P)$
$e_2 = 1, T = P, P = 2P + T_0P = 5P, T_0 = T$
$(P = 5P)$
$e_3 = 1, T = P, P = 2P + y_1P = 10P, T_0 = T$
$(P = 10P)$
$e_4 = 0, T = P, P = 2P + s_2P = 31P, T_0 = T$
$(P = 31P)$
$e_5 = 1, T = P, P = 2P + T_0P = 26P, T_0 = T$
$(P = 72P)$
$e_6 = 1, T = P, P = 2P + T_0P = 175P, T_0 = T$
$(P = 175P)$
Q=175P

We have five different kinds of coordinate systems (A, P, J, J_c, J_m) [10] that we often used. Here we compare the different cost of doubling and addition between different coordinate system. Both computation time of the operation [17] shown in Table 2 and Table 3.

Table 1: 116, 117, 118bit of PTLC

	116bit	117bit	118bit
The first time	3198	3483	1586
The second time	1844	5391	1018
The third time	758	3963	3283
The forth time	1228	4049	3037

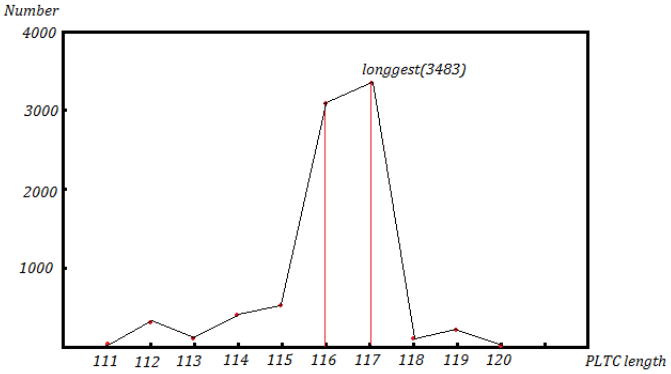


Figure 1: The number of 160-length change

Analyzing of the Table 2 and Table 3, we can obtained that the point addition operation under the J coordinates is the most time-saving operation, at the same time, the point doubling operation under the J^c coordinate is the most time-saving way. The resulting mixed coordinates are shown in Table 4.

In Algorithm 2, the calculation has one addition and one doubling each time. The length of addition is $\frac{1}{2}l$, and the length of doubling is $\frac{1}{2}l(l = 117)$. So we get the following formula.

$$\sharp[m] = l(7[M] + 7[S]).$$

The cost time of PLTC can be calculated as $1474[m]$. To effectively illustrate the advantages of PLTC algorithm. we choose to compare the number with other algorithms in the same coordinate and get Table 5.

From Table 5, we can see that under the same coordinate, PLTC is 22.7% faster than the GRAC, 7.9%, 17.2% and 29.9% faster than the 4-NAF, NAF and Double-and-add. At the same time, the reduction of chain length is considerable, and the results are shown in Table 6.

From Table 6, we can see the length of PLTC is shorter than other kind of algorithms. Even under the same length number with DFAC-160, PLTC-160 is 26.9%shorter than DFAC-160. Compared with other algorithms, PLTC is more suitable for the environments such as security chips and smart cards, which are more demanding about memory space.

Table 2: Doubling cost on different coordinates

doubling	
operation	**costs**
$2A{=}J$	$2[M]+4[S]$
$2A=J^m$	$3[M]+4[S]$
$2J^m=J$	$3[M]+4[S]$
$2A=J^c$	$3[M]+5[S]$
$2J^m$	$4[M]+4[S]$
$2J^m=J^c$	$4[M]+5[S]$
$2J$	$4[M]+6[S]$
$2J^c$	$5[M]+6[S]$
$2P$	$7[M]+5[S]$

Table 3: Addition cost on different coordinates

addition	
operation	**costs**
$P+P$	$12[M]+2[S]$
J^m+J^m	$13[M]+6[S]$
$J+A$	$8[M]+3[S]$
$J^m+A=J^m$	$9[M]+5[S]$
$J^m+A=J$	$8[M]+3[S]$
$J^c+J=J$	$11[M]+3[S]$
$J^c+J^c=J^m$	$11[M]+4[S]$
$J^c+J^c=J$	$10[M]+2[S]$
J^c+J^c	$11[M]+3[S]$
$J^c+A=J^m$	$8[M]+4[S]$
$J^c+A=J^c$	$8[M]+3[S]$
$J+A=J^m$	$9[M]+5[S]$
$A+A=J^m$	$5[M]+4[S]$
$A+A=J^c$	$5[M]+3[S]$
$J+J$	$12[M]+4[S]$
$J^c+J=J^m$	$12[M]+5[S]$
$J^m+J^c=J^m$	$12[M]+5[S]$

Table 4: Mixed coordinate

	Addition	**Doubling**
Operation	$A+A=J^c$	$2A{=}J$
Cost	$5[M]{+}3[S]$	$2[M]{+}4[S]$

Table 5: Mixed coordinate

Algorithm	Coordinate	$\sharp[m]$
Double-and-Add [17]	Mixed	2104
NAF [16, 17]	Mixed	1780
4-NAF [17]	Mixed	1600
GRAC-258 [12]	Mixed	1907
PLTC-117	Mixed	1474

Table 6: The chain length for algorithms

Algorithm	Chain Length
Fibonacci-add-add	358
Signed Fib-add-add	322
Window Fib-add-add	292
EAC-320	320
GRAC-258	258
DFAC-160	160
PLTC-160	117

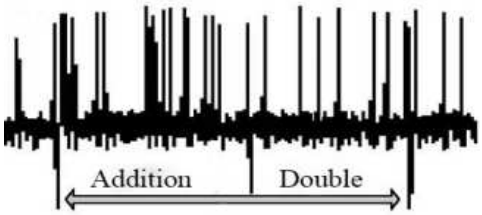


Figure 2: The power waveform of e=0

5.2 Resist SPA Analysis

The key obtained by PLTC algorithm is composed of "0" and "1", The power consumption waveforms obtained in both cases are shown in Figure 2 and Figure 3. Because the key is longer, randomly select 8 bits (1001 0001) used for PLTC coding. The power consumption waveform of a scalar multiplication in Figure 4, which collected from the power consumption analysis platform.

We can see from the three figures, that each bit has same waveform, no matter it's "0" or "1", both contains one addition and one doubling, the waveform of power is same when attacker see from outside.integrated into Figure 4, it is very hard to distinguish the energy curve, can't know the exactly number of the channel, even select a part of information. so it can resist against SPA.

6 Conclusion

This is the first study to combine Pell Lucas Type sequence with elliptic curve cryptography. With the advantages of the pell-Lucas sequence, we can improve the

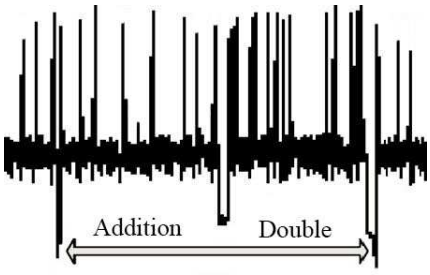


Figure 3: The power waveform of e=1

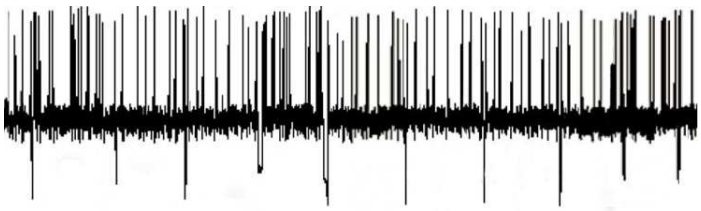


Figure 4: The power waveform of (1001 0001)

ratio between the numbers of the chain and the efficiency significantly.

For further study, we need to address the problem, although S has very little impact on the calculation and can be ignored. It accounts for about 25 %of the total chain. Therefore it will increase the burden of coding, decoding and transmission and add operations for the analysis of the password. the numbers in TABLE S can be bigger when the main chain gets longer. If we could reduce the storage space of, PLTC could be applied to elliptic curve cryptosystems more efficiently where memory is involved, such as smart card.

Acknowledgments

The support of NSFC (National Natural Science Foundation of China, No. 61272525), Jiangxi Natural Science Foundation (No. 2009GQN0094), natural science foundation research project by Shaanxi province (No. 2017JQ6010).

References

[1] D. Adachi and T. Hirata, "Combination of mixed coordinates strategy and direct computations for efficient scalar multiplications," in *IEEE Pacific Rim Conference on Communications, Computers and signal Processing*, pp. 117–120, 2005.

[2] E. Al-Daoud, R. Mahmod, M. Rushdan, and A. Kilicman, "A new addition formula for elliptic curves over GF(2n)," *IEEE Transactions on Computers*, vol. 51, no. 8, pp. 972–975, 2002.

[3] L. M. Batten, *Public Key Cryptography: Applications and Attacks*, Wiley, 2013.

[4] T. Caddy, *Differential Power Analysis*, Springer, Boston, MA, 2005.

[5] C. C. Chang, Y. Liu 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.

[6] T. Chen, F. Yu, K. Wu, H. Li, "Simple power analysis on elliptic curve cryptosystems and countermeasures: Practical work," *IEEE*, 2009.

[7] D. V. Chudnosky and G. V. Chudnovsky, *Sequences of Numbers Generated by Addition in Formal Groups and New Primarily and Factorization Tests*, 1986. (`https://core.ac.uk/download/pdf/82012348.pdf`)

[8] M. Ciet, M. Joye, K. Lauter, and P. L. Montgomery, "Trading inversions for multiplications in elliptic curve cryptography," *Designs, Codes and Cryptography*, vol. 39, no. 2, pp. 189–206, 2006.

[9] H. Cohen, "Analysis of the sliding window powering algorithm," *Journal of Cryptology*, vol. 18, no. 1, pp. 63–76, 2005.

[10] H. Cohen, A. Miyaji, T. Ono, "Efficient elliptic curve exponentiation using mixed coordinates," in *International Conference on the Theory and Applications of Cryptology and Information Security: Advances in Cryptology*, pp. 51–65, 1998.

[11] K. Fong, D. Hankerson, J. Lopez, and A. Menezes, "Field inversion and point halving revisited," *IEEE Transactions on Computers*, vol. 53, no. 8, pp. 1047–1059, 2004.

[12] R. R. Goundar, K. I. Shiota, and M. Toyonaga, "Spa resistant scalar multiplication using golden ratio addition chain method," *Iaeng International Journal of Applied Mathematics*, vol. 38, no. 2, pp. 83–88, 2008.

[13] K. Javeed, X. Wang, "Efficient montgomery multiplier for pairing and elliptic curve based cryptography," *International Symposium on Communication Systems*, 2014. DOI: 10.1109/CSNDSP.2014.6923835

[14] J. Kar, "Id-based deniable authentication protocol based on diffie-hellman problem on elliptic curve," *International Journal of Network Security*, vol. 15, no. 5, pp. 357–364, 2013.

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

[16] T. C. Lin, "Algorithms on elliptic curves over fields of characteristic two with non-adjacent forms," *International Journal of Network Security*, vol. 9, no. 2, pp. 117–120, 2009.

[17] N. Meloni, "New point addition formulae for ecc applications," in *International Workshop on the Arithmetic of Finite Fields*, pp. 189–201, 2007.

[18] V. S. Miller, "Uses of elliptic curve in cryptography," *Lecture Notes in Computer Science Springer-Verlag*, vol. 218, pp. 417–426, 1986.

[19] V. Natarajan, M. Lavanya, "Improved elliptic curve arithmetic over gf(p) using different projective coordinate system," *Applied Mathematical Sciences*, vol. 9, no. 45, pp. 2235–2243, 2015.

[20] Reddy and E. Kesavulu, "Elliptic curve cryptosystems and side-channel attacks," *International Journal of Network Security*, vol. 12, no. 3, pp. 151–158, 2011.

[21] J. T. Tate, "The arithmetic of elliptic curves," *Inventiones Mathematicae*, vol. 23, pp. 179–206, 1974.

[22] X. C. Yin and H. X. Hou, "Improved sliding window scalar multiplication algorithm," *Journal of Chinese Computer Systems*, vol. 29, no. 5, pp. 863–866, 2008.

[23] D. Yong, Y. F. Hong, W. T. Wang, Y. Y. Zhou, and X. Y. Zhao, "Speeding scalar multiplication of

elliptic curve over gf(2)," *International Journal of Network Security*, vol. 11, no. 10, pp. 70–77, 2010.

Biography

Shuang-Gen Liu was born in 1979, associate professor. He graduated from Xidian University in 2008 with a major in cryptography, PhD, a member of the Chinese Institute of computer science, and a member of the Chinese code society.

Hui Zhao is a graduate student of Xi'an University of post and telecommunications. She is mainly engaged in the research of elliptic curve cryptosystem.

Recent Trends in Development of DDoS Attacks and Protection Systems Against Them

Vladimir Galyaev, Evgenia Zykova, Dmitry Repin, and Denis Bokov
(Corresponding author: Vladimir Galyaev)

Laboratory of Mathematical Modelling and Information Technologies,
State Institute of Information Technologies and Telecommunications (SIIT&T "Informika")
Brusov per., 21 str.2, Moscow, Russian Federation, 125009
(Email: v.galiaev@informika.ru)
(Received Mar. 30, 2018; Revised and accepted July 7, 2018; First Online May 28, 2019)

Abstract

Distributed denial of service (DDoS) attacks are considered the most common and often the most destructive among all threats to network infrastructure. For the last decade the number of DDoS attacks constantly grows. They become more elaborate and sophisticated making standard security techniques went out of date quickly. In the review we collected, investigated and link together data from academic publications and information security reports provided by top companies in the field. We marked out tendencies in evolution of DDoS attacks, characterized protection systems, and summarized the last achievements and future developments in application of intellectual methods for network security.

Keywords: DDoS Attack; Intellectual Methods; Network Protection; Network Security

1 Introduction

Distributed Denial of Service (DDoS) attack is a type of cyberattacks that aims to exhaust network resources (server capacity, information channel bandwidth) causing resource deny from providing access to legitimate system users. It uses a number of compromised or vulnerable hosts distributed in the Internet to create malicious traffic and send it to a victim.

DDoS attacks are destructive for the network infrastructure and can very quickly disable a server or an entire network. There are following aspects that determine the effectiveness of DDoS attacks [5]:

- There is a large number of interdependencies in the network architecture.
- The resources of network devices are limited.
- Compromised devices may participate in two or more botnets belonging to different attackers and can be used against several target servers or networks.
- Information and resources that can be used to prevent impending attacks are under control of different people.
- Simple and direct routing principles are commonly used in the Internet infrastructure.
- There are inconsistencies in the architecture of different local networks. The speed difference between network devices of the core and the boundary usually occurs.
- Network management is often low-level.
- In general, the useful practice of sharing information and technical resources has its drawbacks.

Any organization may become a target for DDoS attacks, regardless of its size or business scope. Few years ago most common victims for DDoS attacks were top corporations with income highly dependent on network resources: financial institutions, hosting companies and providers of cloud services, major media outlets. Nowadays attacks may also affect small and medium-size enterprises in any sphere of business, from public health institutions and social insurance to e-sport organizations. From surveys involving all around the world organizations from various spheres and with different outcome, it follows that most of the victims suffered financial losses up to $255,000 per hour of an attack [41]. Records are breaking almost every quarter: the maximum duration of one continuous attack is up to 277 hours and the maximum number of attacks per day is 1497 attacks [24]. With increase of average duration, volume and the number of DDoS attacks targeted to a company cost of the damage from DDoS attacks is growing day by day making protection systems much in demand.

Various organizations choose different tactics to protect their resources from DDoS attacks. Some assess their risks as minimal and do not take any additional measures believing that a correctly built network infrastructure can

withstand most threats. However, the development of DDoS as a service increases the chance for a company to become a victim of DDoS attacks. There are no universal means for countering DDoS attacks. In general there are three basic approaches used to provide security measures:

1) Network infrastructure improvement with aim to increase its stability and survivability;

2) Application of specialized hardware and software solutions;

3) Resource protection as a service by top system integrators. Each of the approaches has its advantages and drawbacks and can be used both independently and together (see more details in Section 3).

Taking into account the complexity of DDoS attacks, their multiple vectoring, volume and constant modification, only implementations that are able to adapt to changing conditions will be able to successfully cope with them. Therefore, the mathematical and algorithmic basis for software and hardware solutions are intelligent methods of data analysis. Over the last 10 years a number of research works have been published in this field, suggesting to use as a mathematical basis statistical, signature, heuristic analysis, expert systems, queuing networks, multi-agent systems, genetic and behavioral algorithms. However, the major disadvantages of most solutions are the narrow specialization of the developed methods, as well as determinacy of incoming traffic classification.

This work aims to bring a systematic view of recent DDoS attacks developments, introduce main defence strategies and highlight possible protection mechanisms improvements. We have analysed data for the last year from DDoS attacks reports provided by a number of companies in the field and defined main trends and possible further evolvement.

The rest of the paper is organized as follows. In Section 2 we introduced social and economic aspects of DDoS attacks, attacks classification and statistics of threats, taking in consideration quarter and annual statistics that is available in information security reports of top companies in the field. In Section 3 we paid attention to most commonly used DDoS protection mechanisms for network security and highlighted their advantages and drawbacks. Section 4 is devoted to in-depth traffic analysis via application of intellectual systems and sophisticated statistical algorithms. Finally, we present the concluding remarks in Section 5.

2 Main Tendencies in Development of DDoS Attacks

2.1 Social and Economic Aspects of DDoS Attacks

DDoS attacks evolve from demonstrative actions into a prominent niche of the shadow market. It bears on unfair competition (temporal blockage of the competitor, reputational damage), fraud on electronic stock exchanges and e-sports events, blackmailing and extortion with the threat of an attack on company's resources. In particular, the DDoS for Ransom strategy keeps developing: demonstrating a DDoS attack with a promise of continuing it if the ransom is not paid. The business model explains some DDoS attacks that might look like an attempt to set a new record: attackers demonstrate their capabilities on popular websites to frighten potential victims [23]. Apart from the main goal of the intruder a DDoS attack can also serve as a mask for hacking information resources, penetrating a protected perimeter and stealing confidential data or money.

Often DDoS attacks are used to draw attention and even for revenge. A group of attackers conducted a powerful DDoS attack on the site of famous American journalist Brian Krebs, who writes popular analytical articles about information security and cybercriminals, — KrebsOnSecurity.com. The attack was carried out in September 2016 and by that time became the largest of the officially recorded: volume of malicious traffic reached 620 Gbit/s and the attack duration was almost 2 days. Most of the traffic consisted of generic routing encapsulation data packets, and the attack was carried out using a botnet of hundreds of thousands of IP cameras and video game consoles.

Originally, botnets based on infected computers were used to implement DDoS attacks. Recently new technologies for infection and use of network devices appeared, and botnets on the basis of "smart" things included in Internet of Things (IoT) are being used more frequently. Especially, two record-breaking DDoS attacks on the French hosting company OVH and the American DNS provider Dyn were conducted with help of botnets based on IP cameras, printers and other devices. There is information about revealed vulnerabilities of a number of household appliances, for instance "smart" dishwashers Miele and kitchen stoves AGA [8]. Usually such devices are equipped with inexpensive samples of operating systems with free software and no security support. Most devices have a password and a login "by default", so their hacking is easy. At the same time, the variety of "smart" devices is constantly expanding, and each of them can potentially be used for illegal purposes. The issue is called in press as the phenomenon of "microwave threat".

According to forecasts of top IT companies, the proportion of IoT devices will grow rapidly in the coming years and exceed the number of other devices at 2020 [11]. It is clear that the number of botnets created on the basis of weakly protected mobile devices and IoT elements will grow as well. Major companies have started to pay attention to security issues and are developing security tools for their IoT devices. However market is overflowed with cheap products (most of them are made in China) without any protection systems. Meanwhile the number of scumware samples for "smart" devices is constantly growing creating huge basis for IoT botnets (see Figure

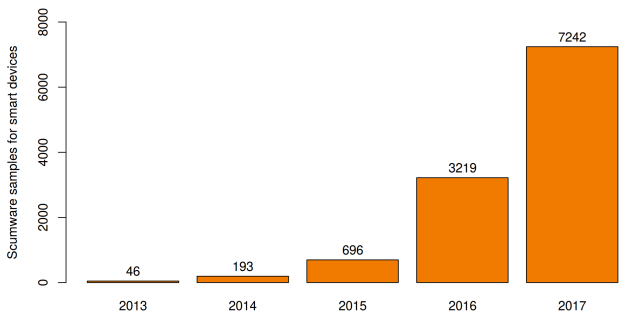


Figure 1: The number of scumware samples for "smart" devices is constantly growing

1, [35]).

Another trend is complication of DDoS attack mechanisms. Intruders are using a few types of DDoS attacks in a single act (multi-vector attack), combining vulnerabilities of different network protocols. In this case, two areas of implementation can be clearly identified. The first one includes ineffective, but massive and easily self-upgrading attacks based on finished products. The second one includes specialized developments that are created individually to attack a particular company, taking into account its specific vulnerabilities and the architecture of its information system. The developments are much more expensive, but as a result they are more effective: in most cases they bypass standard protection systems.

Meanwhile, the cost of conducting a DDoS attack is significantly lower than the cost of protective actions. The price for DDoS attacks is decreasing, and the number of services is constantly growing. In 2017 a DDoS attack with use of 1,000 workstations has a cost price about $7 per hour. A customer could order a one-hour DDoS attack of this level on a special web-service for only $25 [32].

Based on a survey conducted among US and Canadian companies in 2016, Incapsula came to a conclusion that 45% of American and Canadian organizations were a subject to attacks, 91% of them faced this threat at least once a year. It also revealed that 43% of respondents have lost credibility of clients, 51% recorded a decline in profits, some companies informed about losses of intellectual property (19%), personal data of customers (33%) or financial information (26%). Incapsula analysts have also estimated an approximate amount of financial losses of a company (direct and indirect) [17]. There is still a consistent trend for reputational costs to prevail over other forms of financial losses. The Kaspersky Security annual report in 2016 have shown results of the survey attended by more than 4,000 companies around the world [1]. It follows from the survey that most companies estimate their reputational costs as the most significant, while the reaction time to the incident plays a critical role: a direct dependence of the financial losses on reaction time of the company have been revealed.

2.2 Classification of DDoS Attacks

In the last publications, including reviews and analytical reports of top companies in the field of information security, various classifications of DDoS attacks are given. Attacks can be distinguished by differences in functionality, final action, use of protocols, and other characteristics. One of the most complete classifications was proposed by Mirkovich *et al.* [34], it takes into account the type of attack, the degree of automation, the frequency of attack, the type of impact, *etc.*

In general there are two types of DDoS attack mechanisms: direct and amplificated (reflected) attacks. Direct attacks try to overload the information resource or communication channel by directly sending a large number of packets to the target (packet flood). Amplification attacks are based on another principle: attackers send packets with small queries to vulnerable resources aimed to get large size responses from them redirected to the victim resource. Direct attacks are dependent on large computational resources and require botnet of a proper size to be used. In contrary amplification attacks are less demanding in resources, however, searching for network vulnerabilities and their correct usage is critical.

Companies specializing in information security use the following classification of DDoS attacks [11, 17, 29, 36]:

- **HTTP flood.** During HTTP flood attacks a great amount of HTTP requests GET are sent on the 80th port of a victim server. It leads to server overloading and inability to handle other requests. Attacks of this type can be aimed at failing the server, as well as overflowing the network bandwidth. In recent years significant complication of HTTP flood DDoS attacks took place: requests can be dynamically self-modified according to certain rules, queries might address not the root of the website, but scripts, consuming a large amount of resources, as well as they can simulate the simplest actions of the user. It significantly impedes HTTP flood attacks detection.

- **SYN flood.** The attacks employ features of the so-called three-way handshake — the procedure that is used to establish a connection between two nodes in the network. Infected computers send multiple SYN requests for connection, and at the same time ignore response requests sent by the victim, thereby creating a queue of "half-open" connections on the target server.

- **UDP flood.** UDP flood attacks overflow the communication channel by sending multiple UDP-packets to the ports of various UDP services.

- **TCP flood.** The attacks are aimed at overwhelming the session/connection tables. It makes legitimate server requests rejected as well.

- **ICMP flood.** An ICMP flood attack appear to be a simple and easy-to-implement method for bandwidth

overflow through multiple sending of ICMP ECHO requests (so-called "pings"). It is usually not very effective, but it works well with resources that are not prepared for DDoS attacks.

- **DNS flood.** DNS flood attacks exploit the vulnerability of DNS systems using UDP. The attacks are based on sending multiple requests to the DNS server overflowing the victim's server with requests and consuming its resources.
- **DNS amplification.** During an attack the intruder's DNS server sends requests in which the target computer is specified as the source address. Thus, the DNS server of the victim suffers from a critical situation.
- **SSDP amplification.** An SSDP amplification attack uses a vulnerability of the SSDP protocol — the feature that is intended to provide network clients the capability to recognize various network services. It initiates a dispatch for the UDP port 1900 by substituting the sender's address in the SSDP protocol request.
- **NTP amplification.** Attacks use the functionality of the monlist request to the NTP server: a list of the last 600 ntpd clients is returned on request. As a result a small request with a fake IP address sends a large UDP stream to the victim.

2.3 Statistics of Threats

Quite regularly (quarterly or annually) a number of top companies in the field of information security publish reports and analytical reviews with results of their research [1, 18, 19, 22–25, 44]. Comparing reviews about DDoS attacks by periods and countries we have defined the following statistically significant dependencies and trends:

1) China, South Korea and the United States are leading in the number of attacks, the number of targets and command and control servers (see Table 1).
2) The number of DDoS attacks approximately doubles every year.
3) The number of simple attacks at the application level is reducing, it gives way to attacks at the level of network protocols and mixed types. This increases the number of multi-vector attacks at the application level, taking into account the specificity of a particular organization resources. The number of vectors in the attacks can reach 5 or more (see Figure 2).
4) There is an increasing demand on the DDoS as a service, the most popular are attacks aimed at putting pressure on the security service rather than causing real harm to the company. The number of attacks for blackmail is growing rapidly and exceeds at least 18% of the total.

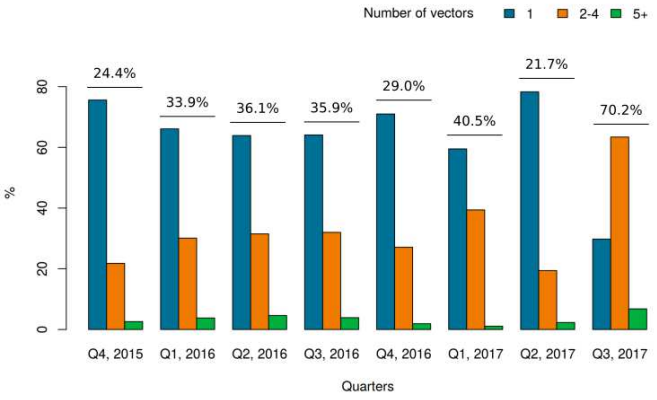


Figure 2: Number of vectors in DDoS attacks for the last two years. Percents of multivector attacks are given on the top of bars. Data were taken from Incapsula reports [18, 19]

5) In the last two years there were a few peak attacks with volume from 400 Gbit/s to 1200 Gbit/s. The average volume of DDoS attacks to major information resources is close to 100 Gbit/s.
6) The maximum duration of attacks is gradually increasing with records of 291 hours of continuous attack (IV quarter of 2016) and 277 hours (II quarter of 2017). However, the average time of attacks remains at a fairly low level — about 3 hours. Short-term attacks account for about 75% of all attacks.
7) A new approach called Pulse Wave technology has been developed. It is capable of increasing the power of a DDoS attack by use of vulnerabilities in hybrid and cloud technologies. Pulse Wave technology implies series of powerful but short duration attacks in the small period of time.
8) Cybercriminals primarily create botnets on the basis of Linux-devices included in the IoT — the share of the networks increased reaching 75%.
9) Botnets participating in DDoS attacks become more sophisticated. Some advanced bots are able to emulate browser behavior, for example, they are able to store cookies and handle JavaScript. The share of such bots reached 42% and tends to increase.
10) The number of attacks with encryption is growing.
11) In 2017 the most popular were SYN flood attacks, they took up to 60% from the total. TCP flood DDoS attacks became less frequent — the share of the attacks decreased from 28% to 12%. Conversely there is a growing demand for HTTP flood DDoS attacks — the share of the attacks increased from 5% to 11%.

Based on the analysis of reports it can be stated that the proportion of simple attacks (primitive to conduct)

Table 1: Percents of targets and command and control servers for DDoS attacks in various countries for the period from the IV quarter (Q4) of 2016 to the III quarter (Q3) of 2017

#	Country	Targets / Attacking servers			
		Q1	Q2	Q3	Q4
1	China	71.6% / 77.0%	47.8% / 55.1%	47.4% / 58.1%	**51.6% / 63.3%**
2	USA	9.1% / 7.3%	13.8% / 11.4%	18.6% / 14.0%	**17.3% / 13.0%**
3	South Korea	9.4% / 7.0%	26.6% / 22.4%	16.4% / 14.2%	**11.1% / 8.7%**
4	Russia	1.7% / 1.8%	1.6% / 1.6%	1.3% / 1.2%	**2.2% / 1.6%**
5	United Kingdom	0.5% / 0.3%	1.1% / 0.8%	2.1% / 1.4%	**2.0% / 1.4%**
6	Hong Kong	1.2% / 0.8%	1.6% / 1.4%	1.0% / 2.4%	**1.6% / 1.3%**
7	Germany	0.6% / 0.8%	0.8% / 0.6%	0.9% / 0.5%	**1.4% / 1.2%**
8	Other	6.0% / 5.3%	6.9% / 6.8%	12.2% / 8.2%	**12.8% / 9.6%**

is reducing. The growth of DDoS attacks complexity indicates that intruders are quick enough at identifying vulnerabilities of network devices and network protocols, they find new ways to exploit vulnerabilities for conducting multi-vector attacks and quickly master at applying new developing techniques (such as botnets based on IoT). It makes it possible to set new records on maximum volume and duration of DDoS attacks. At the same time, the absolute number of simple DDoS attacks does not decrease, as in the global network there are constantly appearing both special services and free tools for their organization. So ordering of simple DDoS attacks becomes available for everyone who is interested in it.

3 Countermeasures for DDoS Attacks

Over the last decade several books and major research works have been published on the subject. The book [2] recommends actions that can be taken before, during and after the attack. The author described the main steps in preparation and conducting of DDoS attacks and discussed how to anticipate attacks and provide protection for computers and networks, minimizing potential devastating consequences. In [37] authors specially paid attention to protection techniques applied in real time on high speed packet transmission with wide channel width. They described a set of possible options for managing web services during the DDoS attack. In the monograph [40] features of network protocols are considered from different points of view, under different conditions of use, including a large number of new scenarios. In [5] various types of DDoS attacks and their implementation are considered, the main stages and mechanisms of creating botnets are given. Particular attention is paid to the methods of statistical analysis and machine learning used to detect and prevent DDoS attacks. DDoS attacks and defences in cloud infrastructure are described in the detailed survey [6].

Various classifications of DDoS attacks protection mechanisms are used for assessment of performance and applicability. They can consider a number of factors, however the most common classifications are based on time of reaction, activity level, deployment location and cooperation degree [33,42]. For example, by time of reaction protection mechanisms can be divided into preventive (can prevent the fact of attack or significantly reduce its damage), real-time (identify the type of attack and filter traffic), and post factum (investigate the incident to improve the means of protection). Classification by deployment location includes protection mechanisms with outer, border and inner location in the network infrastructure.

Here we follow the classification of DDoS protection mechanisms based on "areas of responsibility", in other words we group approaches by party that takes responsibility for applying countermeasures against DDoS attacks. The classification includes:

- **Protection at the level of information resources management.** System administrators of an organization are responsible for countermeasures. The effectiveness and reliability of protection are fully determined by their professional level and network infrastructure capabilities.

- **Protection by specialized hardware and software products.** Responsibility lies with companies developing hardware and software solutions, and in this case, protection level can change only with purchasing new equipment or updating software and it is unlikely to be improved during the attack.

- **Protection by involving security services.** The company that provides security services is responsible for countermeasures. Due to the company's large resources it can vary the protection level according to the situation and use a wide range of protection measures, for example, the channel capacity can be increased with growing attack volume.

There is a brief description for each defence mechanism below; the most prominent business and academia solu-

tions are named. In Subsection 3.4 we discuss advantages and disadvantages of the mechanisms in general.

3.1 Protection at the Level of Information Resources Management

Protection at the level of information resources management includes the analysis of resources demand and bottleneck identification, carrying most of the traffic load in normal network state. Taking into account peculiarities of the server and/or network segment load it is possible to determine potential attack vectors and provide additional network resources for a client, *e.g.* extension of the communication channel bandwidth in advance, increase of server resources, and allocation of resources between several devices.

With the development of cloud technologies and hosting services usability, it is possible to conclude an agreement on providing a client with the necessary amount of resource depending on the load on the infrastructure. It can be a reaction, both to a temporary increase in the number of legitimate users, and to undesired malicious requests. Many providers implement a bandwidth cap, a restriction imposed on the transfer of data over the network, and only a certain type of traffic can consume resources over the time.

Rerouting is another solution that may be used by ISP providers. Most commonly the "black hole" option is used — after filtering malicious traffic is sent to a non-existent interface, which, in effect, leads to its removal. As a result server resources will not be overloaded, however, incoming traffic will still overload the communication channel [14, 20].

In terms of network equipment, many routers allow to configure access control lists (ACLs) to filter out unwanted traffic. The settings provide protection against simple and known DDoS attacks, for example, from ICMP flood attacks. Also, firewalls can be used as additional barriers and confine external networks from internal ones. However the direct protection from DDoS attacks is not included in their functionality.

3.2 Protection by Specialized Hardware and Software Products

Most specialized hardware and software solutions use the concept of protection from DDoS attack "clean pipes" developed by Cisco Systems. It includes the following steps:

- **Baseline Learning.** Traffic profiling with learning intrinsic traffic characteristics.
- **Detection.** Identification of attacks and anomalies.
- **Diversion.** Traffic redirection to the cleaning device.
- **Mitigation.** Filtering traffic to mitigate attacks.
- **Injection.** Entering traffic back into the network and sending to the client.

Cisco Systems used the technology in their products implemented as separate devices or modules.

The Cisco Intrusion Prevention System (IPS) module, deployed on a subnet, is able to eliminate the threat of DDoS attacks that occurs below the location of the sensor device. The system recognizes different signatures of flood attacks and then automatically implements countermeasures specified for them, such as resetting the connection, dropping packets so that they do not reach the target, modifying ACLs on the edge router or the switch next to the affected zone. IPS can also establish a rationing policy, *i.e.* limit the amount of data transferred per unit of time on the edge router.

The Cisco Guard, the product for DDoS attacks protection, consists of two components: a traffic anomaly detector (Cisco Traffic Anomaly Detector) and an anomaly protection tool (Cisco Anomaly Guard). Both components can be implemented as server applications, switch modules or "older" series of Cisco routers (7XXX and above). During the initial deployment, it is required to train the system to capture normal traffic parameters. Subsequently, the trained module is able to detect DDoS attacks by protocols and functionality, and transmit information to the Cisco Anomaly Guard in order to take further action. It should be noted that, although this solution is still on the market, Cisco Systems has stopped further development in this direction, relying on partner companies, and does not support the products since 2014.

Arbor Networks, also actively supporting the "clean pipes" concept, were developing their hardware and software solutions in parallel with the analogues from Cisco Systems. Due to withdraw of the main competitor from the market Arbor Networks significantly expanded the product line and took leading positions, both in development and in the production of solutions for DDoS attacks protection. Similarly to the Cisco Guard their products consists of two main modules:

- Peakflow SP CP is a platform for collecting and analyzing routing information. It differs from Cisco Detector by a feature to control the sampling frequency in analysis of information flows, which allows to use Peakflow SP CP in telecom operators networks and backbone channels.
- Peakflow SP TMS is a threat management system. It suppresses DDoS attacks by a multistage filtering procedure rest upon data received from Peakflow SP CP. Preliminary training of the system is carried out on the basis of statistical data prepared by the laboratory of ASERT, a subsidiary of Arbor Networks.

The Radware offers a comprehensive protection solution DefensePro — a device intended to deal with attacks in real-time including overloading of the Internet channel, attacks on authorization pages, CDN DDoS attacks and

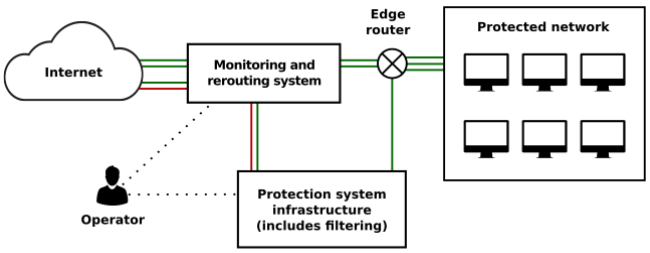


Figure 3: Principle scheme for the DDoS attack protection by involving security services

powerful attacks based on SSL. DefensePro uses a specially designed hardware platform based on OnDemand Switch from Radware with support for network bandwidths up to 40 Gbit/s. It has two hardware components built in:

1) The mechanism for massive DoS and DDoS attacks defence, as well as bulk attacks without affecting the legitimate traffic;

2) The mechanism aimed to accelerate the detection of signatures.

The software of the APSolute Vision device offers centralized management, monitoring and reporting on numerous DefensePro devices. The functions as intrusion detection, network behavior analysis, protection from DDoS, protection against SSL attacks are implemented.

3.3 Protection by Involving Security Services (SaaS Model)

SaaS model security services can be rendered only by large companies with a branched network infrastructure. Kaspersky Lab, one of the leaders in the market, offers the Kaspersky DDoS Prevention — a protecting complex against most types of DDoS attacks, that appears as distributed in the Internet infrastructure of data-clearing centers. The central element of the Kaspersky DDoS Prevention is a sensor installed in the immediate vicinity of the client's information infrastructure. The sensor comes as the software running on a standard x86 architecture server with Ubuntu operating system, it performs traffic analysis without redirecting/changing traffic and examining contents of packages. The statistics is then transferred to the cloud infrastructure of the Kaspersky DDoS Prevention, where customer-specific statistical "profiles" are created based on the collected metadata. The profiles reflect information exchange patterns that are typical for the client taking into account time and calendar fluctuations. Later, during traffic analysis deviations of current characteristics from the statistical profiles serve as indicators of a possible attack.

The second element of the Kaspersky DDoS Prevention is data-clearing centers connected to the largest Internet highways, they are geographically distributed with duplication of functionality in each region of presence. The data-clearing centers are integrated into the cloud infrastructure, however the traffic passing through them remains within the original region. If an attack is detected, the infrastructure allows to divide traffic over several threads decreasing the attack volume and processing each thread separately.

Another key mechanism of DDoS defence is traffic filtering on the provider side. The provider does not only provide an access to the Internet channel, but also filters out "junk" traffic with help of the Kaspersky DDoS Prevention, including traffic that is generated during most flood DDoS attacks. This also makes it difficult to merge DDoS flows into a single powerful attack and reduces the load on data-clearing centers.

If during monitoring of the current traffic deviations from the client's statistical profiles are observed, a warning signal is sent to the DDoS expert of Kaspersky Lab on duty. In case the expert confirms the fact of the attack, the client is notified and the malicious traffic is rerouting to data-clearing centers. Next, the type of attack is determined and type- and resource-specific clearing rules are applied. Traffic comes to servers of data-clearing centers, where filtering by a set of characteristics is applied, *e.g.*, filtering by blacklisting IP addresses, geography, according to statistical criteria or information from HTTP headers. During filtering the sensor continues to analyze client's incoming traffic and if signs of a DDoS attack are still observed, the sensor reports this to data-clearing centers, and the traffic become a subject to in-depth behavioral and signature filtering. Thus, especially complex attacks such as HTTP flood can be detected and neutralized during which the common actions of users on the site are simulated.

When the attack is over, the traffic is again redirected to the client's servers. The Kaspersky DDoS Prevention switches to the standby mode, and the client is provided with detailed report on the incident including a description of the attack progress, diagrams illustrating traffic dynamics, and geographical distribution of attack sources. Another company in the field is the Qrator Labs. The filtering nodes of Qrator Labs are connected to the channels of the largest backbone Internet providers in the USA, Russia, Western and Eastern Europe, and Southeast Asia. The network infrastructure is designed for extreme loads, and an attack on one of the resources should not affect the performance of other resources.

The filtering nodes use the BGP anycast technology to announce their IP addresses. If there is a need to protect client subnets, the corresponding client prefixes can be added to the BGP anycast. Traffic of clients constantly, regardless of the presence/absence of an attack, goes through the Qrator Labs network and is analyzed. "Clean" traffic is redirected to the protected site. This technology allows the filtration nodes to determine which traffic profiles are typical for each resource, and in the event of any deviations, respond immediately. All nodes

of the Qrator Labs network work independently, and if there is a failure of one of them, the traffic of the protected site will not be lost. It will automatically be redirected to the nearest filtering node.

3.4 Practical Aspects of Countermeasures Application

Companies choose defence strategies against DDoS attacks depending on the criticality of protected resources, available financial means, and company-specific security policies. There is no universal tool suitable for every organization. For some cases it will be fair enough to find a qualified stuff for information resources management and entrust network attack protection on it. Later on with business improvement or changes in security policies the protection level can be enforced by more convenient solutions. So, each of the proposed protection mechanisms has its advantages associated with flexibility of use, cost and quality of the staff, but also all of them have a number of substantial shortcomings.

Choosing a solution based on the protection at the level of information resources management one could experience its economic inexpediency, caused by spending money on processing traffic including malicious requests. In addition, legitimate and illegitimate traffic are treated the same, so useful traffic can also be rejected by mistake. Thus, protection by sustainable information resources management allows only to "absorb" DDoS-attacks of low intensity due to a well-designed infrastructure, but does not provide any countermeasures to serious threats. For example, if a DDoS attack is used as diversion for information stealing, there will be no proper actions to prevent data leaks.

Software and hardware solutions are less flexible than other protection systems in some aspects. They are highly dependent on updating and can easily become outdated for newly developed attacks. During a DDoS attack software and hardware products provide just few options to control defence mechanisms and if they fail to negate the attack there is no additional countermeasures. Also, for some companies the price for software and hardware solutions, their maintenance and updates is unaffordable.

Protection from DDoS involving security services is only suitable for very large companies, including providers. The efficiency of these solutions is achieved through the redistribution of computing resources involved in the overall system of protection. However there is no assurance that applied clearing algorithms are appropriate and optimal in each case.

4 In-depth Traffic Analysis and Intellectual Systems

Algorithms for in-depth traffic analysis and intellectual systems represent an advanced field of science and technology, they are developed exponentially and bring promising result for modern challenges in the network security. The methods often serve as mathematical basis for solutions offered by major IT-companies in the field. However they could be possibly used independently for network traffic monitoring and network infrastructure maintenance.

There are two distinct strategies for in-depth traffic analysis: comparison of network traffic characteristics with known templates of attacks (misuse detection systems) and tracing of deviations from common system states (anomaly detection systems). Currently, most studies are aimed at developing anomaly detection methods. Attackers by all means try to complicate cyberattacks detection and bypass security systems, for example, by adding random packets in malicious traffic or using special algorithms for botnets exploitation. Therefore, methods for misuse detection, and methods for anomaly detection require sophisticated intellectual algorithms. Both strategies are discussed in details below.

4.1 Misuse Detection Systems

For misuse detection it is required to determine some abnormal states of the network and describe their characteristics. For each type of attacks a specific pattern, so-called attack signature, is created taking into account the basic parameters of the attack. Any state of the system that does not match any of known patterns is considered normal.

Misuse detection systems have high speed and relatively high accuracy, however, in most cases they only able to detect already known attacks. So the relevance of the system training set is extremely important for good detection performance. If the attack is characterized by a previously unknown set of system parameters, in other words, does not correspond to any of specified signatures, then the attack will be missed. Another drawback of misuse detection systems — these methods require significant amounts of memory for storing signature databases.

As primary characteristics of traffic data streams, the number of packets from different sources, the amount of incoming traffic, the amount of incoming UDP traffic, the amount of incoming TCP traffic, *etc.*, can be used. Some works propose to apply the basic statistics, *i.e.* logical or algebraic functions of initial parameters, to form secondary characteristics. Well-known detection methods differ in approaches to the formation of a space of secondary characteristics that well describe the flows of telemetric traffic data, as well as measures of comparison of these characteristics. As a comparison measure for secondary characteristics, Shannon entropy variants, collision entropy or Renyi's quadratic entropy, Kulback-Leibler discrepancy, generalized entropy or information distance, Jeffreys divergence, squared Hellinger distance and Sibson's information radius are most often used [4, 27, 38, 49].

4.2 Anomaly Detection Systems

Anomaly detection systems try to establish the normal state of the system or its elements, for example, a specific user or a service. If a profile of normal network system functioning is determined, than any system state that is significantly different from the created profile will be defined as anomalous, and a warning for the administrator will be generated. The main advantage of anomaly detection systems is the ability to detect previously unknown DDoS attacks. There is no need to collect, describe and store all types attacks: every system behaviour deviating from common usual will be considered as unwanted and malicious.

However, in contrast to misuse detection systems anomaly detection techniques have worse performance: It shows less accuracy in detection and is memory-consuming, since there is a need to store statistics on large volumes of legitimate traffic. There are a set of possible normal system states if the traffic characteristics are distributed unequally over week, month or year (for example, e-shops traffic will be significantly different for periods of sales and after presentation of new collection than in ordinary days) and this information have to be properly saved and addressed in future. Also a problem may occur if the protected resource become extremely popular in a short period of time. Then detection systems by mistake can consider situation as potentially dangerous and block incoming traffic, as consequences legitime users will partly or totally lost the access to the resource. The effect is known as flash crowds. It is required to apply subtle sensitive algorithms to not confuse flash crowds with malicious traffic.

4.3 Algorithms for In-depth Traffic Analysis

Intelligent systems can be built on the basis of various methods of data mining and machine learning, both approaches specified in Subsections 4.1 and 4.2 are applied for them. The range of methods used is quite wide, to demonstrate the possible directions of intellectual systems development we name a number of works studied various techniques applied to the problem.

For traffic anomaly detection based on deviations from templates of normal system states many techniques have been applied, including principal component analysis [30], wavelets [43], histogram-based modelling [26], support vector machines [7, 48], detection of shifts in spatial-temporal traffic patterns [51]. For more in-depth analysis secondary characteristics are used, they are formed as logical [12], entropy [38], correlation [50] and structural [15] functions of primary traffic characteristics. Various probability measures and special metrics are used to differentiate DDoS attacks from legitimate traffic (including the effect of flash crowds) [8, 21].

Signature analysis, implemented for detecting DDoS attacks, requires to accumulate measurable amount of data for all diverse attacks types and is similar to analogous virus detection tools [33]. A compact and effective technique for formation of network protocols fingerprints was proposed in work [12], preliminary result demonstrated accurate traffic classification.

Artificial neural networks (ANN) are widely used as a part of complex DDoS detection and protection systems. The ability to identify hidden regularities in packets data and create accurate pattern recognition system make them attractive for researchers. However, the accuracy of ANN mainly relies on the relevance of the training set. Chen *et al.* [10] reported high accuracy for an algorithm judging legitime users behaviour from malicious traffic based on auto Turing test and ANN. In work [16] ANN were successfully applied for estimation of the attacking botnet size. Saied *et al.* [39] presented an ANN algorithm for TCP, UDP and ICMP protocols attacks detection based on characteristic traffic patterns. Packets headers were used for training process including source addresses, ID and sequence numbers coupled with source destination port numbers. The algorithm achieved 98% accuracy and performed well on unknown attacks: it failed to detect less than 5% of new attacks from the testing set.

Another approach is relying on the group-related anomalous behavior that botnet exhibits in contrary to normal random communication of ordinary resource users. Traffic source IPs [3] and packet IDs [46] deviations could be used for easy to implement and low cost DDoS detection methods. Chen & Lin [9] proposed a detection system that efficiently identifies anomalous traffic by patterns in hosts homogenous response and group activity. The system applies two-level correlation analysis to reveal sets of hosts with same communication pattern over a long duration, and it may detect malicious traffic produced by even small number of infected hosts. In works [2,13,31,45,47] different clustering techniques were used to group malicious traffic packets and detect bots.

It is challenging to find works describing algorithms for the formation of a space of secondary characteristics, corresponding to the dynamic nature of the network information channel. In the works mentioned above network traffic is considered as a set of static values (*e.g.*, packets per second) without taking into account the dynamic structure of network traffic. In fact, any information channels and their traffic conditions appear to be dynamical systems [28]. We believe, that for their adequate description it is necessary to consider the rates of change in packets flows, and not just the instantaneous values of their loads. Also algorithms allowing automatic adjustment of threshold values of secondary characteristics are rare implemented. It leads to the ongoing need of manual setting of these threshold values and, as a result, to errors in the identification of attack types. No work has been found with preset probabilities of type I and type II errors for identification of attack types.

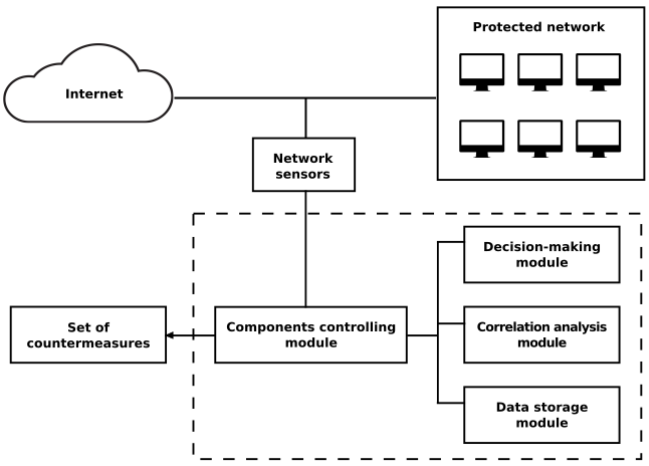


Figure 4: Principle scheme for application of intellectual algorithms in DDoS attack protection systems

4.4 Principle Scheme for Intellectual Methods Application

The intellectual methods discussed above can be implemented both in hardware and software solutions, and in the SaaS model security. Intellectual component can be introduced into the protection mechanism as a separate module. Most of these solutions are based on a similar generalized architecture which includes:

1) **Network sensors.** providing recording of traffic characteristics and revealing some of their patterns;

2) **Components controlling module.** providing interactions between all modules of the system, as well as being an element of interaction between systems for responding to emerging threats and an interface for the system operator;

3) **Decision-making module.** determining whether a package belongs to legitimate or malicious traffic based on identified features;

4) **Data storage module.** containing both signatures for legitimate and malicious traffic detected earlier;

5) **Correlation analysis module.** inspecting for significance newly detected network features, analysing data obtained earlier for current set of features from the decision-making module.

As a basis for decision-making modules and correlation analysis modules researchers use almost the entire spectrum of intellectual decision-making methods.

All of the intelligent algorithms discussed above can be used in various combinations within the overall detection complex. The entire architecture remains the same, but the decision module can be built both on the basis of parallel and sequential study of traffic for the presence of anomalies and their typing. It generally increases the accuracy of incoming information processing, however it may possibly affect the processing speed.

5 Conclusions

DDoS attacks are becoming more sophisticated and massive and cause significant damage to loyal users. The development of attacking techniques is very dynamic and does not keep up with the general pace of development in information technologies. The work aimed at analyzing recent years development of DDoS attacks in order to identify trends with most significant adverse effects to the network infrastructure.

Scientific articles studying DDoS attacks evolution and protection mechanisms against them may lag to some extent and do not reflect the current state of affairs. In addition, this problem is so extensive that research works has become narrowly focused: many articles are devoted to solving one specific problem that arises in a certain situation providing detailed techniques and their applications. To be up-to-date with the last trends in DDoS attacks development it is required to monitor research reports by top IT companies publishing recent statistics and key accidents quarterly or annually as well.

The number of devices connected to the Internet is growing day by day giving wider opportunities for intruders to create large botnets and conduct massive DDoS attacks. The main promising direction for development of DDoS attacks protection systems is their consistency and intellectualization. In this regard, it becomes urgent to develop methods and algorithms for filtering traffic based on in-depth analysis using intelligent systems allowing such analysis for large traffic volumes (more than 100 Gbit/s). Probably different approaches should be combined together for development of new-generation intellectual protection systems taking best from already existing solutions. It could be concluded from the analysis of research articles that there is a tendency to study network traffic as dynamical system with parameters changing in time. Most of the algorithms take in consideration only primary traffic characteristics, however, the study of secondary characteristics may assure DDoS attacks detection using less amount of data or shorter time intervals compared to classical approaches. This hypothesis requires additional research.

Acknowledgments

We thank our colleges A. Krasnov, E. Nadezhdin and D. Nikol'skii for their constructive feedback about the work and helpful comments that greatly improved the manuscript. The work was supported by the Ministry of Education and Science of Russia by lot code 2017-14-579-0002 on the topic: "The development of effective algorithms for detection network attacks based on identifying of deviations in the traffic of extremely large volumes arriving at the border routers of the data network and

creating a sample of software complex for detection and prevention of information security threats aimed at denial of service". The agreement No. 14.578.21.0261 on granting a subsidy at September, 26, 2017, a unique identifier of the work (project) is RFMEFI57817X0261.

References

[1] Kaspersky Security Bulletin 2016, "Review of the year: Overall statistics for 2016," 2016. (https://securelist.com/kaspersky-security-bulletin-2016-executive-summary/76858/)

[2] D. V. V. Sindhu Arumugam and M. V. P. Sumathi, "Detection of botnet using fuzzy c-means clustering by analysing the network traffic," *International Journal of Scientific and Engineering Research*, vol. 6, no. 4, pp. 475–479, 2015.

[3] R. C. Baishya, N. Hoque, and D. K. Bhattacharyya, "DDoS attack detection using unique source IP deviation," *International Journal of Network Security*, vol. 19, no. 6, pp. 929–939, 2017.

[4] D. Balamurugan, S. Chandrasekar, D. Jaya Prakash, and M. Usha, "Analysis of entropy based DDoS attack detection to detect UDP based DDoS attacks in IPv6 networks," *International Journal of Information and Computation Technology*, vol. 3, no. 10, pp. 25–28, 2013.

[5] Dhruba K. Bhattacharyya and Jugal K. Kalita, *DDoS Attacks. Evolution, Detection, Prevention, Reaction and Tolerance*, Boca Raton, USA: Taylor and Francis, 2016.

[6] A. Bonguet and M. Bellaiche, "A survey of denial-of-service and distributed denial of service attacks and defenses in cloud computing," *Future Internet*, vol. 9, no. 3, p. 43, 2017.

[7] C. A. Catania, F. Bromberg, and C. G. Garino, "An autonomous labeling approach to support vector machines algorithms for network traffic anomaly detection," *Expert Systems With Applications*, vol. 39, no. 2, pp. 1822–1829, 2012.

[8] S. Chawla, M. Sachdeva, and S. Behal, "Discrimination of DDoS attacks and flash events using Pearson's product moment correlation method," *International Journal of Computer Science and Information Security*, vol. 14, no. 10, pp. 382–389, 2016.

[9] C. M. Chen and H. C. Lin, "Detecting botnet by anomalous traffic," *Journal of information security and applications*, vol. 21, pp. 42–51, 2015.

[10] J. H. Chen, M. Zhong, F. J. Chen, and A. D. Zhang, "DDoS defense system with Turing test and neural network," in *IEEE International Conference on Granular Computing (GrC'12)*, pp. 38–43, Hangzhou, China, Aug. 2012.

[11] Aruba. Hewlett Packard Enterprise Company, "The Internet of Things: Today and tomorrow," 2016. (https://www.arubanetworks.com/assets/eo/HPE_Aruba_IoT_Research_Report.pdf)

[12] M. Crotti, M. Dusi, F. Gringoli, and L. Salgarelli, "Traffic classification through simple statistical fingerprinting," in *ACM SIGCOMM Computer Communication Review*, vol. 37, pp. 5–16, Kyoto, Japan, August 2007.

[13] C. J. Dietrich, C. Rossow, and N. Pohlmann, "CoCoSpot: clustering and recognizing botnet command and control channels using traffic analysis," *Computer Networks*, vol. 57, no. 2, pp. 475–486, 2013.

[14] C. Dietzel, A. Feldmann, and T. King, "Blackholing at ixps: On the effectiveness of ddos mitigation in the wild," in *International Conference on Passive and Active Network Measurement*, pp. 319–332, Heraklion, Crete, Greece, Mar. 2016.

[15] V. S. Galayev, A. E. Krasnov, D. N. Nikol'skii, and D. S. Repin, "The space of structural features for increasing the effectiveness of algorithms for detecting network attacks, based on the detection of deviations in traffic of extremely large volumes," *International Journal of Applied Engineering Research*, vol. 12, pp. 10781–10790, 2017.

[16] B. B. Gupta, R. C. Joshi, and M. Misra, "ANN based scheme to predict number of zombies in a DDoS attack," *International Journal of Network Security*, vol. 14, no. 2, pp. 61–70, 2012.

[17] Imperva, Inc, "Incapsula's 2014 DDoS impact report," 2014. (https://lp.incapsula.com/ddos-impact-report.html)

[18] Imperva, Inc, "Global DDoS threat landscapy," Q2 2017. (https://www.incapsula.com/ddos-report/ddos-report-q2-2017.html)

[19] Imperva, Inc, "Global DDoS threat landscapy," Q3 2017. (https://www.incapsula.com/ddos-report/ddos-report-q3-2017.html)

[20] K. Kalkan and F. Alagöz, "A distributed filtering mechanism against DDoS attacks: ScoreForCore," *Computer Networks*, vol. 108, pp. 199–209, 2016.

[21] L. Ke, Z. Wanlei, L. Ping, and L. Jianwen, "Distinguishing DDoS attacks from flash crowds using probability metrics," in *IEEE Third International Conference on Network and System Security*, pp. 9–17, Shanghai, China, Oct. 2009.

[22] A. Khalimonenko and O. Kupreev, "DDOS attacks in Q1 2017. kaspersky lab," 2017. (https://securelist.com/ddos-attacks-in-q1-2017/78285/)

[23] A. Khalimonenko, O. Kupreev, and T. Ibragimov, "DDoS attacks in Q2 2017. kaspersky lab," 2017. (https://securelist.com/ddos-attacks-in-q2-2017/79241/)

[24] A. Khalimonenko, O. Kupreev, and K. Ilganaev, "DDoS attacks in Q3 2017. kaspersky lab," 2017. (https://securelist.com/ddos-attacks-in-q3-2017/83041/)

[25] A. Khalimonenko, J. Strohschneider, and O. Kupreev, "DDOS attacks in Q4 2016. Kaspersky Lab," 2016. (https://securelist.com/ddos-attacks-in-q4-2016/77412/)

[26] A. Kind, M. P. Stoecklin, and X. Dimitropoulos, "Histogram-based traffic anomaly detection," *IEEE Transactions on Network and Service Management*, vol. 6, no. 2, pp. 110–121, 2009.

[27] A. Koay, A. Chen, I. Welch, and W. K. Seah, "A new multi classifier system using entropy-based features in DDoS attack detection," in *Information Networking (ICOIN), 2018 IEEE International Conference*, pp. 162–167, Chiang Mai, Thailand, Jan. 2018.

[28] A. E. Krasnov, E. N. Nadezhdin, V. S. Galayev, E. A. Zykova, D. N. Nikol'skii, and D. S. Repine, "DDoS attack detection based on network traffic phase coordinates analysis," *International Journal of Applied Engineering Research*, vol. 13, no. 8, pp. 5647–5654, 2018.

[29] V. Kuskov, M. Kuzin, Ya. Shmelev, D. Makrushin, and I. Grachev, "Honeypots and the Internet of Things. securelist. Kaspersky Lab.," 2017. (`https://securelist.com/honeypots-and-the-internet-of-things/78751/`)

[30] Y. Liu, L. Zhang, and Y. Guan, "Sketch-based streaming PCA algorithm for network-wide traffic anomaly detection," in *Distributed Computing Systems (ICDCS), 2010 IEEE 30th International Conference*, pp. 807–816, Genova, Italy, June 2010.

[31] W. Lu, G. Rammidi, and A. A. Ghorbani, "Clustering botnet communication traffic based on n-gram feature selection," *Computer Communications*, vol. 34, no. 3, pp. 502–514, 2011.

[32] D. Makrushin, "The cost of launching a DDoS attack. securelist. Kaspersky Lab," 2017. (`https://securelist.com/the-cost-of-launching-a-ddos-attack/77784/`)

[33] L. Malina, P. Dzurenda, and J. Hajny, "Testing of DDoS protection solutions," in *Security and Protection of Information*, pp. 113–128, Brno, Czech Republic, May 2015.

[34] J. Mirkovic and P. Reiher, "A taxonomy of DDoS attack and DDoS defense mechanisms," in *ACM SIGCOMM Computer Communication Review*, vol. 34, pp. 39–53, Portland, USA, 2004.

[35] MIT, *The Internet of Things*, Business Reports, MIT Technology Review, Aug. 2014.

[36] Radware, "DDoS attack definitions — DDoSPedia: glossary that focuses on network and application security terms with many DDoS-related definitions," 2017. (`https://security.radware.com/ddos-knowledge-center/ddospedia/rudy-r-u-dead-yet/`)

[37] S. V. Raghavan and E. Dawson, eds., *An Investigation into the Detection and Mitigation of Denial of Service (DoS) Attacks: Critical Information Infrastructure Protection*, India: Springer Science & Business Media, 2011.

[38] V. S. Reddy, K. D. Rao, and P. S. Lakshmi, "Efficient detection of DDoS attacks by entropy variation," *IOSR Journal of Computer Engineering*, vol. 7, no. 1, pp. 13–18, 2012.

[39] A. Saied, R. E. Overill, and T. Radzik, "Detection of known and unknown DDoS attacks using artificial neural networks," *Neurocomputing*, vol. 172, pp. 385–393, 2016.

[40] Chris Sanders, *Practical packet analysis: Using Wireshark to solve real-world network problems*, San Francisco: No Starch Press, 2017.

[41] A Neustar security solution exclusive report, "Worldwide DDoS attacks & cyber insights research report: Taking back the upper hand from attackers," 2017. (`https://hello.neustar.biz/201705-Security-Solutions-DDoS-SOC-Report-LP.html`)

[42] K. Singh, K. S. Dhindsa, and B. Bhushan, "Distributed defense: An edge over centralized defense against DDoS attacks," *International Journal of Computer Network and Information Security*, vol. 9, no. 3, pp. 36, 2017.

[43] V. Srihari and R. Anitha, "DDoS detection system using wavelet features and semi-supervised learning," in *Security in Computing and Communications. SSCC 2014. Communications in Computer and Information Science*, pp. 291–303, Delhi, India, Sept. 2014.

[44] Akamai Technologies, "Q3 2016 state of the internet: Security report," 2016. (`https://content.akamai.com/pg7407-soti-security-report-q3-en.html`)

[45] D. S. Terzi, R. Terzi, and S. Sagiroglu, "Big data analytics for network anomaly detection from netflow data," in *Computer Science and Engineering (UBMK), 2017 IEEE International Conference*, pp. 592–597, Antalya, Turkey, Oct. 2017.

[46] T. M. Thang and V. K. Nguyen, "Synflood spoof source DDoS attack defence based on packet ID anomaly detection — PIDAD," *Software Networking*, vol. 2017, no. 1, pp. 213–228, 2017.

[47] K. Wang, C.-Y. Huang, S.-J.Lin, and Y.-D. Lin, "A fuzzy pattern-based filtering algorithm for botnet detection," *Computer Networks*, vol. 55, no. 15, pp. 3275–3286, 2011.

[48] J. Ye, X. Cheng, J. Zhu, L. Feng, and L. Song, "A DDoS attack detection method based on SVM in software defined network," *Security and Communication Networks*, vol. 2018, pp. Article ID 9804061, 8 pages, 2018.

[49] S. Yu, W. Zhou, R. Doss, and W. Jia, "Traceback of DDoS attacks using entropy variations," *IEEE Transactions on Parallel and Distributed Systems*, vol. 22, no. 3, pp. 412–425, 2011.

[50] S. Yu, W. Zhou, W. Jia, S. Guo, Y. Xiang, and F. Tang, "Discriminating DDoS attacks from flash crowds using flow correlation coefficient," *IEEE Transactions on Parallel and Distributed Systems*, vol. 23, no. 6, pp. 1073–1080, 2012.

[51] J. Yuan and K. Mills, "Monitoring the macroscopic effect of DDoS flooding attacks," *IEEE Transactions on Dependable and secure computing*, vol. 2, no. 4, pp. 324–335, 2005.

Biography

Vladimir Galyaev received his specialist's degree in Math and his Ph.D. degree in Information Technologies from Dagestan State University (Makhachkala, Russia) in 2000 and 2004, respectively. Since 2008 he is the Head of the Department of Information Technologies and Information Security of Dagestan State University of National Economy (Makhachkala, Russia). In 2013 he became an associate professor in Dagestan State University of National Economy, he gives lectures for students specialized in information security. Since 2017 he works as senior research fellow in State Institute of Information Technologies and Telecommunications. His research interests include network security, data protection, steganography and e-learning.

Evgenia Zykova received her specialist's degree in Biochemical Physics from Siberian Federal University (Krasnoyarsk, Russia) in 2011, and she received her Ph.D. degree in Biophysics from Institute of Cell Biophysics RAS (Pushchino, Russia) in 2016. She is a research fellow in State Institute of Information Technologies and Telecommunications. Her research interests include data statistical analysis, machine learning and data mining techniques.

Dmitry Repin graduated from Kaliningrad Higher Engineer School of Engineer Troops (Kaliningrad, Russia) in 1992 and from Moscow Power Engineering Institute in 2005, he have a major in computer networks and telecommunications. In 2008 he received his Ph.D. degree in Technical Sciences in Moscow State Mining University. In 2018 he graduated from the Academy of Information Systems by specialty of information security. He is a Deputy Director of State Institute of Information Technologies and Telecommunications. For many years Dr. Repin participated in federal projects devoted to development and implementation of advanced information technologies for high-performance data processing, storage and high-speed data transmission in computer telecommunications networks of Russian Ministry of Education and Science. His main areas of expertise are telecommunication technologies, network security, software-defined networking, National Research and Education Networks located in or connected to Russia.

Denis Bokov received his specialist's degree in Comprehensive information security of automated systems from Moscow Engineering Physics Institute in 2007, and he received his Ph.D. degree in Social Philosophy from Moscow State Regional University in 2011. Also he received a specialist's degree in Jurisprudence from Institute for Socio-Economic Forecasting and Modeling (Moscow, Russia) in 2006. He is the Director of State Institute of Information Technologies and Telecommunications. Dr. Bokov is managing a few federal programs aimed at development of the information infrastructure to ensure coordination of activities in the field of informatization of subordinate institutions of the Ministry of Education and Science of Russia. His main areas of expertise are technical, social and law aspects of higher education informatization, e-learning, data protection and National Research and Education Networks located in or connected to Russia.

A Secure and Reliable Data Transmission Scheme in Wireless Body Area Network

Huaijin Liu, Yonghong Chen, Hui Tian, Tian Wang, and Yiqiao Cai
(Corresponding author: Yonghong Chen)

College of Computer Science and Technology, Huaqiao University
Xiamen 361021, China
(Email: lhjhqdx@163.com)

(Received Jan. 5, 2018; Revised and Accepted Apr. 21, 2018; First Online Mar. 9, 2019)

Abstract

In view of the privacy protection and shadow effect of wireless body area network (WBAN), we propose a secure and reliable data transmission scheme. In the first place, on the basis of the characteristics of WBAN, we propose a reasonable relay transmission strategy, which uses the time-varying model to model the channel and select the relay node based on the principle of load balancing, to solve the problem of how reasonable and efficient use of relay nodes, thus improving the energy efficiency of relay transmission. In addition, in order to solve the problem of secure transmission of physical data in WBAN, a new authentication and key agreement protocol is proposed. Through in-depth analysis, it is verified that the proposed scheme conforms to the highest security level defined by IEEE 802.15.6 body area network standard, which can ensure the confidentiality and integrity of information while satisfying the demand of data reliability, and has high application value.

Keywords: Load Balancing; Privacy Protection; Reliability; Time-Varying Model; Wireless Body Area Network

1 Introduction

In recent years, with the rapid development of wireless communications, micro-sensor equipment and artificial intelligence, an emerging, human-centered wireless sensor network-wireless body area network (WBAN) came into being. WBAN is mainly composed of a variety of sensor nodes attached to the human body that continuously perceive human physiological data and a coordinator that collects and processes various perceived data.

Due to the asymmetry between the coordinator and the sensor nodes, a standard single-hop star topology is widely used in traditional WBAN. However, in the actual situation, the human body will cause the wireless link between the sensor node and the coordinator to be blocked, resulting in reduced data transmission reliability. In order to reduce the shadow effect of the human body on the channel, the use of relay transmission mechanism can greatly reduce the link outage probability.

However, the introduction of a relay transmission mechanism will bring additional energy overhead, which will further shorten the lifetime of WBAN. Therefore, how to use the relay node reasonably and efficiently is great importance to improve the energy efficiency of relay transmission. In addition, WBAN in the transmission of data, security is also very important. Since the data transmitted by WBAN are physiological parameters that are closely related to the human body, the confidentiality and integrity of the data are indispensable.

In order to ensure the reliability and security of data transmission in WBAN, we propose a secure and reliable data transmission scheme for WBAN. The main contributions of the scheme are the following:

1) Using the time-varying model to establish the wireless human body channel, according to the time-varying prediction model to determine whether the sensor node needs to allocate the relay node, to solve the problem of relay timing judgment.

2) A relay transmission strategy based on load balancing is proposed to solve the problem of relay node selection and improve the energy efficiency.

3) According to the transmission mode of different links, this paper proposes a new authentication and key agreement protocol, which solves the problem of data security transmission.

The rest of this article is organized as follows. Section 2 reviews the related work. Section 3 describes the system model and design goals. Section 4 presents this proposed safety and reliability scheme. Section 5 describes the safety analysis, and Section 6 describes the simulation results. Summarized in Section 7.

2 Related Work

A large number of personal data collected by WBAN are important information about the security and privacy of users, and it is of great significance to explore how to ensure that these data are transmitted securely to the relevant medical institutions. In the literature [3,19], the security requirements of WBAN are analyzed, and the security objectives of WBAN system are mainly to ensure the confidentiality, integrity, authenticity and freshness of the data. Because the sensor nodes have strict low power limits, it is challenging to meet these security requirements. If the use of complex security encryption measures, will inevitably lead to excessive energy consumption, and easy to affect the normal communication of the sensor nodes. IEEE 802.15.6 body area network standard defines a multi-level security level of communication, each of which corresponds to a different level of protection and frame format [15]:

1) Level 0: Unsafe communication, no data is authenticated during communication, and no integrity protection;

2) Level 1: Only authentication, data transmission in the security authentication mode, but the data is not encrypted;

3) Level 2: Authentication and encryption, which is the highest security level of communication mode. In order to ensure the safe transmission of data, the literature [28,30] through asymmetric encryption technology to encrypt the data, but these schemes have high computational complexity, not suitable for WABN. Literature [13,16] proposed a number of lightweight security encryption scheme, which can effectively ensure the safe transmission of private data, but these methods require a large storage space and does not meet the reliability of the data. In addition, in order to resist the presence of attacks in WBAN, the literature [9,20] proposed to use the time-varying human physiological signal to establish the symmetric key, reduces the key management of symmetric encryption algorithm, but this method is limited to the human body sensor symmetry of the network topology.

On the other hand, the traditional WBAN usually uses the standard star topology to transmit the data, but in the actual process, because of the shadow effect of the human body structure, in the signal transmission process will cause great path loss [2, 23]. In order to optimize the topology of WBAN, the relay nodes can be introduced into the network to improve the reliability of data. Gorce *et al.* [10] conducted a theoretical study on the reliability of relay transmission mechanism in WBAN, and then compared the relay transmission mechanism with the single-link transmission mechanism and the two-hop transmission mechanism respectively. It is proved that WBAN adopts relay transmission mechanism can be more effective than the other two mechanisms to improve reliability. Errico *et al.* [6] proposed a performance evaluation method of relay transmission mechanism for WBAN, and based on the measured data of the wireless human body channel under the daily activities of the human body, it is proved that relay transmission mechanism can greatly reduce the link outage probability. However, the literature [6, 10] does not give the specific implementation strategy of relay transmission in WBAN. Abbasi *et al.* [1] proposed a relay transmission strategy to improve the reliability of WBAN. The strategy uses a dynamic contention-based relay node selection mechanism, that is, the first relay node that makes feedback on the request from the source node is selected as the relay node of the source node. The results show that the strategy can effectively improve the reliability of transmission while reducing the delay. Hara *et al.* [12] also proposed a relay transmission strategy to improve the communication reliability of WBAN. This strategy is based on the principle of ?low interrupt correlation? to make a more reasonable choice of relay nodes. The results show that this kind of relay node selection method can improve the reliability in the weaker dynamic scenario. Although the research [1,12] proposed a specific relay transmission strategy for WBAN, they only verified the reliability of the strategy and did not examine the energy efficiency of the strategy. The study [18] evaluated the energy consumption of the proposed relay transmission strategy and found the high energy consumption problem of the relay node, but did not give the corresponding solution to the problem.

By analyzing and summarizing the above research results, we can see that only symmetric encryption technology is suitable for WBAN sensor nodes with low power consumption and limited storage resources. In addition, the main reason for the high energy consumption of relay nodes is that the relay nodes are not allocated reasonably in the relay transmission process. Therefore, in this paper, we propose a load balancing based relay transmission strategy to solve the problem of high energy consumption of relay nodes. At the same time, combined with the proposed authentication and key agreement mechanism, it provides the security guarantee for data transmission.

3 System Models and Design Goals

3.1 Network Model

WBAN mainly includes intra-body and extra-body two parts of the application structure, as shown in Figure 1. In this paper, we mainly study the safety and reliability of intra-body network. The intra-body part is mainly composed of a coordinator and each sensor node attached to the human body surface. Each sensor node continuously senses physiological information and periodically transmits the perceived data to the coordinator. The co-

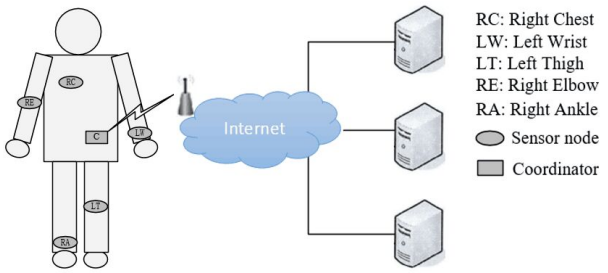


Figure 1: Wireless body area network application structure diagram

ordinator is responsible for collecting the data perceived by each sensor node and then processing and transmitting the data to the external network. The extra-body part mainly includes base stations, communication networks and remote servers. The coordinator sends the collected physiological information to the base station in extra-body, and the base station sends the information to the remote storage server through the external network.

3.2 Channel Model

In the study, based on the recommendation of the IEEE 802.15.6 Working Group [8], we use the time-varying model [5] to model the channel. The short-term average channel gain $\bar{G}(n)$ is a random variable that describes the slow fading characteristics of channel due to the human shadow effect. Based on the statistical analysis of a large number of measured channel gains, it is proved that the short-term average channel gain $\bar{G}(n)$ follows the normal distribution:

$$\bar{G}(n) \mid_{dB} \sim N(\mu_{\mathbb{S}}, \sigma_{\mathbb{S}}^2). \tag{1}$$

Where $\mathbb{S}$ is a specific application scenario, $\mu_{\mathbb{S}}$ and $\sigma_{\mathbb{S}}^2$ are the mean and standard deviation for the specific scenario, respectively. Assume that the link between the sensor nodes S_i and S_j is denoted as S_{ij} and the short-term average channel gain of the link is represented by the random variable $\bar{G}_{ij}$. The random variables $\bar{G}_{ij}(m_i)$ and $\bar{G}_{ij}(m_i+k)$ represent the average channel gain of the link in slot m_i and slot $m_i + k$, respectively. According to Equation (1), the two random variables are Normal distribution:

$$\bar{G}_{ij} \mid_{dB} \sim N(\mu_{ij}, \sigma_{ij}^2), \bar{G}_{ij}(m_i+k) \mid_{dB} \sim N(\mu_{ij}, \sigma_{ij}^2). \tag{2}$$

Due to the temporal autocorrelation of the channel, there is a certain temporal correlation between the two variables, so their joint probability distributions can be expressed as follows:

$$\begin{aligned}(\bar{G}_{ij}(m_i), \bar{G}_{ij}(m_i+k)) \mid_{dB} \sim N(\mu_{ij}, \mu_{ij}, \sigma_{ij}^2, \sigma_{ij}^2, \rho_{ij}(k)),\\ \rho_{ij}(k) = \frac{E\{[\bar{G}_i(m_i) - \mu_{ij}][\bar{G}_i(m_i+k) - \mu_i j]\}}{\sigma i j^2}\end{aligned} \tag{3}$$

We call Equation (3) denote the time-varying model, where $\rho_{ij}(k)$ represents the correlation coefficient between $\bar{G}_{ij}(m_i)$ and $\bar{G}_{ij}(m_i+k)$. Under the premise of known $\bar{G}_{ij}(m_i)$, the probability distribution of the random variable $\bar{G}_i(m_i+k)$ can be obtained by further derivation:

$$\begin{aligned}\bar{G}_i(m_i+k) \mid_{dB} \sim N((1-\rho_i(k)) \cdot \mu_i + \rho_i(k) \cdot \bar{G}_i(m_i),\\ (1-\rho i^2(k))\sigma_i^2).\end{aligned} \tag{4}$$

Equation (4) shows that the outage probability $Pout_i(m_i+k)$ in the next transmission slot can be predicted based on the channel state $\bar{G}_i(m_i)$ in the current time slot:

$$\begin{aligned}Pout_i(m_i+k) &= Prob(\bar{P}_i(m_i+k) < \bar{P}^*)\\ &= Prob(\bar{G}_i(m_i+k) + P_t < \bar{P}^*)\\ &= Prob(\bar{G}_i(m_i+k) < \bar{G}^*)\\ &= \int_{-\infty}^{\bar{G}^*} f(\bar{G}_i(m_i+k))d\bar{G}_i\\ &= \phi(\frac{\bar{G}^* - (1-\rho_i(k)) \cdot \mu_i - \rho i(k) \cdot \bar{G}_i(m_i)}{\sqrt{(1-\rho i^2(k))} \cdot \sigma i})\end{aligned} \tag{5}$$

We call Equation (5) denote the time-varying prediction model, where $\bar{P}_i(m_i+k)$ represents the average received signal power, $\bar{P}^*$ represents the predefined receive power threshold, $\bar{G}^*$ is expressed as the link interrupt threshold, and satisfies $\bar{G}^* = \bar{P}^* - P_t$, $f(\bar{G}_i(m_i+k)$ denotes the probability density function of $\bar{G}_i(m_i+k)$, and $\phi(\cdot)$ denotes the standard normal distribution function.

3.3 Threat Model

Because of the openness of wireless communication and the importance of transmitting information, WBAN is vulnerable to attack. These security threats are mainly from the following attacks.

Eavesdropping attacks: Since the openness of wireless channel transmission, so the attacker can eavesdrop any messages transmitted between nodes and obtains sensitive or valuable information by analysis.

Tampering attack: An attacker can remove or replace the eavesdropping message, and then send the tampered message to the original recipient to achieve some illegal purpose.

Camouflage attack: If the attacker eavesdropped to the legitimate sensor node or coordinator identity information, then he can be disguised as a legal node through the identity information to deceive.

Replay attack: The attacker to use network monitoring or other ways to steal data packets and resend a destination host has received packets, to achieve the purpose of deception system.

Man-in-the-middle attack: The attacker use a variety of technologies to intercept network data flow, and then to steal the information and illegal tampering, thus deceiving both ends of the authorized client.

Denial of service attack: An attacker sends a large number of packets to consume the network bandwidth and resources of the target server so that it can run out of power and can not continue to work.

3.4 Design Goals

In WBAN, because of the particularity of node structure, the particularity of function and the particularity of its application environment, WBAN not only to meet the basic security objectives of the network, but also to ensure the reliability of the data. A secure and reliable WBAN architecture should be able to provide the following services.

Data reliability: Due to the particularity of the wireless human channel, the human body's own blocking effect on the wireless channel will lead to a strong shadow effect, thus reducing the arrival rate of data packets, affecting the reliability of data.

Data confidentiality: Patient information in the transmission process should be encrypted, can not directly to the user's privacy information leaked to internal or external users.

Data integrity: If there is no relevant security mechanism to protect the integrity of the data, the attacker is easy to tamper with or forge the original data segment to destroy the integrity of the data.

Authentication: Since the coordinator collects the perceptual information from each sensor node in the body, the coordinator must have the ability to validate the data source.

4 The Proposed Scheme

In this section, we propose a secure and reliable data transmission scheme for WBAN. First of all, the scheme uses the time-varying prediction model to judge the relay timing, and then select the relay node according to the principle of load balancing to ensure the energy efficiency of relay transmission on the premise of reliability. At the same time, according to the different ways of link transmission, respectively, a two-party and three-party authentication and key agreement protocol are proposed to ensure the secure communication of data.

4.1 Judgment of Relay Timing

In the time-varying model, the coordinator C determines whether or not a relay node needs to be assigned to the sensor node in the next superframe according to the channel state. We use the typical TDMA superframe structure to allocate time slots, as shown in Figure 2, each superframe is divided into three parts, namely, the transmission period, the forwarding period and the sleep period. During the transmission period, the sensor node sends the perceptual data to the coordinator and the relay node in the allocated time slot. During the forwarding period, the relay node will forward the received packets to the coordinator C within the allocated forwarding time slot. During sleep, all nodes go to sleep.

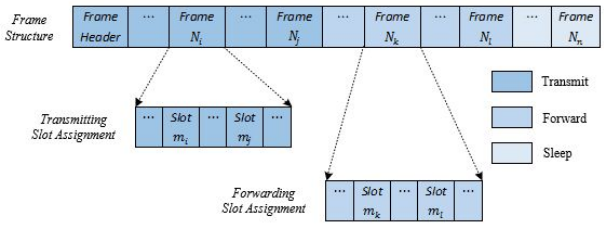


Figure 2: TDMA superframe structure

It is assumed that the sensor node S_i has transmitted the data to the coordinator C in the m_i-th time slot of the current superframe. The coordinator C obtains the average channel gain value of the link $S_i - C$ in the current transmission slot according to the RSSI (received signal strength indicator) value of the received packet, denoted $\bar{G}_i(m_i)$. If the coordinator C assigns the (m_i+k)-th time slots in the next superframe as the next transmission slot to S_i, the outage probability $Pout_i(m_i+k)$ in the next transmission slot can be predicted according to Equation (6):

$$\begin{aligned} Pout_i(m_i+k) &= \int_{-\infty}^{\bar{G}^*} f(\bar{G}_i(m_i+k))d\bar{G}_i \\ &= \phi(\frac{\bar{G}^* - (1-\rho_i(k))\cdot\mu_i - \rho i(k)\cdot\bar{G}_i(m_i)}{\sqrt{(1-\rho i^2(k))}\cdot\sigma i}) \end{aligned} \tag{6}$$

Where $f(\bar{G}_i(m_i+k_i))$ is the probability density function of the random variable $\bar{G}_i(m_i+k)$. When the coordinator C calculates the outage probability $Pout_i(m_i+k)$, it is possible to determine whether S_i needs to allocate the relay node in the next transmission slot according to Equation (7):

$$\begin{cases} Pout_i(m_i+k) > \sigma, & \text{allocate relay nodes} \\ Pout_i(m_i+k) \le \sigma, & \text{do not allocate relay nodes} \end{cases} \tag{7}$$

Where δ is the predefined threshold for relay allocation.

4.2 Selection of Relay Node

Suppose there are N sensor nodes, denoted as $R = \{S_i \mid i = 1, 2, \cdots, N\}$. The coordinator C predicts the link quality of the next transmission slot of the sensor node set R according to Equation (7) to obtain the set $R_1 = \{S_i \mid Pout_{S_i} > \delta, S_i \in R\}$ that needs to allocate the relay node and the set $R_2 = \{S_j \mid Pout_{S_j} < \delta, S_j \in R\}$ that does not need to allocate the relay node. For each sensor node in the set R_1, we use load balancing principle to allocate the relay nodes to maximize the energy efficiency of the relay nodes. Assume that the coordinator

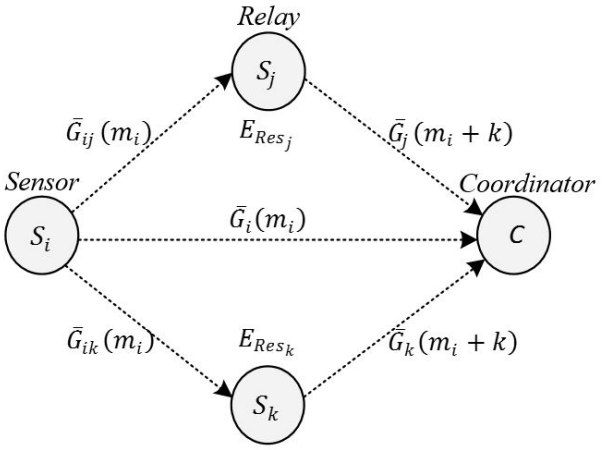


Figure 3: Relay transmission model based on load balancing

C needs to allocate a relay node for the sensor node S_i. First, the node S_i needs to broadcast a request so that the node who received the request will upload the unique identifier id and the remaining energy E_{Res} to the coordinator C. The coordinator C obtains the node set R_3 which makes the request feedback, and then performs the intersection operation with R_2 to get the candidate set of relay nodes R_4:

$$R_4 = R_2 \cap R_3 \tag{8}$$

Finally, the coordinator C allocates the lowest cost node as the optimal relay node to S_i according to the residual energy E_{Res} of the node in the candidate set R_4, thus extending the network lifetime. Figure 3 shows a simple diagram based on load balancing relay transmission model in which it is assumed that the residual energy E_{Res_j} the relay node S_j is greater than the residual energy E_{Res_k} the relay node S_k, so the coordinator C chooses S_j as the optimal relay node to allocate to S_i. When the link $S_i - C$ and the link $S_i - S_j - C$ are interrupted at the same time, the joint link of S_i is interrupted. Therefore, the outage probability $Pout_i(j)$ of S_i is:

$$\begin{aligned} &Pout_i(j) \\ &= Prob(\bar{G}_i(m_i) < \bar{G}^*) \times Prob(\bar{G}_{ij}(m_i) < \bar{G}^*) \\ &\quad + Prob(\bar{G}_i(m_i) < \bar{G}^*) \times Prob(\bar{G}_{ij}(m_i) \geq \bar{G}^*) \\ &\quad \times Prob(\bar{G}_j(m_i + k) < \bar{G}^*) \end{aligned} \tag{9}$$

4.3 Secure Transmission of Messages

After the coordinator C assigns the relay node to the sensor node S_i, S_i uploads the perceptual data to C. During data upload, the sensor node S_i and the coordinator C need to perform authentication and key agreement to ensure the security of data transmission. The system needs to be initialized before the key agreement. Therefore, we divide the data security transmission into system initialization phase, authentication and key agreement phase and data transmission phase.

4.3.1 System Initialization Phase

In the initialization phase, the system administrator (SA) needs to deploy some parameters for each sensor node S_i and coordinator C. The specific steps are as follows:

Step 1: SA assigns a unique identifier id_i and id_c to each sensor node S_i and coordinator C.

Step 2: SA selects a preshared key K_{ic} for each sensor node S_i and coordinator C.

Step 3: SA defines a one-way hash function $h(\cdot)$ and a keyed message authentication code $MAC_k(\cdot)$.

Step 4: SA selects a symmetric encryption algorithm $E_k(\cdot)$ and a pseudo-random function $f(\cdot)$.

Step 5: Finally, SA assigns the parameters $\{K_{ic}, H(\cdot), MAC_k(\cdot), E_k(\cdot), f(\cdot)\}$ to S_i and C.

4.3.2 Authentication and Key Agreement Phase

In the single link transmission process, we assume that the sensor node S_i communicates with the coordinator C. The proposed two-party authentication and key agreement protocol for single link transmission is shown in Figure 4. In the relay transmission process, it is assumed that the sensor node S_i communicates with the coordinator C through a relay node S_j. The proposed three-party authentication and key agreement protocol for relay link transmission is shown in Figure 5, described as follows:

Step 1: S_i Generate a random number k, calculate $x = Enc_{K_{ic}}(id_i, k)$ and $H_1 = h(id_i, k)$, then send the message $Mes_1 = (id_i, x, H_1)$ to S_j.

Step 2: S_j after receiving the message Mes_1, calculate $H_2 = MAC_{K_{jc}}(id_i, id_j, x, H_1)$ and send messages $Mes_2 = (id_i, id_j, x, H_1, H_2)$ to C.

Step 3: C after receiving the message Mes_2, calculate $H_2^* = MAC_{K_{jc}}(id_i, id_j, x, H_1)$ and verify that $H_2^* = H_2$ is equal. If the authentication fails, stop the session, otherwise C decrypt $Dec_{K_{ic}}(x) = id_i, k$, and then calculate $H_1^* = h(id_i, k)$ and verify that $H_1^* = H_1$ is equal. If the authentication fails, stop the session, otherwise C will generate a random number $r \in Z_p$, calculate $SK = f(k, r, id_i, id_c, K_{ic})$, $y = Enc_{K_{ic}}(id_i, r)$, $H_3 = h(id_i, k, r)$ and $H_4 = MAC_{K_{jc}}(id_i, id_j, id_c, y, H_3)$, and finally send the message $Mes_3 = (id_i, id_j, id_c, y, H_3, H_4)$ to S_j.

Step4: S_j after receiving the message Mes_3, calculate $H_4^* = MAC_{K_{jc}}(id_i, id_j, id_c, y, H_3)$ and verify that $H_4^* = H_4$ is equal. If equal, send the message $Mes_4 = (id_i, id_c, y, H_3)$ to S_i.

S_i		C
$< id_i, K_{ic} >$		$< id_i, id_c, K_{ic} >$
Generate a random nonce k. Compute $x = Enc_{K_{ic}}(id_i, k)$. $H_1 = h(id_i, k)$.		
	$Mes_1 = (id_i, x, H_1)$. →	
		Compute $Dec_{K_{ic}}(x) = \{k, h(k)\}$. $H_1^* = h(id_i, k)$. Verify if $H_1^* = H_1$. Generate a random nonce r. $SK = f(k, r, id_i, id_c, K_{ic})$. $y = Enc_{K_{ic}}(id_i, r)$. $H_2 = h(id_i, k, r)$.
	← $Mes_2 = (id_i, id_c, y, H_2)$.	
Compute $Dec_{K_{ic}}(y) = \{id_i, r\}$. $H_2^* = h(id_i, k, r)$. Verify if $H_2^* = H_2$. $SK = f(k, r, id_i, id_c, K_{ic})$.		

Figure 4: Two-party authentication and key agreement protocol for single link transmission

S_i	S_j	C
$< id_i, K_{ic} >$	$< id_j, K_{jc} >$	$< id_i, id_j, id_c, K_{ic}, K_{jc} >$
Generate a random nonce k. Compute $x = Enc_{K_{ic}}(id_i, k)$. $H_1 = h(id_i, k)$. $Mes_1 = (id_i, x, H_1)$. →		
	Compute $H_2 = MAC_{K_{jc}}(id_i, id_j, x, H_1)$. $Mes_2 = (id_i, id_j, x, H_1, H_2)$. →	
		Compute $H_2^* = MAC_{K_{jc}}(id_i, id_j, x, H_1)$. Verify if $H_2^* = H_2$. $Dec_{K_{ic}}(x) = \{id_i, k\}$. $H_1^* = h(id_i, k)$. Verify if $H_1^* = H_1$. Generate a random nonce r. $SK = f(k, r, id_i, id_c, K_{ic})$. $y = Enc_{K_{ic}}(id_i, r)$. $H_3 = h(id_i, k, r)$. $H_4 = MAC_{K_{ic}}(id_i, id_j, id_c, y, H_3)$. $Mes_3 = (id_i, id_j, id_c, y, H_3, H_4)$.
	$H_4^* = MAC_{K_{jc}}(id_i, id_j, id_c, y, H_3)$. Verify if $H_4^* = H_4$. $Mes_4 = (id_i, id_c, y, H_3)$. ←	
Compute $Dec_{K_{ic}}(y) = \{id_i, r\}$. $H_3^* = h(id_i, k, r)$. Verify if $H_3^* = H_3$. $SK = f(k, r, id_i, id_c, K_{ic})$.		

Figure 5: Three-party authentication and key agreement protocol for relay link transmission

Step 5: S_i after receiving the message Mes_4, decrypt $Dec_{K_{ic}}(y) = id_i, r$ and calculate $H_3^* = h(id_i, k, r)$, and then verify that $H_3^* = H_3$ is equal. If equal, calculate the temporary session key $SK = f(k, r, id_i, id_c, K_{ic})$.

4.3.3 Data Transmission Phase

After the sensor node S_i, the relay node S_j and the coordinator C complete authentication and temporary session key establishment, the sensor node S_i uses SK to encrypt the perceptual data M to obtain $E_{SK}(M, h(M))$, and then through the relay node S_j transmitted to coordinator C. Coordinator C decrypts the data to get M.

5 Security Analysis and Proof

Theorem 1. *The proposed scheme can provide mutual authentication.*

Proof. In our scheme, the coordinator C can authenticate the sensor node S_i and the relay node S_j by the preshared key K_{ic} and K_{jc}, respectively. At the same time, S_i and S_j can also authenticate C through K_{ic} and K_{jc}. Therefore, our solutions are able to provide mutual authentication. □

Theorem 2. *The proposed scheme can resist denial of service attacks.*

Proof. Denial of service attack is the most common type of attack on the network. This kind of attack utilizes the asymmetry of information exchange resources and consumes a large amount of the limited resources of the victim, thus undermining the network usability. For example, an attacker could repeatedly send a forged Mes_1 to C, and in the absence of any protection measures, C will think that this is the retransmission message Mes_1 from S_i. Therefore, C will continue to repeat the calculation of temporary session key SK, and store all the calculated SK and the corresponding random number r. But in our proposed scheme, denial of service attack is invalid. Since C receives the forged Mes_1, it does not generate and store the random number r and the temporary session key SK after verifying H_1 failure. Similarly, S_i is the same. Therefore, our scheme can resist denial of service attacks. □

Theorem 3. *The proposed scheme can resist man-in-the-middle attacks.*

Proof. Man-in-the-middle attack means that the attacker can intercept, replace or tamper with the information in the interaction process. In the proposed scheme, it is impossible for an attacker to arbitrarily forge and tamper with the information, because it can not obtain a pre-shared key between the sender and the receiver. For example, suppose an attacker S_k intercepts the interaction between S_i and C and replaces the authentication request (x, H_i) with (y, H_k). However, this man-in-the-middle attack is still unsuccessful because the attacker does not have a pre-shared key K_{ic} and can not produce a correct $y = Enc_{K_{ic}}(id_i, k)$. Therefore, the proposed scheme can resist man-in-the-middle attacks. □

Theorem 4. *The proposed scheme can resist replay attacks.*

Proof. Replay attack is the attacker intercepts the message before the communication process, and then replays the intercepted message in the future interactive communication process. The solution proposed in this paper can resist the attack because of the addition of random numbers k and r to ensure the freshness of the message. If an attacker replays the previous interactive message, the interaction will be stopped because the failure of the random number verification. In addition, except the sender, only the receiver can obtain the random number by the preshared key decryption, and the attacker does not have a preshared key can not get the random number. Therefore, the proposed scheme can resist replay attacks. □

6 Performance Analysis

In the simulation experiment, we use a commonly WBAN settings, as shown in Figure 1. The human body wears five sensor nodes that transmit the perceived data to coordinator C in real time and have a relay forwarding function. At the same time, the indoor walk as the default body movement. Correspondingly, reasonable time-varying model parameters can be determined based on the measurement results of the wireless body channel in [5, 10, 21], as shown in Table 1 and Table 2. In addition, the predetermined reception power threshold $\bar{P}^*$ is set to -85dBm, and the transmission power P_t is set to -10dBm, which is the recommended transmission power level of the medical special node. Therefore, the link interrupt threshold $\bar{G}^*$ is -75dB. In the simulation, we use the superframe structure of the time slot length and superframe length were set to 5ms and 250ms. At the same time, the same time correlation coefficient ρ_i is considered for the link between all sensor nodes and coordinator C, and Table 1 gives the time correlation coefficient within 500ms. In order to examine the reliability of single link communication, the outage probability of all direct links are calculated, as shown in Table 1 and Table 2. It can be seen from Table 1 that the outage probability of link $S_{RA} - C$ and link $S_{LW} - C$ exceeds 5%, which means that it is necessary to assign the relay nodes to the two links to ensure the reliability of communication.

In order to prove the effectiveness of the proposed relay transmission strategy, we compare the load balancing relay transmission strategy with random selection relay transmission [6], optimal selective relay transmission [7] and maximum effort relay transmission [12]. At the same time, we compare the performance of the proposed authentication and key agreement protocol with some typ-

ical authentication and key agreement protocols. Two-party authentication and key agreement protocols including Guiying protocol [11], Saeed protocol [22], Yi protocol [27] and Xie protocol [25], three-party authentication and key agreement protocols including Lv protocol [17], Yang protocol [26], He protocol [14] and Chang protocol [4].

In the simulation experiment, we use the outage probability and the lifecycle of relay node to test the proposed relay transmission strategy, which represent the network reliability and energy efficiency. The lifecycle of the relay node selects the lifetime of the first relay node as the lifetime of the network, which reflects the starting time of network performance deterioration. At the same time, we use the two metrics of calculation overhead and energy consumption to evaluate the proposed authentication and key agreement protocol. The initial energy of each sensor node is set to 1000mJ/s.

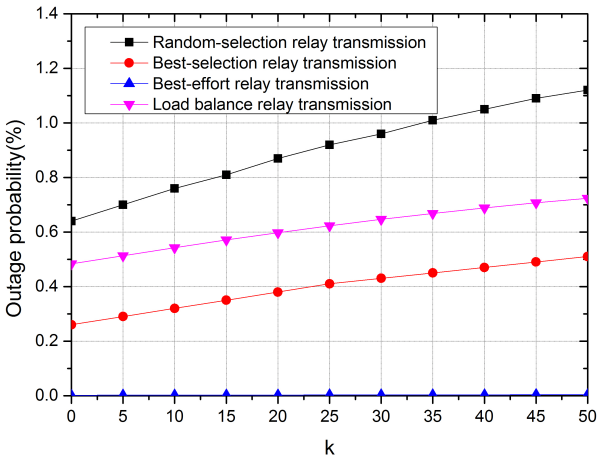


Figure 6: The change curve of the outage probability of link RA-C with the time slot

6.1 Outage Probability

Figure 6 and Figure 7 shows the relationship between the outage probability of the link $S_{RA}-C$ and $S_{LW}-C$ using the relay transmission strategy and the next transmission time slot. As can be seen from Figure 6 and Figure 7, the outage probability increases with the increase of the next transmission time slot. This is because the time correlation coefficient decreases with the transmission time slot increases. In addition, we can find that the four relay transmission strategies significantly reduce the outage probability of link $S_{RA}-C$ and link $S_{LW}-C$, and prove the efficiency of the relay transmission strategy. It can be seen from Figure 6 and Figure 7 that the outage probability of best-effort relay transmission strategy is the lowest, but the energy consumption of the relay node is the largest. Figure 7 shows that the outage probability of our scheme is higher than that of the other three schemes, but the outage probability of our scheme is no more than 1%, and the link $S_{LW}-C$ still has high transmission efficiency.

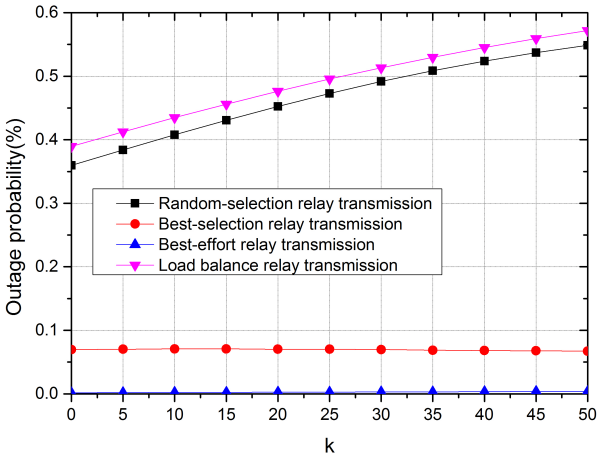


Figure 7: The change curve of the outage probability of link LW-C with the time slot k

6.2 Relay Node Energy Consumption

Figure 8 shows the relationship between the energy consumption of relay nodes and the lifecycle of four relay transmission strategies. From the figure we can see that the lifecycle of best-effort relay transmission strategy is the shortest, this is because the best-effort relay transmission strategy assigns all candidate nodes as relay nodes to the sensor node, thus greatly reducing the service life of the relay nodes. In addition, it can be seen from the figure that the lifecycle of load balancing relay transmission strategy is the longest, which means that the proposed relay transmission strategy is superior to the other three strategies in terms of energy efficiency. This is because in our relay transmission strategy, the coordinator to select the optimal relay node according to the residual energy of nodes, thus significantly improving the energy efficiency of relay transmission while ensuring the reliability of transmission.

6.3 Computation Overhead

Before simulating the running time of the protocol, the running time of the various algorithms used in the protocol is simulated on the single chip. The simulation environment is 512KB memory, clocked at 72MHz 32-bit Cortex-M3 single chip [24]. As shown in Table 3, we run the simulation time for each operation 100 times to get the average result.

In this paper, we use AES-128 algorithm for encryption and decryption, pseudo-random function using HMAC-SHA256 algorithm to calculate the temporary session key, hash function using SHA-256 algorithm, message authentication code using HSHA-256 algorithm, random number

Table 1: Single link parameters for time-varying model in indoor walking scenarios

Link	(μ_i, σ_i)	$Prob(\bar{G}_i < -75dB)$	$\rho_i(5), \rho_i(10), \cdots, \rho_i(100)$
$S_{RA} - C$	(-69.6,6.3)	6.59%	0.95,0.90,0.85,0.80,
$S_{RE} - C$	(-68.0,6.2)	4.74%	0.75,0.70,0.65,0.60,
$S_{LT} - C$	(-66.5,5.5)	2.12%	0.55,0.50,0.45,0.40,
$S_{LW} - C$	(-63.4,7.9)	5.21%	0.35,0.30,0.25,0.20,
$S_{RC} - C$	(-57.7,5.2)	0.11%	0.15,0.10,0.05,0.00.

Table 2: Relay link parameters for time-varying model in indoor walking scenarios

Source node	**Relay node**	(μ_{ij}, σ_{ij})	$Prob(\bar{G}_i < -75dB)$
S_{RA}	S_{RE}	(-64.4,7.6)	5.28%
	S_{LT}	(-59.7,7.1)	1.84%
	S_{RC}	(-71.2,6.2)	8.19%
S_{LW}	S_{RE}	(-68.6,7.8)	9.92%
	S_{LT}	(-65.4,7.1)	4.89%
	S_{RC}	(-59.7,6.6)	1.23%

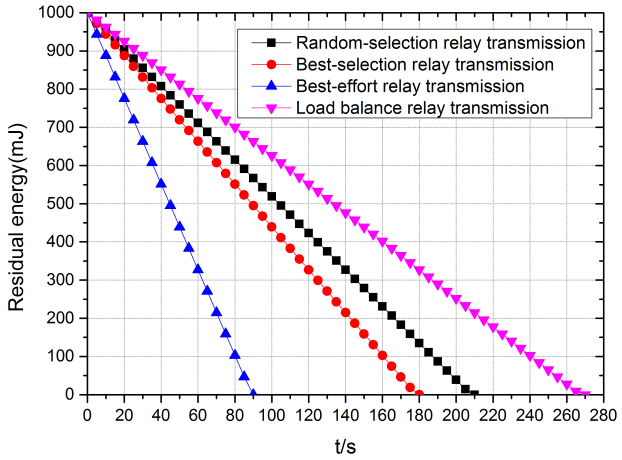


Figure 8: The change curve of the energy of relay node with time t

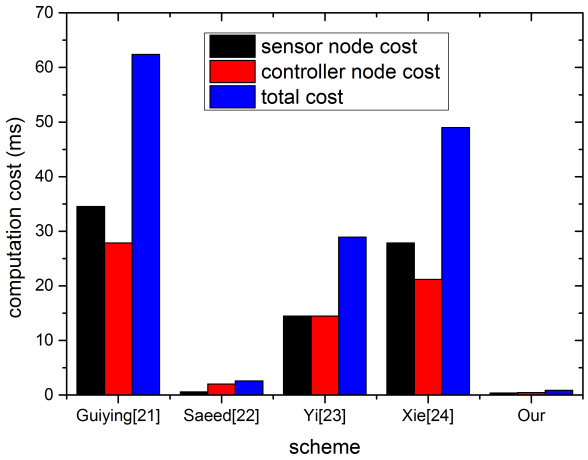


Figure 9: The computational cost of two-party authentication and key agreement protocol

generation contains three AES-128 encryption and two XOR operations. When simulating the running time of the protocol, the intermediate transmission time of the message is ignored, taking into account only the time at which it is calculated at both ends. In the process of single link transmission, we compare the proposed two-party authentication and key agreement protocol with some classical two-party authentication and key agreement protocol. The operation time of each two-party protocol is shown in Table 4, and the corresponding histogram result is shown in Figure 9. In the process of relay transmission, we compare the proposed three-party authentication and key agreement protocol with some classic three-party authentication and key agreement protocol. The operation time of each three-party protocol as shown in Table 5, the corresponding histogram results shown in Figure 10.

It can be seen from Figure 9 that the calculation overhead of Guiying protocol, Yi protocol and Xie protocol is relatively large. We propose that two-party authentication and key agreement protocol have the shortest running time in the five protocols, compared to other protocols is more superior. From Figure 10 we can see that the proposed three-party authentication and key agreement protocol is less time-consuming in this comparison of four protocols. In the other three schemes, the computation of relay nodes is relatively large, which greatly shortens the lifetime of nodes, and is not suitable for WBAN.

6.4 Energy Consumption

The energy consumption of encryption operation is used to evaluate our protocol. For 32-bit Cortex-M3 microcontroller with 72MHz, the current consumption of active mode is 36mA [29] at an ambient temperature

Table 3: Computational time

Notations	Operations	Computation time (ms)
T_{sym}	Symmetric en/decryption	0.031/0.067
T_{Asym}	Asymmetric en/decryption	0.146/1.584
T_{Hash}	One-way hash function	0.032
T_{Ran}	Random number	0.117
T_{HMAC}	Keyed message authentication code	0.043
T_{Pse}	Pseudorandom function	0.156
T_{Exp}	Modular exponentiation	5.542
T_{Bp}	Bilinear pairing	14.316
T_{Ecsm}	Elliptic curve scalar point multiplication	6.697

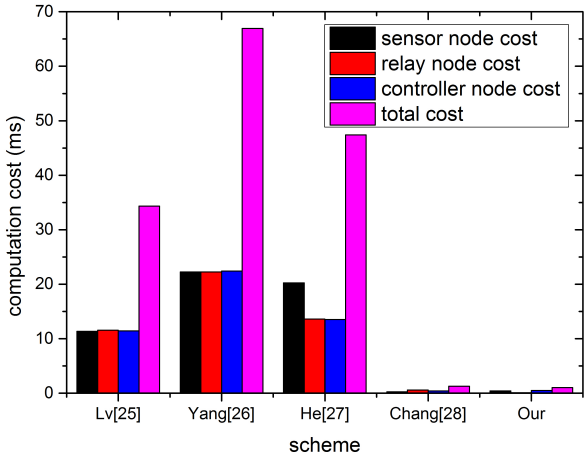


Figure 10: The computational cost of three-party authentication and key agreement protocol

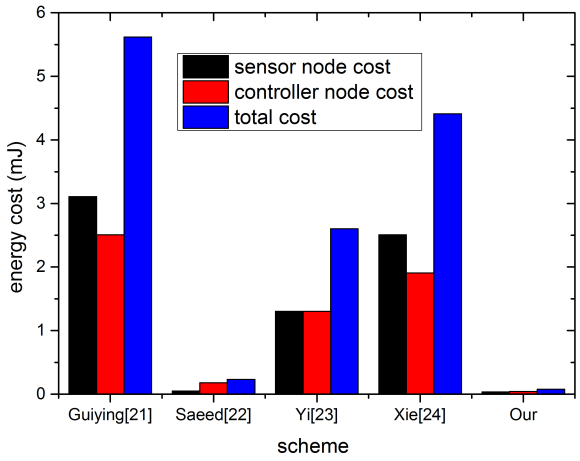


Figure 11: The energy consumption of two-party authentication and key agreement protocol

of 27, and the power consumption of active mode is approximately 90mW at a voltage of 2.2V. Therefore, according to Table 4 and Table 5 running time, we can calculate the corresponding energy loss. For example, a sensor node takes 0.031ms to complete the AES-126 encryption operation, the energy consumption is about $0.031ms \times 90/1000 = 0.003mJ$. The energy consumption of all schemes is shown in Figure 11 and Figure 12. From Figure 10 we can see that the total energy consumption of the proposed two-party authentication and key agreement protocol is the smallest, and the calculated energy consumption of the sensor node is also the smallest, and can meet the limited computing ability of WBAN demand. From Figure 12 we can see that the proposed three-party authentication and key agreement protocol of the sensor nodes, relay nodes and control node are the smallest energy consumption. In WBAN, the relay node needs to be used frequently, so the proposed scheme can meet its needs.

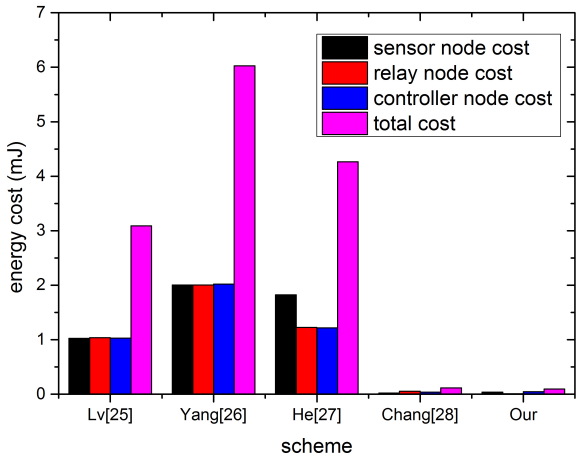


Figure 12: The energy consumption of three-party authentication and key agreement protocol

Table 4: Simulation results of two-party authentication and key agreement protocol

Protocols	Computation cost $[T_{Ran}, T_{Asym}, T_{Sym}, T_{Pse}, T_{Hash}, T_{HMAC}, T_{Bp}, T_{Ecsm}]$			
	Sensor node (ms)	Controller node (ms)	Tatal computation cost (ms)	Total energy cost (mJ)
Guiying [21]	[1,0,0,0,2,0,1,3]	[1,0,0,0,2,0,1,2]	62.413	5.617
Saeed [22]	[1,1,0,1,2,1,0,0]	[1,1,0,1,2,1,0,0]	2.577	0.232
Yi [23]	[1,0,0,0,1,0,1,0]	[1,0,0,0,1,0,1,0]	28.93	2.603
Xie [24]	[1,0,0,0,2,0,1,2]	[1,0,0,0,2,0,1,1]	49.019	4.412
Our	[1,0,2,1,2,0,0,0]	[1,0,2,1,2,0,0,0]	0.87	0.078

Table 5: Simulation results of three-party authentication and key agreement protocol

Protocols	Computation cost $[T_{Ran}, T_{Asym}, T_{Sym}, T_{Pse}, T_{Hash}, T_{HMAC}, T_{Exp}, T_{Ecsm}]$				
	Sensor node (ms)	Relay node (ms)	Controller node (ms)	Tatal computation cost (ms)	Total energy cost (mJ)
Lv [25]	[1,0,3,0,1,0,2,0]	[2,0,4,0,1,0,2,0]	[1,0,3,0,2,0,2,0]	34.338	3.090
Yang [26]	[0,0,0,0,3,0,4,0]	[0,0,0,0,3,0,4,0]	[0,0,0,0,8,0,4,0]	66.952	6.026
He [27]	[0,0,2,0,2,0,0,3]	[0,0,4,0,1,0,0,1]	[0,0,2,0,1,0,0,2]	47.399	4.266
Chang [28]	[1,0,0,0,4,0,0,0]	[2,0,0,0,11,0,0,0]	[1,0,0,0,10,0,0,0]	1.268	0.114
Our	[1,0,2,1,2,0,0,0]	[0,0,0,0,0,2,0,0,]	[1,0,2,1,2,2,0,0]	1.042	0.094

7 Conclusion

In this paper, a new security and reliability scheme is proposed based on the channel characteristics of WBAN. Through the use of time slot allocation and load balancing relay transmission strategy to realize the reliability transmission of data. Then, in the process of data transmission, a new authentication and key agreement protocol is proposed for single-link transmission and relay link transmission mode respectively, which ensures the security transmission of data. Through the security analysis, we prove that the proposed scheme meets the high security level requirements of communication. The simulation results show that our transmission strategy can improve the reliability of data transmission with low computational cost and energy consumption.

Acknowledgments

Above work is supported by National Natural Science Foundation (NSF) of China under grant Nos. 61370007, 61572206, U1405254, Huaqiao University graduate research innovation ability cultivation project of China under grant No. 1511314006, Fujian Provincial Natural Science Foundation of China under grant No. 2013J01241, and Program for New Century Excellent Talents of Fujian Provincial under grant No. 2014FJ-NCET-ZR06.

References

[1] U. F. Abbasi, A. Awang, and N. H. Hamid, "Performance investigation of using direct transmission and opportunistic routing in wireless body area networks," in *IEEE Symposium on Computers and Informatics (ISCI'13)*, pp. 60–65, Apr. 2013.

[2] A. Boulis, D. Smith, D. Miniutti, L. Libman, and Y. Tselishchev, "Challenges in body area networks for healthcare: The mac," *IEEE Communications Magazine*, vol. 50, no. 5, pp. 100–106, 2012.

[3] R. Cavallari, F. Martelli, R. Rosini, C. Buratti, and R. Verdone, "A survey on wireless body area networks: Technologies and design challenges," *IEEE Communications Surveys and Tutorials*, vol. 16, no. 3, pp. 1635–1657, 2014.

[4] C. C. Chang, W. Y. Hsueh, and T. F. Cheng, "A dynamic user authentication and key agreement scheme for heterogeneous wireless sensor networks," *Wireless Personal Communications*, vol. 89, no. 2, pp. 447–465, 2016.

[5] R. D'Errico and L. Ouvry, "Time-variant ban channel characterization," in *IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications*, pp. 3000–3004, Sep. 2009.

[6] R. D'Errico, R. Rosini, and M. A. Maman, "performance evaluation of cooperative schemes for on-body area networks based on measured time-variant channels," in *IEEE International Conference on Communications (ICC'11)*, pp. 1–5, June 2011.

[7] H. Feng, B. Liu, Z. Yan, C. Zhang, and C. W. Chen, "Prediction-based dynamic relay transmission scheme for wireless body area networks," in *IEEE 24th International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC'13)*, pp. 2539–2544, Sep. 2013.

[8] E. Forrister, G. Lee, D. Xue, B. Garner, and Y. Li, "Characterization of narrowband on-body wireless channels using motion capture experimentation," in

Texas Symposium on Wiireless and Microwave Circuits and Systems (WMCS'16), pp. 1–4, Mar. 2016.

[9] L. C. Fourati N. Jamali, "Skep: A secret key exchange protocol using physiological signals in wireless body area networks," in *International Conference on Wireless Networks and Mobile Communications (WINCOM'15)*, pp. 1–7, Oct. 2015.

[10] J. M. Gorce, C. Goursaud, G. Villemaud, R. D'Errico, and L. Ouvry, "Opportunistic relaying protocols for human monitoring in ban," in *IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications*, pp. 732–736, Sep. 2009.

[11] L. Guiying, H. Mengbo, Z. Chuan, and X. Qiuliang, "A two-party certificateless authenticated key agreement protocol with provable security," in *9th International Conference on Computational Intelligence and Security (CIS'13)*, pp. 559–563, Dec. 2013.

[12] S. Hara, D. Anzai, K. Yanagihara, K. Takizawa, and K. Hamaguchi, "A cooperative transmission scheme for real-time data gathering in a wireless body area network," in *IEEE 22nd International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC'11)*, pp. 2254–2258, Sep. 2011.

[13] D. He, S. Chan, Y. Zhang, and H. Yang, "Lightweight and confidential data discovery and dissemination for wireless body area networks," *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 2, pp. 440–448, 2014.

[14] D. He and S. Zeadally, "Authentication protocol for an ambient assisted living system," *IEEE Communications Magazine*, vol. 53, no. 1, pp. 71–77, 2015.

[15] B. E. Ieee, "Ieee standard for local and metropolitan area networks - part 15.6: Wireless body area networks," *IEEE Standard for Information Technology*, vol. 802, no. 6, pp. 1–271, 2012.

[16] J. Iqbal, N. ul Amin, A. I. Umar, N. Din, and A. Waheed, "Secure lightweight authentication and key agreement for wireless body area networks," *International Journal of Computer Science and Information Security*, vol. 14, no. 5, p. 196, 2016.

[17] C. Lv, M. Ma, H. Li, J. Ma, and Y. Zhang, "An novel three-party authenticated key exchange protocol using one-time key," *Journal of Network and Computer Applications*, vol. 36, no. 1, pp. 498–503, 2013.

[18] M. Maman and L. Ouvry, "Batmac: An adaptive tdma mac for body area networks performed with a space-time dependent channel model," in *5th International Symposium on Medical Information and Communication Technology (ISMICT'11)*, pp. 1–5, Mar. 2011.

[19] S. Movassaghi, M. Abolhasan, J. Lipman, D. Smith, and A. Jamalipour, "Wireless body area networks: A survey," *IEEE Communications Surveys and Tutorials*, vol. 16, no. 3, pp. 1658–1686, 2014.

[20] S. C. Mukhopadhyay, "Wearable sensors for human activity monitoring: A review," *IEEE Sensors Journal*, vol. 15, no. 3, pp. 1321–1330, 2015.

[21] H. Ren, M. Meng, and C. Cheung, "Experimental evaluation of on-body transmission characteristics for wireless biosensors," in *IEEE International Conference on Integration Technology (ICIT'07)*, pp. 745–750, Mar. 2007.

[22] M. Saeed, H. S. Shahhoseini, A. Mackvandi, M. R. Rezaeinezhad, M. Naddafiun, and M. Z. Bidoki, "A secure two-party password-authenticated key exchange protocol," in *IEEE 15th International Conference on Information Reuse and Integration (IRI'14)*, pp. 466–474, Aug. 2014.

[23] D. B. Smith, D. Miniutti, and L. W. Hanlen, "Characterization of the body-area propagation channel for monitoring a subject sleeping," *IEEE Transactions on Antennas and Propagation*, vol. 59, no. 11, pp. 4388–4392, 2011.

[24] S. Wang, Z. H. Wu, P. Hu, and Z. LI, "Design and implement of bootloader based on pxa270 processor," *Journal-Sichuan university natural science edition*, vol. 44, no. 3, p. 578, 2007.

[25] Y. Xie, L. Wu, Y. Zhang, and Z. Xu, "Strongly secure two-party certificateless key agreement protocol with short message," in *International Conference on Provable Security*, pp. 244–254, Nov. 2016.

[26] H. Yang, Y. Zhang, Y. Zhou, X. Fu, H. Liu, and A. V. Vasilakos, "Provably secure three-party authenticated key agreement protocol using smart cards," *Computer Networks*, vol. 58, pp. 29–33, 2014.

[27] T. Yi, M. Shi, and W. Shang, "Personalized two party key exchange protocol," in *IEEE/ACIS 14th International Conference on Computer and Information Science (ICIS'15)*, pp. 575–579, June 2015.

[28] X. Yi, A. Bouguettaya, D. Georgakopoulos, A. Song, and J. Willemson, "Privacy protection for wireless medical sensor data," *IEEE Transactions on Dependable and Secure Computing*, vol. 13, no. 3, pp. 369–380, 2016.

[29] D. Zhao, Y. Wang, J. Tan, "Design and experiment on a biomimetic robotic fish inspired by freshwater stingray," *Journal of Bionic Engineering*, vol. 12, no. 2, pp. 204–216, 2015.

[30] J. Zhou, Z. Cao, X. Dong, N. Xiong, and A. V. Vasilakos, "4s: A secure and privacy-preserving key management scheme for cloud-assisted wireless body area network in m-healthcare social networks," *Information Sciences*, vol. 314, pp. 255–276, 2015.

Biography

Huaijin Liu received the B.S. degree from Huaqiao University, China, in 2015, where he is currently pursuing the master's degree. His current research interest includes wireless sensor network security, wireless body area network security and privacy protection, wireless vehicle network security.

Yonghong Chen received the B.S. degrees from Hubei National University, and M.Eng. and Ph.D. degree degrees from Chognqing University, Chongqing, China, in 2000 and 2005 respectively. He is currently the professor of of College of Computer Science and Technology, Huaqiao University, Xiamen, China. His research interests include network security, watermarking and nonlinear processing.

Hui Tian received his BSc and MSc degrees in Wuhan Institute of Technology,Wuhan, China in 2004 and 2007, respectively. He received his PhD degree in Huazhong University of Science and Technology, Wuhan, China. He is now an associate professor in the National Huaqiao University of China. His research interests include network and multimedia information security, digital forensics and information hiding.

Tian Wang received his BSc and MSc degrees in Computer Science from the Central South University in 2004 and 2007, respectively. He received his PhD degree in City University of Hong Kong in 2011. Currently, he is a professor in the Huaqiao University of China. His research interests include wireless sensor networks, fog computing and mobile computing.

Yiqiao Cai received the B.S. degree from Hunan University, Changsha, China, in 2007, and the Ph.D. degree from Sun Yat-sen University, Guangzhou, China, in 2012. In 2012, he joined Huaqiao University, Xiamen, China, where he is currently a lecturer with the College of Computer Science and Technology. He is interested in differential evolution, multiobjective optimization, and other evolutionary computation techniques.

Security Analysis and Enhancements of A Remote User Authentication Scheme

Shou-Qi Cao, Qing Sun, and Li-Ling Cao
(Corresponding author: Li-Ling Cao)

College of Engineering Science and Technology, Shanghai Ocean University
Shanghai 201306, China
(Email: llcao@shou.edu.cn)

(Received June 6, 2018; Revised and Accepted Sept. 22, 2018; First Online Jan. 23, 2019)

Abstract

Many remote user authentication schemes have been designed and developed to establish secure and authorized communication between the users and the sever over an insecure channel. By employing a secure remote user authentication scheme, the users and the server can authenticate each other and utilize advanced services. In 2012, Hsieh and Leu proposed a remote user authentication scheme. However, we review and analyze Hsieh and Leu's scheme and find that their scheme can't provide user anonymity and is vulnerable to slow wrong password detection, masquerading attack and password guessing attack. In order to solve these drawbacks, we propose a security-improved authentication scheme which can resist all attacks above. Finally, security formal analysis of the proposed scheme using Burrows-Abadi-Needham logic (BAN-logic) is given, which indicates that the proposed scheme can protect against several possible types of attacks with only a slightly high computational cost.

Keywords: Authentication; Hash Function; Masquerading Attack; User Anonymity

1 Introduction

With the rapid development of the internet, computer applications have penetrated into all areas of society, which provide us a lot of services and bring conveniences to our life and work. However, it also brings many negative effects such as the fortified ability of attackers via the internet. At the same time, information security has become an important issue. In order to protect the security of information, many scholars have proposed some authentication schemes to establish secure and authorized communication between the users and the server. After Lamport [13] first proposed a password-based user authentication scheme, many researchers proposed password-based user authentication schemes with key agreement [7,16,20]. However, Lennon et al. [15] and Yen and Liao [25] demonstrated that Lamport's scheme was vulnerable to stolen verifier attacks. Typically, in one-factor authentication schemes, the server maintains a table containing the verifiers of the users [8]. Therefore, the servers tend to be easy objects of attack, because if an adversary achieves the verifier of a user that is stored in the verification table, then he/she can masquerade as the victim user [2, 6, 11].

In order to overcome these problems, some schemes based on smart card which also called two-factor authentication schemes have emerged. In 1991, Chang and Wu [4] developed the first smart-card-based password authentication scheme. Then, many improvements were made to enhance its security and efficiency [12,14,17,24]. In 2004, Yoon et al. [26] proposed a scheme which enabled users to change passwords freely and securely without the help of a remote server, while also providing secure mutual authentication. However, Hsiang and Shih [9] found that it can't resist parallel session attack and masquerading attack and password guessing attack. Therefore, Hsiang and Shih proposed their own scheme, but He et al. [5] pointed out that it was still vulnerable to password guessing attack, masquerading attack. Besides, Hsieh and Leu [10] found that an insider can carry out an infringed account attack and a resembling account attack on Hsiang et al.'s solution. However, Wang et al. [22] showed that, under their non-tamper- resistance assumption of the smart cards, Hsieh and Leu's scheme was still prone to offline dictionary attack, in which an attacker could obtain the victim's password when getting temporary access to the victim's smart cards. Wang et al. didn't put forward improved scheme. In this paper, we find that Hsieh and Leu's scheme is still exposed to masquerading attack, password guessing attack and can't provide user anonymity, mutual authentication, fast password detection. Therefore, we propose our improved scheme that can fight against all aforementioned attacks.

The rest of the paper is organized as follows. The review and the analysis of Hsieh and Leu's scheme are presented in Section 2 and 3. In Section 4, we propose a scheme that can resist all attacks mentioned in related researches. Section 5 devotes to making security formal

analysis based on Burrows-Abadi-Needham logic (BAN-logic) and comparing the proposed scheme with some existing ones. The result indicates that our modified scheme has a slightly high computational cost and can protect against some possible attacks. Finally, we conclude this paper in Section 6.

2 Review of Hsieh and Leu's Scheme

Hsieh and Leu's scheme is a remote user authentication scheme which uses hash functions. It contains four phases: registration, login, authentication and password change.

2.1 Registration Phase

In this phase, the initial registration is different from the re-registration. The process of the initial registration is depicted as follows.

R1. User U chooses a random number b and computes $h(b \oplus PW_u)$.

R2. U sends the message ID_u, $h(PW_u)$ and $h(b \oplus PW_u)$ to S.

R3. In the account database, the server S creates an entry for U and stores $n = 0$ in this entry.

R4. S performs the following computations: $P = h(EID \oplus x)$, $EID = h(h(ID_u)||n)$, $R = P \oplus h(b \oplus PW_u)$, $V = h(h(PW_u) \oplus h(x))$ which is stored in the entry of U.

R5. S gives a smart card to U containing R and $h(\cdot)$.

R6. When receiving the smart card issued by the server S, U inputs b into his smart card. Finally, the smart card contains b, R and $h(\cdot)$. After this phase, U does not need to remember b.

If U misses her/his smart card and wants to re-register to S, the process of the re-registration is as the follows.

RR1. User U chooses a new random number b' and computes $h(b' \oplus PW_u)$.

RR2. U sends the message ID_u, $h(PW_u)$, $h(b' \oplus PW_u)$ to S.

RR3. S computes $V' = h(h(PW_u) \oplus h(x))$ and compares V with V'.

RR4. If V' is equal to V, S sets $n = n+1$ in the existing entry. Then S performs the following computations: $P_{new} = h(EID \oplus x)$, $EID = h(h(ID_u)||n)$, $R_{new} = P \oplus h(b' \oplus PW_u)$.

RR5. Finally, S performs Steps R5 and R6 shown in the initial registration process.

2.2 Login Phase

When U wants to login to the remote server S, the following operations will be performed.

L1. U inserts his smart card into the smart card reader and enters his ID_u and PW_u.

L2. The following computations are performed by U's smart card: $C_1 = R \oplus h(b \oplus PW_u)$, $C_2 = h(C_1 \oplus T_u)$, where T_u denotes U's current timestamp .

L3. U sends $C = \{ID_u, T_u, C_2\}$ to S.

2.3 Authentication Phase

After receiving the login request message $C = \{ID_u, T_u, C_2\}$, the remote server S and U's smart card perform the following operations.

A1. If either ID_u or T_u is invalid or $T_s - T_u \leq 0$, S rejects U's login request. Otherwise, S computes $C_2' = h(h(EID \oplus x) \oplus T_u)$ and compares C_2' with the received C_2. If $C_2' = C_2$, S accepts U's login request and computes $C_3 = h(h(EID \oplus x) \oplus h(T_s))$, where T_s is S's current timestamp. Otherwise, S rejects U's login request.

A2. S sends T_s and C_3 to U.

A3. If either T_s is invalid or $T_s = T_u$, this session will be terminated by U. Otherwise, U computes C_3' and compares C_3' with the received C_3, $C_3' = h(C_1 \oplus h(T_s))$. If C_3' is equal to C_3, U authenticates S successfully.

2.4 Password Change Phase

When U wants to change his password, the following process will be performed.

P1. U inserts his smart card into the smart card reader and enters ID_u, PW_u and new password PW_{new}.

P2. U sends the message ID_u, $h(PW_u)$, $h(PW_u) \oplus h(PW_{new})$, $h(b \oplus PW_{new})$ to S.

P3. S computes $V' = h(h(PW_u) \oplus h(x))$ and compares V' with V in the account database.

P4. If V' is equal to V, then S computes $h(PW_u) \oplus h(PW_{new}) \oplus h(PW_u)$ to get $h(PW_{new})$. Next, S performs the following computations.

$$\begin{aligned} P &= h(EID \oplus x) \\ EID &= h(h(ID_u)||n) \\ R_{new} &= P \oplus h(b \oplus PW_{new}) \\ V_{new} &= h(h(PW_{new}) \oplus h(x)) \end{aligned}$$

which is stored in U's entry.

P5. S sends R to U.

P6. Finally, U's smart card replaces R with R_{new}.

3 Cryptanalysis of Hsieh and Leu's Scheme

In this section, we analyze the security of Hsieh and Leu's scheme on the basis of the following assumptions:

1) An attacker can eavesdrop, intercept, and modify any message in the channel.

2) An attacker may either (1) obtain a user's password or (2) extract the secret information of the smart card, but can't achieve both (1) and (2) at the same time.

3.1 User Anonymity

Whenever a legal user U sends a login request message, the login request message contains the identity ID_u of the user. Therefore, Hsieh and Leu's scheme can't protect the anonymity of the user.

3.2 Slow Wrong Password Detection

Slow wrong password detection refers to instances in which the user can't know of a mistake immediately when inputs the wrong password, and the user can know when server S notifies there is a wrong user password. In Hsieh and Leu's authentication scheme, the user's smart card can't verify the accuracy of the user password during the login phase. Only S verifies a legal user by comparing the similarities between C_2' and C_2 during authentication phase. Concretely, U inputs ID_u and PW_u, then if U selects a wrong password PW_u^*, the smart card is unaware that the password is incorrect. The smart card does not check the PW_u^* and it computes various values $\{C_1^*, C_2^*\}$ using PW_u^* for login and authentication. The smart card then sends $\{ID_u, T_u, C_2^*\}$.

S is unable to immediately confirm the wrong password after receiving the message $\{ID_u, T_u, C_2^*\}$. First, S checks the validity of ID_u and T_u, then computes $C_2' = h(h(EID \oplus x) \oplus T_u)$. Then, because C_2^* is not the same as C_2', S eventually confirms that the received messages are not normal, and maybe U could have input the wrong password. Finally, S sends the wrong password notification to U. Hsieh and Leu's authentication scheme requires a lengthy phase that includes value computation and message transmission before confirming that the user input the wrong password. Therefore, a smart card needs a fast wrong password detection technique during login. When U inputs the wrong password during the login phase, the smart card needs to quickly identify the incorrect password and should immediately notify U of the mistake.

3.3 User Masquerading Attack & Replay Attack

When an attacker steals the smart card and intercepts the login request message $\{ID_u, T_u, C_2\}$ from U, he may send the replaying message $\{ID_u, T_u, C_2\}$ to S in a new session during authentication phase. Then S will compute C_2' and find C_2' is the same as C_2. Then S regards the attacker as legal user and accepts the login request.

3.4 Password Guessing Attack

Hsieh and Leu pointed out that Hsiang and Shih's scheme could not resist offline password attack. However, we find that Hsieh and Leu's scheme also fails to solve the problem. Then the attacker can guess the password in the following two conditions.

1) An attacker can know the information $\{R, h(\cdot), b\}$ stored in a smart card.

2) An attacker can intercept the login request message $\{ID_u, T_u, C_2\}$ over the communication channel.

The specific steps are as follows:

1) An attacker selects a password PW_u^*.

2) Computes $C_1^* = R \oplus h(b \oplus PW_u^*)$ and $C_2^* = h(C_1^* \oplus T_u)$.

3) Verifies the correctness of PW_u^* by checking if the computed C_2^* is equal to the intercepted C_2.

4) Repeats 1) $\sim$ 3) of this procedure until the correct value of PW_u is found.

Once the smart card is stolen or picked up, the corresponding password factor can be guessed. So Hsieh and Leu's scheme is not a two-factor scheme and is insecure.

3.5 Mutual Authentication

Generally, if authentication scheme is secure, it can resist user impersonation attack and server masquerading attack. However, the authentication scheme can't resist user impersonation attack as described in Section 4.3. Therefore, Hsieh and Leu's scheme fails to provide mutual authentication.

4 Proposed Scheme

In this section, we propose an enhanced scheme based on Hsieh and Leu's scheme which can resist all the attacks mentioned in Section 3. The enhanced scheme contains four phases: registration, login, authentication, password change phase.

4.1 Registration Phase

The registration phase of the proposed scheme is shown in Figure 1.

1) Initial registration.

R1, R2, R3. The same as Hsieh and Leu's scheme.

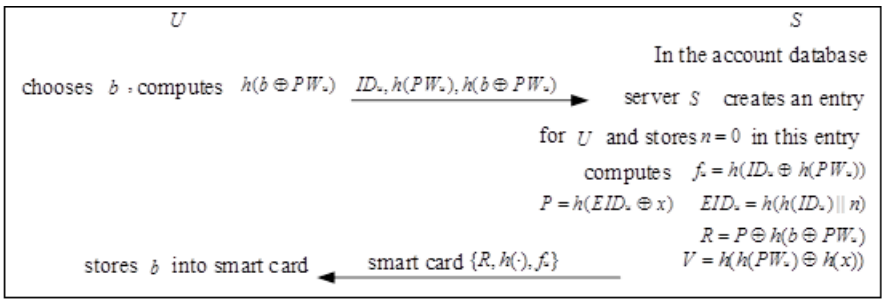


Figure 1: Registration phase

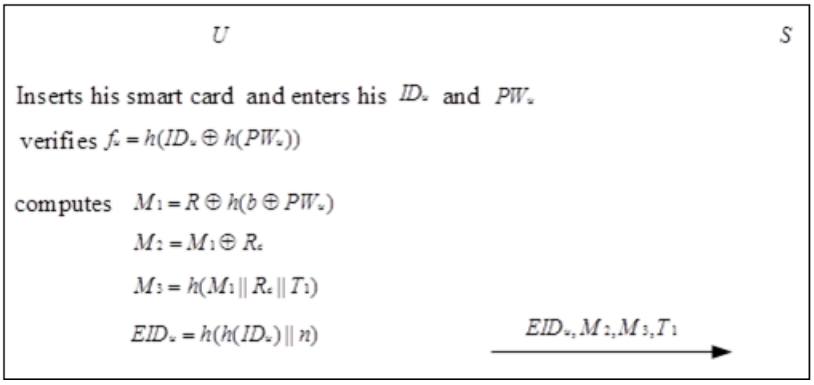


Figure 2: Login phase

R4. S computes f_u, P, R, V as follows.

$$\begin{aligned} f_u &= h(ID_u \oplus h(PW_u)) \\ P &= h(EID_u \oplus x) \\ EID_u &= h(h(ID_u)||n) \\ R &= P \oplus h(b \oplus PW_u) \\ V &= h(h(PW_u) \oplus h(x)). \end{aligned}$$

R5. S sends U a smart card containing f_u, R, $h(\cdot)$. Then U stores b in the smart card.

2) Re-registration.
This phase is the same as Hsieh and Leu's scheme.

4.2 Login Phase

The user U should execute the following steps when he logins to the remote server S (See Figure 2).

L1. The same as Hsieh and Leu's scheme.

L2. The smart card computes f_u and compares the computed f_u with the stored f_u.

L3. If they are the same, U generates the current timestamp T_1 and a random number R. Then U computes M_1, M_2, M_3, EID_u as follows:

$$\begin{aligned} M_1 &= R \oplus h(b \oplus PW_u) \\ M_2 &= M_1 \oplus R_c \\ M_3 &= h(M_1||R_c||T_1) \\ EID_u &= h(h(ID_u)||n). \end{aligned}$$

L4. U sends the login request message $\{EID_u, M_2, M_3, T_1\}$ to S.

4.3 Authentication Phase

After receiving the login request message $\{EID_u, M_2, M_3, T_1\}$, the remote server S and U's smart card perform the following operations and in Figure 3.

A1. S checks whether EID_u is the same as the EID_u stored in the database.

A2. If they are the same, S computes M_4 and M_5 as follows.

$$\begin{aligned} M_4 &= h(EID_u \oplus x) \\ M_5 &= M_2 \oplus M_4. \end{aligned}$$

A3. S compares the M_3 with $h(M_4||M_5||T_1)$. If they are equal, S computes M_6 and M_7.

$$\begin{aligned} M_6 &= M_4 \oplus R_5 \\ M_7 &= h(M_4||R_s||T_2), \end{aligned}$$

where T_2 and R_s respectively denotes S's current timestamp and random number. Then S sends $\{EID_u, M_6, M_7, T_2\}$ to U.

A4. U computes $M_8 = M_6 \oplus M_1$ and verifies whether $M_7 = h(M_1||M_8||T_2)$ or not. If they are the same, U

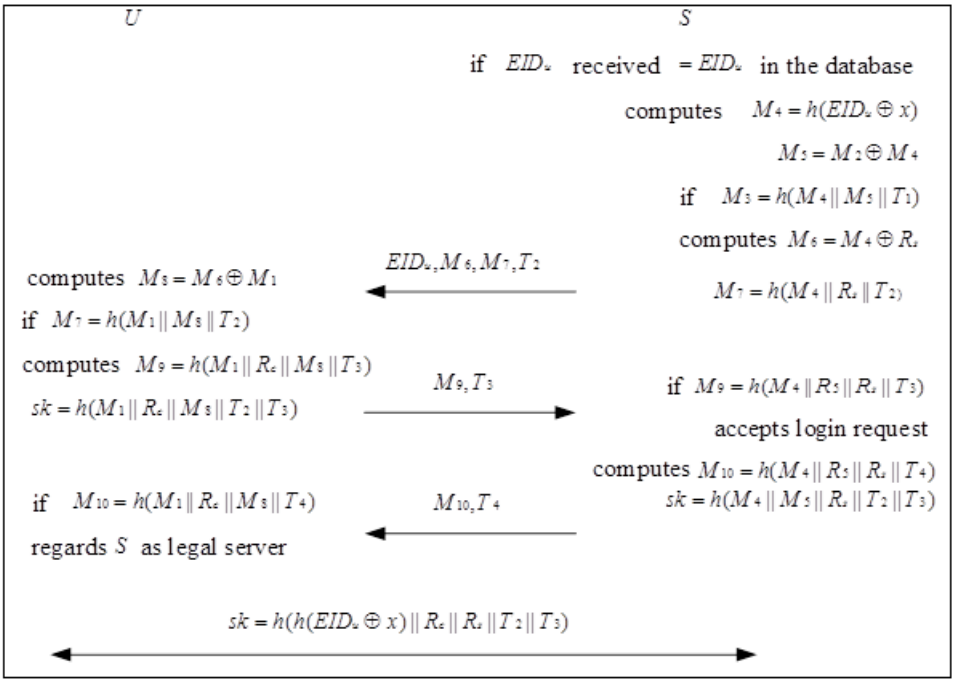


Figure 3: Authentication phase

computes M_9 and sk as follows.

$$\begin{aligned} M_9 &= h(M_1||R_c||M_8||T_3) \\ sk &= h(M_1||R_c||M_8||T_2||T_3) \end{aligned}$$

where T_3 is a timestamp.

A5. U sends $\{M_9, T_3\}$ to S.

A6. S receives the message and starts to verify whether the M_9 is equal to $h(M_4||M_5||R_s||T_3)$. If they are equal, S accepts the login request. Then S computes M_{10} and sk as follows.

$$\begin{aligned} M_{10} &= h(M_4||M_5||R_s||T_4) \\ sk &= h(M_4||M_5||R_s||T_2||T_3). \end{aligned}$$

A7. S sends $\{M_{10}, T_4\}$ to U.

A8. U receives the message and starts to verify whether the M_{10} is equal to $h(M_1||R_c||M_8||T_4)$. If they are equal, U regards S as legal server.

A9. Finally, they share a same session key $sk = h(h(EID_u \oplus x)||R_c||R_s||T_2||T_3)$.

4.4 Password Change Phase

For the proposed scheme, the password change phase is executed when U loses the smart card or wants to update the password.

P1. U selects and inputs ID_u, PW_u, $PW_{u_{new}}$ and generates a new random number b'. Then U submits ID_u, $h(PW_u)$, $h(b' \oplus PW_u)$, $h(PW_{u_{new}})$, $h(b \oplus PW_{u_{new}})$ to S through a secure channel.

P2. After S receives the message, S checks the database for the ID_u and computes $V' = h(h(PW_u) \oplus h(x))$ and compares it with V in the database.

P3. If V' is equal to V, then S carries out the computations as follows:

$$\begin{aligned} f_{u_{new}} &= h(ID_u \oplus h(PW_{u_{new}})) \\ P &= h(EID_u \oplus x) \\ EID_u &= h(h(ID_u)||n) \\ R_{new} &= P \oplus h(b' \oplus PW_{u_{new}}) \\ V_{new} &= h(h(PW_{u_{new}}) \oplus h(x)). \end{aligned}$$

P4. S sends a new smart card to U that contains $f_{u_{new}}$, R_{new}, $h(\cdot)$. Then U stores a new b' in the smart card.

5 Analysis of the Proposed Scheme

In this section, we first analyze the security of our proposed authentication scheme based on the assumptions stated in Section 3. Then, we show that the proposed scheme withstands all attacks mentioned in Hsieh and Leu's scheme.

5.1 Security Analysis Using BAN Logic

Burrows [3] proposed BAN logic in 1990. Although there are some controversial publications about BAN-logic [1, 18, 19, 23], it is the first formal analysis. As a method of analyzing authentication schemes, its simplicity and

intuitiveness have attracted the attention of scholars. The analysis of an authentication scheme using the BAN-logic tool consists of four steps [21] and the formal analysis of the security of the proposed scheme is described as follows.

Step 1. The goals of mutual authentication in the proposed scheme are shown as follows:

$$G1: U \models U \overset{sk}{\longleftrightarrow} S$$
$$G2: S \models U \overset{sk}{\longleftrightarrow} S$$
$$G3: U \models S \models U \overset{sk}{\longleftrightarrow} S$$
$$G4: S \models U \models U \overset{sk}{\longleftrightarrow} S.$$

Step 2. The idealization forms of the messages in the proposed scheme are shown as follows:

Message 1.

$$U \rightarrow S: \langle R_c \rangle_{h(EID_u \oplus x)}, (R_c, T_1)_{h(EID_u \oplus x)}, T_1.$$

Message 2.

$$S \rightarrow U: \langle R_s \rangle_{h(EID_u \oplus x)}, (R_s, T_2)_{h(EID_u \oplus x)}, T_2.$$

Message 3.

$$U \rightarrow S: (R_c, R_s, T_3)_{h(EID_u \oplus x)}, U \overset{sk}{\longleftrightarrow} S, T_3.$$

Message 4.

$$S \rightarrow U: (R_c, R_s, T_4)_{h(EID_u \oplus x)}, U \overset{sk}{\longleftrightarrow} S, T_4.$$

Step 3. The initial states of the proposed scheme can be assumed as follows:

$$P1: U \models \#(T_1)$$
$$P2: U \models \#(T_2)$$
$$P3: U \models \#(T_3)$$
$$P4: U \models \#(T_4)$$
$$P5: U \models U \overset{h(EID_u \oplus x)}{\longleftrightarrow} S$$
$$P6: S \models U \overset{h(EID_u \oplus x)}{\longleftrightarrow} S$$
$$P7: U \models S \Rightarrow U \overset{sk}{\longleftrightarrow} S$$
$$P8: S \models U \Rightarrow U \overset{sk}{\longleftrightarrow} S.$$

Step 4. According to the initial state assumptions and BAN-logic inference rules, the main analysis of the proposed scheme is stated as follows:

According to Message 3, we can get A1:

$$S \lhd \{(R_c, R_s, T_3)_{h(EID_u \oplus x)}, T_3, U \overset{sk}{\longleftrightarrow} S\}.$$

According to the assumption P6 and the message meaning rule, we can get A2:

$$S \models U \vdash \{(R_c, R_s, T_3)_{h(EID_u \oplus x)}, T_3, U \overset{sk}{\longleftrightarrow} S\}.$$

According to P3 and the freshness conjunction rule, we can get A3:

$$S \models \#\{(R_c, R_s, T_3)_{h(EID_u \oplus x)}, T_3, U \overset{sk}{\longleftrightarrow} S\}.$$

According to A2, A3 and the nonce verification, we can get A4:

$$S \models U \models \{(R_c, R_s, T_3)_{h(EID_u \oplus x)}, T_3, U \overset{sk}{\longleftrightarrow} S\}.$$

According to A4, we apply the belief rule, we can get A5:

$$G4: S \models U \models U \overset{sk}{\longleftrightarrow} S.$$

According to P8, A5 and the jurisdiction rule, we can get A6:

$$G2: S \models U \overset{sk}{\longleftrightarrow} S.$$

According to Message 3, we can get A7:

$$U \lhd \{(R_c, R_s, T_4)_{h(EID_u \oplus x)}, T_4, U \overset{sk}{\longleftrightarrow} S\}.$$

According to the assumption P5 and the message meaning rule, we can get A8:

$$U \lhd S \vdash \{(R_c, R_s, T_4)_{h(EID_u \oplus x)}, T_4, U \overset{sk}{\longleftrightarrow} S\}.$$

According to P4 and the freshness conjunction rule, we can get A9:

$$U \lhd \#\{(R_c, R_s, T_4)_{h(EID_u \oplus x)}, T_4, U \overset{sk}{\longleftrightarrow} S\}.$$

According to A8, A9 and the nonce verification, we can get A10:

$$U \models S \models \{(R_c, R_s, T_4)_{h(EID_u \oplus x)}, T_4, U \overset{sk}{\longleftrightarrow} S\}.$$

According to A10, we apply the belief rule, we can get A11:

$$G3: U \models S \models U \overset{sk}{\longleftrightarrow} S.$$

According to P7, A11 and the jurisdiction rule, we can get A12:

$$G1: U \models U \overset{sk}{\longleftrightarrow} S.$$

5.2 Comparison in Security Properties and Efficiency

The improved security properties of the proposed scheme, which is an extension of the Hsieh and Leu's scheme, are described as follows.

1) Identity preservation.
The adversary can easily intercept the user's login request $\{ID_u, T_u, C_2\}$ in Hsieh and Leu's scheme. In order to protect the identity of the legal user, we use the EID_u instead of the ID_u, what's more, the ID_u is encrypted by hash function. The content of M_2, M_3, M_6, M_7 are dynamic and different in each session by using R_c and R_s. Therefore, the proposed scheme can provide identity preservation.

Table 1: Security comparison of the proposed scheme with other related ones

Security features	Hsiang and Shih	Hsieh and Leu	Ours
Provide user anonymity	X	X	O
Mutual authentication	X	X	O
Resist user impersonation attack	X	X	O
Resist server masquerading attack	X	O	O
Resist slow wrong password detection	X	X	O
Resist offline password guessing attack	X	X	O

Table 2: Performance comparison of the proposed scheme with other related ones

Schemes	Login phase	Verification phase	Total
Hsiang and Shih's scheme	$2T_h$	$6T_h$	$8T_h$
Hsieh and Leu's scheme	$2T_h$	$6T_h$	$8T_h$
Proposed scheme	$5T_h$	$10T_h$	$15T_h$

2) Resist slow wrong password detection.
The proposed scheme can check the user's password during the login request. Therefore, it can quickly know whether the password is true or not. In the proposed scheme, when a user wants to login, he inputs his own ID_u, PW_u. On the basis of these, the smart card computes $f_u = h(ID_u \oplus h(PW_u))$ and compares it with the f_u stored in smart card. If the password is wrong, the computed f_u and stored f_u will be different, so the user can't login. At the same time, the user can quickly know he needs to input the correct password.

3) Resist user masquerading attack.
We suppose that an adversary can get the smart card and intercept the login request. If an adversary wants to masquerade as a legal user, he has to send the appropriate response to the server's request. When the adversary replays the login request $\{EID_u, M_2, M_3, T_1\}$ to the sever, the legal server responses $\{EID_u, M_6, M_7, T_2\}$ to the adversary, the adversary accepts it and must response the appropriate $\{M_9, T_3\}$ to the sever. However, he can't compute the correct $\{M_9, T_3\}$ without knowing ID_u, x and R_s, because the ID_u is encrypted by hash function and the x is only known by legal server and the R_s in the database.

4) Resist sever masquerading attack.
If an adversary wants to masquerade as a legal server, he has to send the appropriate response to the user's request. When the user sends $\{M_9, T_3\}$ to the adversary, he has to compute the appropriate $\{M_{10}, T_4\}$ to identify he is the legal server. However, he can't compute the correct $\{M_{10}, T_4\}$ without knowing ID_u, x and R_c, because the ID_u is encrypted by hash function and the x is only known by the legal server and the R_c in the database.

5) Resist password guessing attack.
If an adversary gets the smart card, he can extract all information stored in smart card. If he wants to guess the password, he can guess it by $f_u = h(ID_u \oplus h(PW_u))$. Although the adversary knows f_u, the ID_u is encrypted by hash function, so he can't get the password.

6) Provide mutual authentication.
The proposed scheme can provide mutual authentication because it can resist the user masquerading attack and the server masquerading attack. The security comparison of the proposed scheme with other related ones is presented in Table 1. O denotes that scheme provides the property; X denotes that scheme fails to provide the property. The result obviously indicates that our scheme is more secure.

The computation costs of the proposed scheme and other related ones are calculated in Table 2. In Table 2, T_h presents the computation time for hash function and T_s stands for the computation time for symmetric encryption operation. The computation time for $\oplus$ and $||$ can be ignored because the time is very short.

The results in Table 1 and Table 2 indicate that our scheme provide all security properties with only a slightly high computational cost.

6 Conclusion

This article first reviews Hsieh and Leu's scheme and then analyses the security of Hsieh and Leu's scheme. Secondly, we point the shortcomings of the scheme. Finally, we propose a new scheme to protect against all attacks. The results show that the proposed scheme has more secure properties than some other related ones.

Acknowledgments

This study was supported by 2017 "innovative action plan" of Science and Technology Commission of Shanghai Municipality (17050502000), 2017 cooperative project on Industry-Academy-Research of Shanghai Lingang Administrative Committee (Key technology research and demonstration line construction of advanced laser intelligent manufacturing equipment), the Doctoral Scientific Research Foundation of Shanghai Ocean University (A2-0203-00-100361). The authors gratefully acknowledge the anonymous reviewers for their valuable comments.

References

[1] N. Agray, W. V. D. Hoek and E. D. Vink, " On BAN Logics for Industrial Security Protocols," *Lecture Notes in Computer Science*, vol. 2296, pp. 29-36, 2001.

[2] H. Arshad and M. Nikooghadam, "An efficient and secure authentication and key agreement scheme for session protocol using ECC," *Multimedial Tools Applications*, vol. 75, no. 1, pp. 181-197, 2016.

[3] M. Burrows, M. Abadi and R. Needham, "A logic of authentication," *ACM Transactions on Computer System*, vol. 8, no. 1, pp. 18-36, 1990.

[4] C. C. Chang and T. C. Wu, "Remote password authentication with smart cards," *IEE Proceedings - Computer Digital Technology*, vol. 138, no. 3, pp. 165-168, 1991.

[5] D. B. He, J. H. Chen and J. Hu, "Weaknesses of a remote user password authentication scheme using smart card," *International Journal of Network Security*, vol. 13, no. 1, pp. 58-60, 2011.

[6] D. B. He, N. Kumar and J. H. Lee, "Privacy- preserving data aggregation scheme against internal attackers in smart grids," *Wireless Networks*, vol. 22, no. 2, pp. 491-502, 2016.

[7] D. B. He, N. Kumar, H. Shen and J. H. Lee, "One-to-many authentication for access control in mobile pay-tv systems," *Science China-Information Sciences*, vol. 59, no. 5, pp. 1-14, 2016.

[8] D. B. He, H. Wang, L. Wang, J. Shen and X. Yang, "Efficient certificateless anonymous multi-receiver encryption scheme for mobile devices," *Soft Computing*, vol. 21, no. 22, pp. 6801-6810, 2016.

[9] H. C. Hsiang and W. K. Shih, "Weaknesses and improvements of the Yoon-Ryu-Yoo remote user authentication scheme using smart cards," *Computer Communications*, vol. 32, no. 4, pp. 649-652, 2009.

[10] W. B. Hsieh and J. S. Leu, "Exploiting hash functions to intensify the remote user authentication scheme," *Elsevier Advanced Technology Publications*, vol. 31, no. 6, pp. 791-798, 2012.

[11] J. J. Hwang and T. C. Yeh, " Improvement on Peyravian-Zunic's password authentication schemes," *IEICE Transactions on Communications*, vol. 85, no. 4, pp. 823-825, 2002.

[12] M. S. Hwang and L. H. Li, "A new remote user authentication scheme using smart cards," *IEEE Transactions on Consumer Electronics*, vol. 46, no. 1, pp. 28-30, 2000.

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

[14] C. C. Lee, L. H. Li and M. S. Hwang, "A remote user authentication scheme using hash functions," *IEEE Transactions on Consumer Electronics*, vol. 36, no. 4, pp. 23-29, 2002.

[15] R. Lennon, S. Matyas and C. Mayer, "Cryptographic authentication of time-invariant quantities," *IEEE Transactions on Communications*, vol. 29, no. 6, pp. 773-777, 1981.

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

[17] D. Mishra, "Efficient and secure two-factor dynamic ID-based password authentication scheme with provable security," *Cryptologia*, vol. 42, no. 2, pp. 146-175, 2018.

[18] W. Teepe, "BAN Logic is Not 'Sound', Constructing Epistemic Logics for Security is Difficult," in *Proceedings of Famas*, 2006. (`https://pdfs.semanticscholar.org/bf93/01895b281c2ce6645f260d211833e8dbff03.pdf`)

[19] W. Teepe, "On BAN logic and hash functions or: how an unjustified inference rule causes problems," *Autonomous Agents and Multi-Agent Systems*, vol. 19, no.1, pp. 76-88, 2009.

[20] C. S. Tsai, C. C. Lee and M. S. Hwang, " Password authentication schemes: Current status and key issues," *International Journal of Network Security*, vol. 3, no. 2, pp. 101-115, Sept. 2006.

[21] J. L. Tsai, T. C. Wu and K. Y. Tsai, "New dynamic ID authentication scheme using smart cards," *International Journal of Communication Systems*, vol. 23, no. 12, pp. 1449-1462, 2010.

[22] D. Wang and P. Wang, "Offline Dictionary Attack on Password Authentication Schemes Using Smart Cards," *Information Security*, vol. 52, pp. 1212-1217, 2015.

[23] G. Wedel and V. Kessler, "Formal semantics for authentication logics," in *European Symposium on Computer Security (ESORICS'96)*, vol. 1146, pp. 219-241, 1996.

[24] 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,.

[25] S. Yen and K. Liao, "Shared authentication token secure against replay and weak key attack," *Information Process Letters*, vol. 62, no. 2, pp. 78-80, 1997.

[26] E. J. Yoon , E. K. Ryu and K. Y. Yoo , "Further improvement of an efficient password based remote user authentication scheme using smart cards," *IEEE*

Table 3: Main notations of BAN-logic

Notation	Meaning	Notation	Meaning
P, Q	Principals	X, Y	Message variable
K	Shared key	$\langle X \rangle_Y$	X combined with the formula Y
$P \models X$	P believes X	$\#(X)$	X is fresh
$P \lhd X$	P sees X	$P \overset{K}{\longleftrightarrow} Q$	P and Q may use the shared key K
$P \vdash X$	P once said X	(X, Y)	X or Y is one part of the formula (X, Y)
$P \Rightarrow X$	P has jurisdiction over X	$(X)_K$	X hashed under the key K

Transactions on Consumer Electronics, vol. 50, no. 2, pp. 612-614, 2004.

Appendix

In this section, we introduce the content of BAN-logic with symbols P and Q standing for principals and X and Y representing statements. The main notations of the logic are presented in Table 3.

To describe the logic postulates of BAN-logic, we present the following rules:

1) Message meaning rule.

$$\frac{P \models Q \overset{K}{\longleftrightarrow} P, P \lhd \{X\}_K}{P \models Q \vdash X}$$

P believes that Q once said X if P believes that K is the secret shared key with Q, and P sees X encrypted by K.

2) Nonce-verification.

$$\frac{P \models \#(X), P \models Q \vdash X}{P \models Q \vdash X}$$

P believes that Q believes X if P believes that X is fresh and Q has said X.

3) The belief rule.

$$\frac{P \models (X), P \models (Y)}{P \models (X, Y)}$$

P believes (X, Y) if P believes both X and Y.

4) Jurisdiction rule.

$$\frac{P \models Q \Rightarrow X, P \models Q \models X}{P \models X}$$

P believes Q on the validity of X if P believes that Q has jurisdiction over X.

5) Freshness conjunction rule.

$$\frac{P \models \#(X)}{P \models \#(X, Y)}$$

P believes that (X, Y) is fresh if P believes that X is fresh.

Biography

CAO Shouqi received his bachelor's degree in mechanical manufacturing technology and equipment from Sichuan University in 1996. He received his MS degree in mechanical manufacturing and automation from Sichuan University in 1999. He received his post doctoral degree in control science and Engineering in Shanghai University in 2009. Now he is a professor and doctoral supervisor at the College of Engineering Science and Technology, Shanghai Ocean University. His main research interest is marine Internet of things engineering, Fisheries Engineering and automation technology research.

SUN Qing received her bachelor's degree in electrical engineering and automation from Huaiyin Institute of Technology in 2016. Now, she is a student at the College of Engineering Science and Technology, Shanghai Ocean University. Her main research is communication security and Internet of things technology.

Cao Liling received a bachelor's degree in electronic information science and technology from Central South University in 2004. She received her MS degree in physics electronics in Central South University in 2007. She received a Ph.D. degree in testing technology and automation from Tongji University In 2017. Now she is an experimental teacher of College of Engineering Science and Technology, Shanghai Ocean University. Her main research is Network security, authentication protocol.

Detection Algorithm for Sinkhole Attack in Body Area Sensor Networks Using Local Information

Adnan Nadeem and Turki Alghamdi
(Corresponding author: Adnan Nadeem)
Faculty of Computer, Information System, Islamic University of Madinah
Madinah, Saudi Arabia
(Email: adnan.nadeem@iu.edu.sa)
(Received Jan. 20, 2018; Revised and Accepted Apr. 28, 2018; First Online Jan. 11, 2019)

Abstract

Wireless Body Area Sensors Networks (BAN) have emerged as new applied wireless networking technology with the development of wearable and implanted sensors. BAN has novel application in healthcare, sports, human activity monitoring, disability assistance and entertainment. BAN is now using for real time monitoring and assistance of the patients. BAN operations are vulnerable to various security attacks, including basic and advance attacks. In this paper, we introduce and illustrate the sinkhole attack in a BAN. Then we propose our sinkhole detection algorithm that utilizes the information from data aggregation algorithm to detect a sink hole attacker. Finally, we analyze the performance of the BAN in terms of throughput, latency and packet breakdown and the performance of our detection algorithm. Simulation results show that this attack could severely degrade (up to 40%) the overall performance of the network. The propose detection algorithm has good performance in terms of high success (85% on average) and low (6% on average) false alarm rates.

Keywords: Body Area Networking Technology; Performance Analysis of BAN; Security & Privacy; Sinkhole Attack

1 Introduction

Wireless Body Area Sensor Networks (BANs) is an emerging wireless networking technology. It consist of wearable sensors with the capability of monitoring physiological parameters of the body *e.g.* ECG, temperature, heart rate, EMG and blood pressure measurements [4, 19]. BAN has its applications in health-care, fitness, sports and entertainment. Beside these major applications some novel applications areas of BAN has also emerged recently. BAN consists of wearable or implanted sensors, data aggregator and a gateway device called sink, where all the sense information is aggregated for analysis and decision making. All the data sense by the sensors must be routed to the gateway device. However, this process of data aggregation and routing is vulnerable to various attacks. Specifically in health care applications of BAN where it use to monitor and assists patients health the presence of malicious node could be life threatening [6, 18].

Security & Privacy is one of the major concerns for the researchers involve in BAN along with energy efficient operations. Considering the healthcare applications of BAN, security and privacy of information communicated over the network become highly important. BAN like other networks is also vulnerable to a range of security attacks [11] that could seriously degrade the performance of the network. Sink hole attack is one of them, in this the attacker gets attach with the network claiming to be a sink node and causes both security and privacy issues. Therefore, in this paper we first illustrate the sink hole attack in BAN and then propose out sink hole detection algorithm that utilizes the audit data from the data aggregation techniques to detect sinkhole attack. We analyze the affect of this attack on BAN performance and the performance of our sink hole detection algorithm using a simulation based case study.

The rest of the paper is organized as follows: Section 2 presents the overview and classification of security attacks in a BAN. Section 3, we present the illustration of sinkhole attack. In Section 4, we briefly review the related work. Then we present our proposed sink hole detection algorithm in Section 5. In Section 6, we present the performance analysis of BAN under sinkhole attack and the performance of its detection algorithm, including the simulation results. Finally, we summarize our work and highlight possible future work in Section 7.

2 Security Attacks in BAN

Similar to wireless sensor network (WSN), BAN is also vulnerable to various attacks. Authors in [2] have defined the threats and their security requirements in BAN. Table 1 illustrates the threats and the related security requirements. It mainly discusses the classical basic secu-

rity requirements including integrity, confidentiality, authentication, availability which exists in almost all data communication networks.

Table 1: Threats and related security requirements in BAN [2]

Threats	Security Requirements
Data Modification	Integrity
Impersonation	Authentication
Eavesdropping	Confidentiality
Replying	Integrity
DoS	Availability

We classify security attacks in BAN as either basic and advance attacks. Basic attacks include all the attacks with traditional security attributes/ requirements. Whereas advance attack we include the specialized attack that could be launch in BAN by the attacker to achieve the certain goal. Figure 1 presents our classification of attacks in BAN. Attacks in basic attack category has been extensively discussed in the literature therefore, we will only discuss the advance attacks.

2.1 Data Freshness

Decisions made by physicians or health caregivers are mainly dependent on the freshness of data. Therefore, replaying old messages in WBAN could cause serious consequences.

2.2 Reliability

Due to the type of sensors and its energy constraints, operations reliability of nodes and operations in BAN is a major issue. The BAN applications have several Reliability [9] & Quality of Service [13] problems. Considering the health care applications of BAN this issue could be significant. In emergency situations, if the data is not communicated within the specified time period then it can incur serious consequences even a loss of life. Devices implanted inside the human body are prone to absorption and attenuation because of material composition and structure of the human body.

2.3 Trust Management

Energy restrictions make the key distribution between the nodes a major challenge. Public key cryptography, which is majorly used in Digital signatures for key exchange consumes much more energy than Symmetric cryptography. Therefore, authors in [15] propose static node deployment for energy efficient operations. Considering the energy Moreover, as per the new observation, same physiological values monitored from different parts of the body within the same time frame, exhibit similar characteristics, which can put the Trust management procedure on stake.

2.4 Privacy

Several aspects of the Privacy and social issues exist in WBANs. Health records can be stolen upon by the emergency technician in case of emergency for monetary gains. This issue arrives when extra privilege to information is granted, thus leading to theft of data private to the patient. This may include name, social security number, mailing address, medical record history, *etc.* Also, people might not want some data to be made public *e.g.* early stage pregnancy. Below are the attacks which deal with privacy.

2.4.1 Monitoring and Eavesdropping

Monitoring and eavesdropping is an attack for privacy. The attacker can easily gather the data by snooping.

2.4.2 Traffic Analysis

The attacker can read and understand the communication between two parties by getting traffic patterns and can be harmful to legitimate users.

2.4.3 Camouflage Adversaries

An attacker can introduce a new node or tries to compromise the other nodes by hiding it in the sensor network. Sensor nodes pretend themselves as a common network node in order to capture the packets.

2.4.4 Privacy

Privacy issue also exists on the storage server/site as the site is aware of the ownership of records *i.e.*, which record belongs to which patient. Moreover link ability of records can help stealing vast amount of data linked among one another. Furthermore, Location privacy breach can expose the knowledge of patient?s whereabouts and location, calculated by exploiting the capability of the sensors installed. Privacy has a strong association with the security aspects of Access control and Authorization [10]. Biological signals collected from ECG and EEG can reveal information of psychological status and identity of the subject, which can reveal emotion assessment and thus raise privacy concerns [1].

2.4.5 Sinkhole Attack

A sinkhole is a denial of service attack well defined and extensively research in WSN. In this paper, we first describe and illustrate this attack in a BAN. We have considered multi hop scenario of BAN where a malicious node falsely announces itself as a sink node. The entire sensor node sends their information to this node which drops all the information [12]. There are various techniques have been proposed to detect attacks in wireless sensor network some of them using cryptographic techniques such as [3,16,20], however, few researchers have focus on investigating it in body area network. We believe this attack

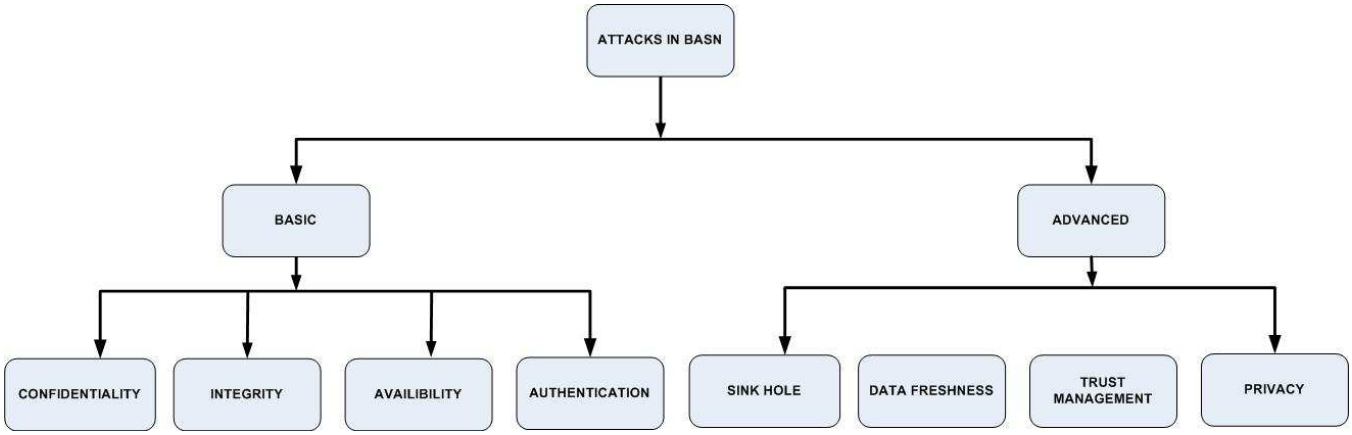


Figure 1: Security attacks in BAN

could seriously degrade the performance of the body area network. This motivates us to analyze the effect of sinkhole attacks on the performance of BAN in this paper.

3 Illustration of Sinkhole Attack in BAN

The sinkhole attack is one of the severe attacks that prevents the legible sink or gateway node in receiving complete and correct information, and creates a severe threat to applications. In a Sinkhole attack [12,14]; A malicious node tries to capture whole traffic from network, by impersonating itself as a sink node in the network. As a result, the attacker gets all traffic that is to be transmitted to legitimate sink node. In this way it can then introduce various severe types of attacks, like selective forwarding, modifying or even dropping the packets coming through.

Wireless body area sensor network plays important role in health-care applications from basic patient monitoring to the specific disease monitoring and detection. Third generation of sensors kits such as ECG and EMG kits are available to use in various healthcare tasks. We assume third generation wearable sensors such as temperature, blood pressure, ECG. The model which we have used is shown in Figure 2. There are six nodes and their placement is as follows.

Table 2: Placement of nodes on the Human body

Node	Placement
0	Right Hip (Sink)
1	Left Arm
2	Right Arm
3	Left Ankle
4	Right Ankle
5	Chest
6	Right Hip

We now consider the network in Figure 2 and illustrate how an attacker can launch sinkhole attack. This network consists of five sensors and a sink node. The nodes in the network operate in a multi hop scenario.

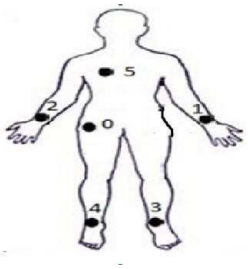


Figure 2: Show the model of BAN

Figure 3 shows the scenario of normal nodes with green lines connected with each other wirelessly and they operate normally. The green boxes show the normal packet flow between the sink node and the other nodes in the network. Figure 4 shows a scenario where an attacker

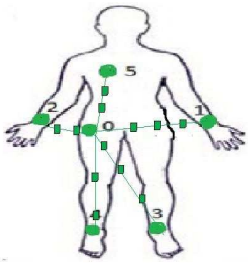


Figure 3: Show the scenario with normal operations of BAN

gets connected with the network. This node which is not an authorized node can act as a sink hole and affects the performance of the network. The malicious node after being the part of network tries to capture network traffic by announcing himself as a sink to all nodes. This is done through sending a false message as shown in Figure 5.

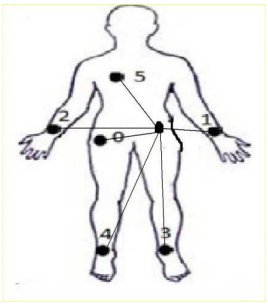


Figure 4: Shows the scenario with an attacker connected to a network

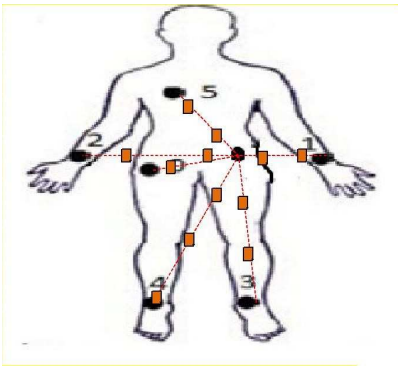


Figure 5: Attacker falsely announcing himself as sink node

When data reaches this malicious sink node, then instead of forwarding the packets to the actual sink it drops the packets as shown in Figure 6. This behavior of attacker prevents data traffic from reaching the legitimate sink node. This could seriously degrade the performance of the network. In this paper we have perform an analysis of the degree of impact that this attack can have on the performance of the network and parameters on which the level of performance degradation depends.

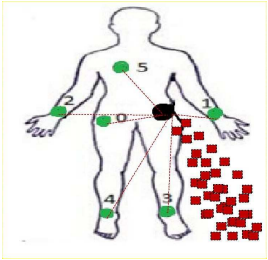


Figure 6: Shows packets dropped by attacker node

4 Related Work

Security and privacy is one of the prime concerns in the BAN research [12,14]. Several studies have suggested different type of detection algorithms in wireless sensor networks with regard to sinkhole attack. The sinkhole attack normally occurs where; there is symmetric traffic among the sensor nodes [14]. Sinkhole attack is devastating because of the weak computation and battery power of the sensor nodes in these networks.

Karlof [8]propose a trust scheme to the routing protocol for detection of sinkhole and wormhole attacks in a sensor network; however activity of nodes in a loose mode is essential. It has been shown that packet restriction can disclose the limit of transfer time and each packet?s distance. It has been suggested that strong authentication mechanism should be used to avoid such types of attacks in wireless sensor networks [12, 14].

Authors in [4,12,14] first suggest a way to detect sinkhole attacks in which the BS in the detection process, causing an increased communication cost for the protocol. The network is flooded by the BS with a request message and the IDs of the nodes which are much affected. These nodes reply to the BS with a message including IDs, next hop ID and its cost. The sinkhole can be detected on the basis of that received information. Other protocols agree to detection methods for sinkhole attacks in sensor networks that are use routing protocols usually Ad Hoc On-demand. Distance Vector Protocol (AODV) and the Dynamic Source Routing (DSR) Protocol [6]. As discussed above, this many-to-one message passing model is susceptible to sinkhole attacks. In sinkhole attack, an adversary usually gets the traffic of whole network by sending broadcast about its presence and pretends to a sink node for all nodes in the network or a node providing shortest path towards sink node. For example, a malicious node, with higher computational resources and communication power as compared to ordinary sensing node, and creates a better-quality single-hop link to nodes existing there. In the end, it broadcasts short routing messages regarding that high quality link, spoofing the neighboring nodes to create a sinkhole (SH). A sinkhole can also be created by using a wormhole, which creates a sinkhole with the attacker being the center; the intruder then forwards the messages toward the sink using a tunnel [4]. Most of the research has investigated and proposed mechanism for sinkhole attack in WSN; in contrast in this paper we investigate the sinkhole attack in BAN which utilizes the local information content from the data aggregation algorithm [7].

5 Detection of Sinkhole Attack

In this section, we propose our detection scheme for the sinkhole attack scenario illustrated in Section 3. We use the terminologies define in Table 3 to presents out idea.

5.1 Sinkhole Launching Strategy

From the attacker?s perspective the most important task to launch this attack in a single BAN scenario are as follows:

- Attacker SNK_Hole impersonates the original sink node *SNK*.

Table 3: Terminologies used by the algorithm

SN_i	Represents sensors nodes where i is its ID
SNK	Represents the original sink node
SNK_Hole	Represents the sink hole attacker
Req_Data	A packet sent from SNK to SNi to request data
TI	Time interval
n	Periods of data aggregation used for detection

- *SNK_Hole* sends a *Req_Data* packet to all SN_i.
- If impersonation successful then *SNK_Hole* will receive data from all nodes as a reply to *Req_Data* and will simply drop them to create the sink hole.
- Legitimate *Req_Data* received from *SNK* later will be processed then.

The data aggregation in BAN could be either

1) Periodic,
2) Event driven,
3) Combination of both periodic and event driven.

In periodic the *SNK* sends the *Req_Data* after a certain time period periodically for example in a general patient monitoring scenario where all body parameters needs to be monitor for maintaining patients history. On the other hand event driven data aggregation will trigger on the occurrence of certain event for example critical level of blood glucose is notice by the sensor. In this case the sensor node will transmit the data to *SNK*, from where it will be transmitted to doctors or to emergency service providers.

In both type of data aggregation schemes the above mention sinkhole launching strategy will work in the scenario illustrated in Section 3. Simply because if there is no means for the SN_i nodes to differentiate between authorize and un authorize sink then the *SNK_Hole* will receive all the data instead of *SNK*. Having a proper authentication procedure in place will certainly stop this type of attack. However, we learn from the literature that the cost of implementing such mechanism is generally are on the higher side for BAN application. Therefore, in this paper we assume there is no authentication service is in place and instead we propose to use the information from the data aggregation protocols to distinguish between and *SNK* and *SNK_Hole*.

5.2 Model Assumptions

We assume the sink hole attack scenario illustrated in Section 3. We assume energy efficient multi hop data aggregation technique such as DARE [17] in place. It uses the concept of relay nodes (a multihop scenario) to efficiently utilize the energy of the nodes in the network. It is a distance aware protocol means before the transmission of data it estimates the residual energy and distance between the sensors, relay and sink node. There are two possible placement of *SNK_Hole* attacker node:

1) On the body of the patient as illustrated in Figures 5 and 6;
2) Outside the body of the patient. We consider the later as in earlier case the patient will notice if extra sensors is attach to the body. We further assume the stationary sink node *i.e.* no mobility.

5.3 Core Functionality of Proposed Method

We now describe the core functionality of our detection mechanism. It mainly consists of two modules data aggregation and Sinkhole Detection.

5.3.1 Data Aggregation

The sink node sends a *Req_Data* to all the SN_i We employ energy efficient multi hop data aggregation technique in [15]. It estimates the transmission and reception energy using the basic radio model proposed in [5] are given below as Equations (1) and (2).

$$E_{TX}(k,d) = E_{TX\,elec} \times k + E_{amp}(n) \times k \times d_n \quad (1)$$

$$E_{RX}(k) = E_{RX\,elec} \times k. \quad (2)$$

Here, E_{TX} in Equation (1) represents the transmission energy and Equation (2) calculates the receiving energy represented by E_{RX}. k represents the number of bits transmitted, d represents the distance. The radio energy dissipates by the transmitter and receiver is represented by $E_{TX\,elec}$ and $E_{RX\,elec}$. E_{amp} is the energy for the transmit amplifier and the d is the distance between sender and receiver.

We consider the scenario shown in Figure 3 and perform the data aggregation in the following steps:

- It first measure the distance between the SN_i and sink *SNK*.
- It then estimates the transmitted energy of sensor and received energy of relay node or sink.
- Based on the estimated energy and distance it selects the multihop path to aggregate data.
- This process continues until sense data from all the SN_i is received.
- It also maintains the residual energy of relay and *SNK* node.

5.3.2 Detection of Sinkhole

We consider the sinkhole launching strategy and data aggregation technique describe earlier. We propose to utilize the parameters related to energy and distance maintain during the data aggregation to identify the *SNK_Hole* attacker. We define the data aggregation is done periodically after each time interval *(TI)* for n periods. We use the concept of anomaly based detection, where we employ two mechanisms training and testing. In training we maintain the expected normal profile of the parameters from data aggregation in *EXPECTED* matrix. Testing process in invoked when training profile is build. In testing the algorithm maintain the current values of the parameters in *OBSERVED* matrix. During the testing *OBSERVED* matrix parameters are statistically compared with *EXPECTED* and in case of significant statistical deviation we declare the node as *SNK_Hole* attacker.

Algorithm 1 illustrate the propose sinkhole detection process in BAN. It requires the maintenance of two matrixes *OBSERVED* and *EXPECTED* with three parameters. Where the later represents the expected parameters values related to distance and energy of sink node and the earlier matrix represents the current information received from the node claiming to be sink. Since we consider the specific placement of sink node on the body, therefore technically these two matrixes should not be significantly different. To reduce the possibility of false detection we calculate the statistical deviation $(S.D)$ based on observation from n periods. The algorithm is general and the detection parameters values such as number of parameters in two matrix?s, n and threshold could be modified to implement the algorithm in different scenarios.

5.3.3 Algorithm

Detection of sinkhole is done in the following steps:

- The detection module maintains the updated information regarding the relay and the *SNK* of data aggregation parameters.
- Repeat after each TI for n periods
  - Updated values of E_{TX}, E_{RX}, d, are kept in the textitSN_i as *EXPECTED* matrix.
    $$EXPECTED=\{E_{TX},\ E_{RX},\ d\}$$
  - When the SN_i receive the *Req_Data*, it will obtain the parameters from the data aggregation algorithm term as *OBSERVED* matrix. $OBSERVED=\{O.E_{TX}, O.E_{RX}, O.d,\}$ SN_i is received.
  - Compare the current values of distance and energy parameters from the algorithm regarding the sink node/ relay node with the previous information store in the table.
- End repeat
- Calculate statistical deviation SD using the Equation (3):

$$S.D = \frac{\sum_{i=1}^{n} OBSERVED_i - EXPECTED_i}{n}. \tag{3}$$

- If $(S.D > threshold)$ then

  we confirm the node as *SNK_HOLE*

  Else

  We conclude the node as genuine *SNK* Update *EXPECTED* matrix using Equation (4)

$$\begin{aligned} & \forall_i(\overline{EXPECTED}_n^i) \\ = \ & \alpha \times OBSERVED_n^i + (1-\alpha) \times \overline{EXPECTED}_n^i. \end{aligned} \tag{4}$$

- End

We use exponentially weighted moving average to update matrices using Equation (4), where $\overline{EXPECTED}_n^i$ and $OBSERVED_n^i$ represents the expected and observed matrix with i parameters and n time interval. Here $\alpha = 2/(n-1)$ is the weighting factor.

5.4 Complexity Estimation of Proposed Method

Now we estimate the running time & complexity of proposed algorithms. We assume a single non-iterative task takes t seconds to complete. Total number of times the algorithm module runs is n *TIs*. Now we consider Algorithm 1 pseudo-code of sink hole detection phase, which can be split into three tasks for estimation of their time complexity.

1) Collecting and maintaining updated values of *EXPECTED* and *OBSERVED* matrices. Running time of this part can be estimated using further dividing into three tasks.
   a. Estimating and storing data for j parameters of *EXPECTED* matrix, so time complexity will be $j * t$.
   b. Obtaining and storing parameters from data aggregation algorithm for j parameters $j*2*t$.
   c. Comparing j parameters of matrices $j*t$.

   So the running time of task 1) is $= jt + j * 2t + jt = 4jt$ As this task repeats for n $TI = n * (4jt)$.

2) Calculation of statistical deviation of j parameters of two matrices.
   Running time of this part can be estimated using further dividing into two tasks.
   a. Calculating $S.D$ using equation in Algorithm 1 $j * n * t$.
   b. Comparison S.D computed and threshold values $j * t$.

So running time estimation for task 2) $= j(nt + t)$.

3) Update *EXPECTED* matrix.
Updating expected values $j * n * t$.
So running time estimation is $= j * n * t$.

Now combining task 1), 2) and 3).
Running Time (complexity) $= n(4jt) + j(nt + t) + jnt$.
Which can be simplified to Running Time (complexity) $= (6nj + j)t$.
If we remove constant then the expression in big-Oh notation will be $O(j(n+1))$. In general we can say that the running time complexity of the detection algorithm will depend on the j *(number of parameters in matrices)* and n *(the number of data aggregation periods)*. This could also give us the estimate of cost of the detection algorithm in the scenario it is implemented.

6 Performance Evaluation

We now present the performance analysis of BAN under sink hole attack and the result of our detection algorithm. We have used Castalia which is based on OMNET ++ platform to simulate the BAN scenario. We consider the BAN in Figure 2 and create a simulation scenario using the simulation parameter in Table 4. Each node in our

Table 4: Simulation parameters

Parameter	Value
No of nodes	6 nodes, node 0 is sink
Transmit Power	-15dB
Simulation Time limit	600 sec
Start up delay	1 sec
Packet rate	30 pkt/sec

scenario sends certain number of packets per seconds for the simulation time. We run our scenarios first with no sink hole attack then intentionally created a sinkhole attack to analyze its effect on network performance using packet received, latency and packet breakdown (errors) as basic parameters. We run these scenario with GTS is turn on or off along with either temporal channel (*i.e.* path loss exists) and no temporal channel (*i.e.* no path loss exists).

All our simulation is performed using the body area network scenarios shown in Figure 2 with six sensors are placed at different parts of the body. We use the simulation parameter in Table 4. The graph in Figure 7 shows the results of the normal scenario (no attack) with packet received per node (all nodes send their data to node zero so the term per node is used). The graph shows variations in the number of packets received by six nodes in the network with respect to the various GTS options. The graphs show GTS on with no temporal has slightly better performance as compare to other GTS options. In this scenario we assume there is no attack in the network, therefore, the graph reflects the normal behavior of nodes in the network. Figure 8 demonstrates the second case

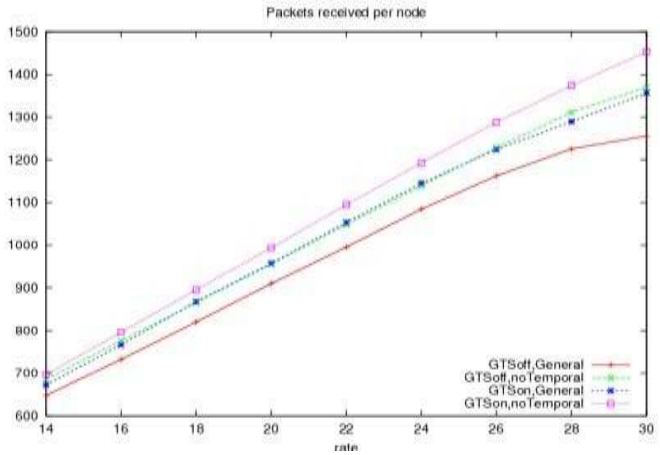


Figure 7: Packets received per node with no sinkhole

in which we have introduced sinkhole attack. In this case where an attacker node acts as a sink node; as a result the GTSon General and GTSoff General curves have fallen drastically because the node 6 is dropping all the packets which it receives from the neighboring nodes. This shows the significant degradation in the network performance. In this scenario we introduce sink hole attack and the drastic change of performance in terms of received packets per node is evident. We have also observe the latency

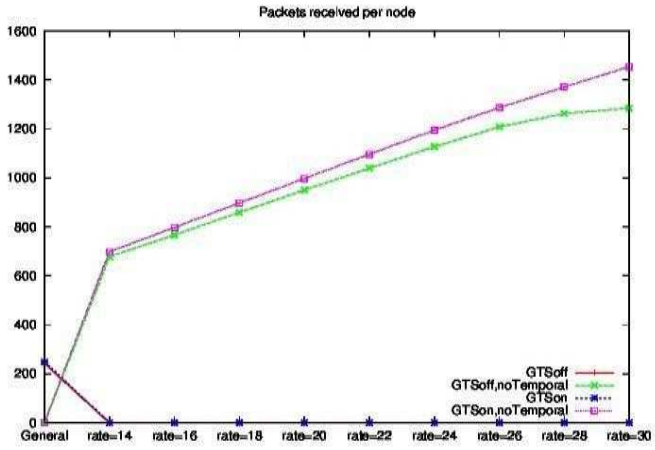


Figure 8: Packets received per node with sinkhole

in both scenarios (*i.e.* with and without sink hole attack), Figure 9 and 10 shows the effects of sinkhole attack on latency. We can see the graph in Figure 9 the latency of majority of the packets is less than 100ms, it shows those packets transmitted in the first MAC frame. In this case no temporal performs better but there is some saturation in temporal case. However, the graph in Figure 10 shows the packets received within the first attempt are quite good in number but later on there are large number of packets with large delay. There is a huge latency shown in general case but a considerable increase is shown in

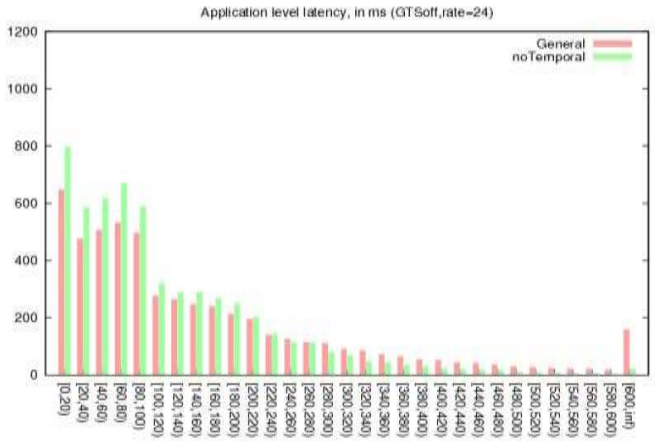


Figure 9: Latency in general and no temporal (GTSoff) with no sinkhole

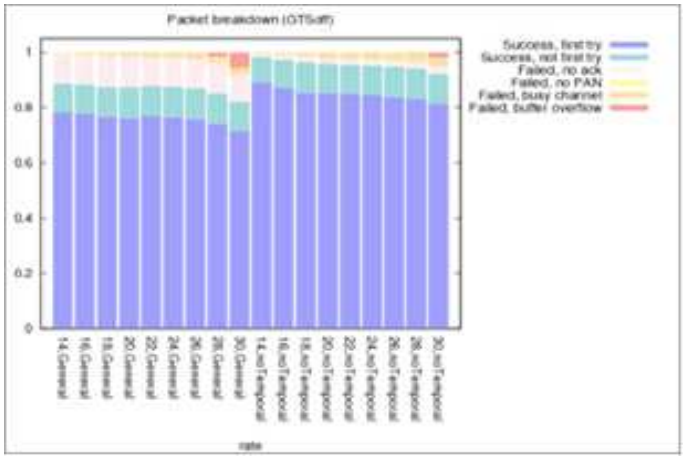


Figure 11: Packet breakdown in noTemporal (GTSoff) without sinkhole

noTemporal case. Analyzing the effect of the presence of

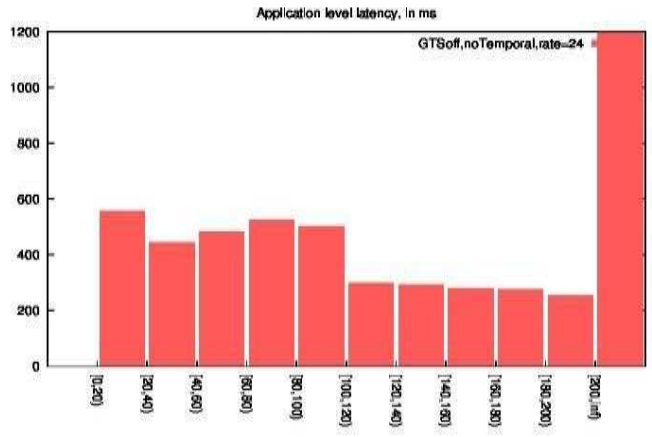


Figure 10: Latency in General and noTemporal (GTSoff) with no sinkhole

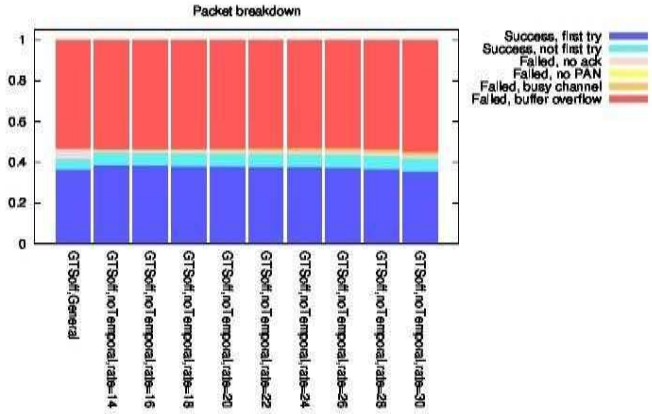


Figure 12: Packet breakdown in noTemporal (GTSoff) with sinkhole

sinkhole attack on packet latency from graphs in Figures 9 and 10 clearly shows the degradation of performance with sink hole attack in the network.

We have also analyzed the Packet breakdown with both scenarios (with and without attack). The graphs in Figure 11 show that the most of the packets failed because of noAck (a direct result of the deep fades in the channel and loss of connectivity) and overflow in the case of high rates. The packet drop rate of busy channel and buffer overflow is negligible, but 90% of the packets are received in successfully. There are almost 80% of total packets received properly in first try because there is no attacker in this case. The graphs in Figure 12 show that the most of the packets failed because of buffer overflow because the attacker is creating such a condition and going to drop the packets and this overflow occurs due to high rates. The packet drop rate of busy channel and buffer overflow is almost 50%. There is 40% more packet loss in the first try because of sinkhole attack and this clearly indicates the degradation of performance in the network.

To sum up this simulation performance evaluation indicates that the sinkhole attack could severely degrade the performance of the network.

In the final set of experiments we implemented our propose sink hole detection algorithm using the same simulation parameters in Table 4. We simulated the scenario with data aggregation technique of [7] and radio model of [5] and introduce the *SNK_HOLE* during the simulation in the network. We perform 10 runs each set of experiments with sink hole attack introduce in the network and observe the detection rates of success and false alarm. The graph in Figure 13 shows success and false alarm rate in the five set of experiments. Success rate here means that the *SNK_HOLE* attacker was detected successfully during the experiments. False alarm rate means the number of time the normal node or genuine sink node is detected by the algorithm as attacker. The graph in general shows the high success and low false alarm rate of our proposed algorithm. The major issue with anomaly based detection scheme is high false alarm rates; therefore we have

declared detection based on the outcome of n periods instead of a single run.

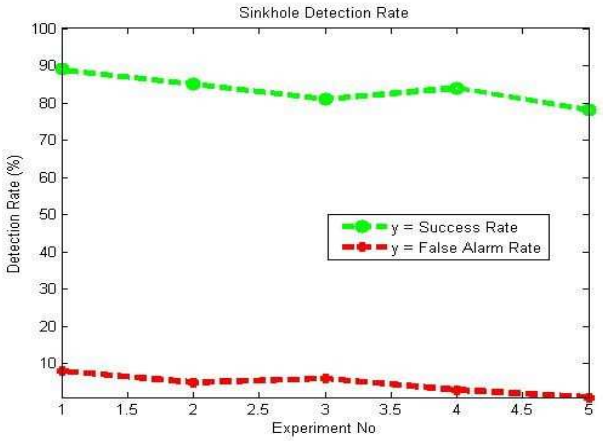


Figure 13: Detection rates of proposed sink hole detection algorithm

7 Conclusion & Future Work

Wireless body area sensor network is emerging technology that has applications in major walks of life especially in healthcare. However, BAN operations are vulnerable to security attacks. Considering the security requirement of healthcare application of BAN, in this paper we have analyze the performance of the BAN under sinkhole attack scenario. We propose the sink hole attack detection algorithm that utilize the distance and energy related information from the data aggregation technique to detect the sink hole attack in BAN. The simulation base study shows that this attack could severely degrade the performance of the network in terms of low throughput, higher delay and packet breakdown. Simulation results show good performance of our detection algorithm in terms of high detection and low false alarm rates.

In future our focus is on investigating security and privacy issues in multi BAN scenario applied to hospital ward. That is to use the multi BAN to remotely monitor all the patients in a ward using the wearable shimmer sensors. Then study, identify and propose solution for the privacy and security issues in this scenario.

Acknowledgments

We acknowledge the efforts of MS student Mr Ghullam Abbas in terms of simulation scenarios of performance analysis.

References

[1] F. Agrafioti, F. M. Bui, and D. Hatzinakos, "On supporting anonymity in a ban biometric framework," in *16th International Conference on Digital Signal Processing*, pp. 1–6, 2009.

[2] M. A. Ameen, J. Liu, and K. Kwak, "Security and privacy issues in wireless sensor networks for healthcare applications," *Journal of Medical Systems*, vol. 36, no. 1, pp. 93–101, 2012.

[3] S. Bouchkaren and S. Lazaar, "Caes cryptosystem: Advanced security tests and results," *International Journal of Network Security*, vol. 20, no. 1, pp. 177–183, 2018.

[4] Boulis, A. Castalia, *A Simulator for Wireless Sensor Networks and Body Area Networks*, ver. 2.2. User's Manual, NICTA: Canberra, Australia, 2009.

[5] B. Braem, B. Latre, I. Moerman, C. Blondia, E. Reusens, W. Joseph, L. Martens, and P. Demeester, "The need for cooperation and relaying in short-range high path loss sensor networks," in *International Conference on Sensor Technologies and Applications*, pp. 566–571, 2007.

[6] M. Deylami and E. Jovanov, "Performance analysis of coexisting ieee 802.15. 4-based health monitoring wbans," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 2464–2467, 2012.

[7] N. Javaid, Z. Abbas, M. S. Fareed, Z. A. Khan, and N. Alrajeh, "M-attempt: A new energy-efficient routing protocol for wireless body area sensor networks," *Procedia Computer Science*, vol. 19, pp. 224–231, 2013.

[8] C. Karlof and D. Wagner, "Secure routing in wireless sensor networks: Attacks and countermeasures," *Ad Hoc Networks*, vol. 1, no. 2-3, pp. 293–315, 2003.

[9] O. U. O. Khan, A. Nadeem, K. Ahsan, and N. Mehmood, "Rprp: Reliable proactive routing protocol for wireless body area sensor network," *Journal of Basic and Applied Scientific Research (JBASR'14)*, vol. 4, pp. 17–25, 2014.

[10] M. Li, W. Lou, and K. Ren, "Data security and privacy in wireless body area networks," *IEEE Wireless Communications*, vol. 17, no. 1, pp. 51–58, 2010.

[11] A. Nadeem and M. P. Howarth, "A survey of manet intrusion detection & prevention approaches for network layer attacks," *IEEE communications Surveys & Tutorials*, vol. 15, no. 4, pp. 2027–2045, 2013.

[12] J. Qi, T. Hong, K. Xiaohui, and L. Qiang, "Detection and defence of sinkhole attack in wireless sensor network," in *IEEE 14th International Conference on Communication Technology (ICCT'12)*, pp. 809–813, 2012.

[13] A. Salam, A. Nadeem, K. Ahsan, M. Sarim, and K. Rizwan, "A novel QoS algorithm for health care applications of body area sensor networks," *Textroad Journal of Basic and Applied Scientific Research*, vol. 4, no. 1, pp. 169–178, 2014.

[14] S. A. Salehi, M. A. Razzaque, P. Naraei, and A. Farrokhtala, "Detection of sinkhole attack in wireless sensor networks," in *IEEE International Conference on Space Science and Communication (IconSpace'13)*, pp. 361–365, 2013.

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

[16] M. Styugin, "Establishing systems secure from research with implementation in encryption algorithms," *International Journal of Network Security*, vol. 20, no. 1, pp. 35–40, 2018.

[17] A. Tauqir, N. Javaid, S. Akram, A. Rao, and S. N. Mohammad, "Distance aware relaying energy-efficient: Dare to monitor patients in multi-hop body area sensor networks," in *Eighth International Conference on Broadband and Wireless Computing, Communication and Applications (BWCCA'13)*, pp. 206–213, 2013.

[18] Y. Tian, Y. Peng, G. Gao, X. Peng, "Role-based access control for body area networks using attribute-based encryption in cloud storage," *International Journal of Network Security*, vol. 19, no. 5, pp. 720–726, 2017.

[19] C. Tumrongwittayapak and R. Varakulsiripunth, "Detecting sinkhole attacks in wireless sensor networks," in *ICCAS-SICE*, pp. 1966–1971, 2009.

[20] I. C. Weng and T. H. Chen, "A novel weighted visual cryptography scheme with high visual quality.," *Internationl Journal Network Security*, vol. 19, no. 6, pp. 922–928, 2017.

Biography

Dr. Adnan Nadeem is currently working as an Associate Professor in the Faculty of Computer Science and Information System, Islamic University in Madinah, KSA since 2016. He is also associated with Federal Urdu University of Arts Science & Technology, Pakistan since March 2011. During this period, he earned several research grants. He was awarded 5th HEC Outstanding Research Award 2013/14 for his paper published in IEEE Journal of Communication Survey and Tutorials (Impact Factor=17.18). During his pedagogical journey he has won several awards and achievements including the Foreign PhD scholarship, Associate Fellowship of Higher Education Academy (AFHEA), UK in 2009 and best paper & best paper of the track award in the ICICTT 2013 and ICEET 2016 conferences, respectively. He was awarded "Nishan-e-Imtiaz" for his outstanding research by Federal Urdu University Pakistan on August 2016. He received his PhD degree from Centre for Communications Systems Research, (CCSR) University of Surrey, UK in 2011. He has published more than 40 papers in international conference and journals. His research interests include WBAN applications in healthcare, agriculture and disability assistance. He also worked in security, routing and QoS in MANET and WSN.

Turki Alghamdi is currently working as an Assistant Professor, the Dean, and the Founder of the Faculty of Computer and Information Systems at Islamic University in Madinah, KSA. He received a BSc in Computer Science from Taif University, KSA in 2005, and MSc in Software Engineering from University of Bradford, UK in 2008 . He received a PhD in Software Engineering from De Montfort University, UK in 2012.

A Searchable CP-ABE Privacy Preserving Scheme

Tao Feng[1], Xiaoyu Yin[1], Ye Lu[2], Junli Fang[1], and Fenghua Li[3]
(Corresponding author: Tao Feng)
School of Computer and Communication, Lanzhou University of Technology, China[1]
College of Electrical and Information Engineering, LanZhou University of Technology, China[2]
The State Key Laboratory of Information Security, Institute of Information Engineering,
Chinese Academy of Sciences, Beijing[3]
(Email: fengt@lut.cn)
(Received Jan. 9, 2018; Revised and Accepted May 22, 2018; First Online Mar. 12, 2019)

Abstract

The separation of users and data in the cloud storage system brings new security threats. The cloud storage scheme with single encryption mechanism has been unable to meet new demands. Aiming at the above problems, the PEKS was introduced into the CP-ABE scheme with multiple authorities to propose a searchable CP-ABE privacy-preserving scheme which supports the direct revocation of users. The access control of users is achieved by the central authority, which avoids the security risks caused by submitting the private keys and access structure to cloud server. The collusion problems are solved by using 1-out-of-n Obvious Transfer algorithm and associating the private key components with random identity token of users. The privacy of keywords is guaranteed though the new improved algorithm. Based on the DBDH assumption, the security of proposed scheme is proved in the random oracle model.

Keywords: Attribute-based Encryption; Cloud Storage; Privacy Preserving; Searchable Encryption

1 Introduction

1.1 Background

Cloud storage services as an emerging data management technology develops rapidly, which achieves data storage, searching and sharing by providing the dynamic network resources for users. However, the separation of data owner and data in the cloud storage system brings new security threats, and the frequent information leakage has triggered the trust crisis of cloud storage service. The privacy preserving gradually becomes the core issue in cloud storage research. In order to apply the cloud storage system to the core data management and realize the secure data sharing mechanism, the privacy preserving mechanism in the attribute-based encryption scheme must be more perfect and the security of the system needs to be further improved. The research about attribute-based encryption cloud storage systems has been relatively mature, but the cloud storage system with single encryption mechanism is unable to meet the users' new demands. The cloud storage system with various encryption mechanisms becomes the research hotspot [5, 8, 12, 13, 20, 24]. Because of the complexity of network environment, there still are some privacy disclosure problems in the cloud storage scheme with hybrid encryption mechanisms including the leak of attribute information, identity information disclosure, and data breaches in the cloud.

1.2 Related Works

Researchers have proposed a number of ciphertext-policy attribute-based encryption (CP-ABE) cloud storage schemes and some privacy protection measures. For the problem of content privacy disclosure, the measures usually adopted are data segmentation and encryption [2]. Besides, the secure revocation mechanism [7] can also ensure that data will not be stolen by illegal users. Attribute privacy preserving is mainly carried out in two aspects: access structure and users' attribute set. By hiding the access strategy [28] and solving the collusion problem the attribute privacy can be well protected. With the deepening of research, the function of cloud storage scheme with attribute-based encryption (ABE) is more perfected, but the privacy disclosure problems still exist in the current ABE cloud storage schemes. The single key generator may lead to users' attribute set leakage and bring vulnerability to the system. For instance, the scheme in literature [29] outsources the calculations of encryption, decryption and key generation to the cloud server, which greatly reduces the computing and communication overhead. Feng *et al.* proposed a decentralized ciphertext-policy attribute based encryption (CP-ABE) scheme in literature [4] to avoid the system vulnerabil-

ity caused by central authority. In literature [19], Li *et al.* also proposed a CP-ABE cloud storage scheme with multiple attribute authorities which outsourcing the bilinear pairing operation to the cloud server. Meanwhile the group keys were introduced into the attribute authority to realize the efficient and fine-grained revocation mechanism. In addition, the multiple attribute authorities (AAs) may recover users' attribute set though collusion, resulting in the disclosure of users' attribute information. Aiming at the collusion problems, Jung *et al.* proposed an attribute-based encryption scheme in [15] by improving the anonymous scheme in [14] and introduced the 1-out-of-n Oblivious Transfer into the multi-authority attribute-based encryption scheme preventing the collusion between the attribute authorities. In addition, users may collude to obtain the private key and decrypt the data without permission, causing the leakage of content privacy in practical applications. Guan *et al.* [6] introduced the attribute management server (AMS) into the scheme to assign the attribute authority for users according to users' attribute set. Attribute name was used to interact and the attribute value was hid in the scheme which still has the privacy disclosure problems caused by collusion. The above schemes have realized the access control to the data, but the privacy protection mechanism is not complete. Besides, the cloud system with single attribute-based encryption mechanism can't realize the search operation of ciphertext.

Searchable encryption is very suitable for ciphertext search environment of the cloud storage and its application prospect is very broad [9,11,16,18]. Users can search and update the data files stored in the cloud though cloud storage system based on searchable encrypted. Wang *et al.* proposed a mixed index structure in [23]. In the scheme, the static index (SI) and dynamic index (DI) were used in the first-time searches and repeated searches respectively, which reduced the complexity of the search operation. Moreover, the scheme also achieved the function of ciphertext updating by means of the dynamic index. The third-party is permitted to obtain the keyword search trapdoor to perform the ciphertext search operation in the public-key encryption with keyword search (PEKS) mechanism. However, in this data sharing mechanism, there still are the privacy disclosure problems brought by the keyword guessing attack that can't be ignored. Xu *et al.* [25] presented a public-key encryption with fuzzy keyword search scheme which can against the keyword guessing attack. The scheme transformed from the anonymous identity encryption scheme. And the mechanism that many keywords sharing one fuzzy search trapdoor solves the problem of privacy leakage caused by third-party stealing keywords. Fuzzy search trapdoor was sent to the untrusted server for ciphertext matching and filtering, and the exact trapdoor was used for local secondary filtering to get the matched ciphertext. With the method of authenticating the keywords, Huang *et al.* prevented the untrusted server from recovering the keywords by keyword guessing attack in literature [10]. The literature [25] and literature [10] have both realized the public-key encryption with keyword search scheme which can resist keyword guessing attack. The researchers established the users' privacy protection mechanism and solved the problem of users' privacy disclosure problem caused by keywords leakage. Researches on cloud storage system based on searchable encryption have been relatively mature, but there still are some security risks in the searchable encryption system as lacking of fine-grained access control.

Secure data sharing in complex network environments requires not only a complete privacy protection mechanism but also efficient and robust system functionalities. In order to achieve both access control and ciphertext search operations, the searchable encryption technology is introduced into the current attribute-based encryption cloud storage system. The researchers have proposed the cloud storage sharing mechanism with multiple encryption technologies. In [26], Yang *et al.* achieved fine-grained access control over searchable encryption schemes through ciphertext-policy attribute-based encryptio. They also achieved the concealment of keywords and the direct revocation of users. But the single key generator may lead to the leakage of users' attribute set. Once the generator is breached, it will bring the inevitable damage to the system. In addition, the access structure is uploaded to the semi-trusted cloud server in the scheme, which may cause the problems of sensitive attribute information disclosure. Wang *et al.* also used the multiple encryption mechanism in [21] and proposed a multi-user, fine-grained searchable encryption scheme, which adopted the hybrid cloud structure. In the structure, public cloud was used to achieve access control and ciphertext search operations. The security of cloud storage services was guaranteed by the re-encryption calculation of private cloud. But the scheme requires users to submit the private key to the cloud server for access control, which inevitably increased the risk of privacy disclosure. The other problem of the scheme is that the trapdoor generation process lacks privacy protection mechanism for keywords. In [17], Li *et al.* proposed a searchable ciphertext-policy attribute-based encryption scheme, in which fine-grained attribute revocation was realized via the version number and the access structure was hid. Meanwhile, the ciphertext search operation was achieved in the scheme and the computation of ciphertext updating was decreased by using homomorphic encryption. But the functions including key updating and re-encryption calculating were performed with a single authority, which brought the system inevitable vulnerability.

1.3 Our Contribution

In order to improve the security and practicability of the existing cloud storage schemes, this paper proposes a searchable ciphertext-policy attribute-based encryption privacy preserving scheme. The capability of cloud storage system is extended with the PEKS and CP-ABE. And the privacy disclosure problems in the current hy-

brid cloud storage systems are solved by optimizing the algorithm and improving the system structure.

1) This scheme solves the problems that the ciphertext can't be searched in attribute-based encryption schemes by introducing the PEKS into the multi-authority CP-ABE cloud storage schemes. Moreover, the scheme adopts the direct revocation to realize the revocation mechanism of users' searching rights.

2) The authority in the proposed scheme is composed of two parts, central authority and attribute authorities. The access control of the users' searching permissions is accomplished by the central authority. To prevent malicious user colluding with each other, the random identity token (RID) of users is introduced into the calculation of privacy key. What's more, the 1-out-of-n Obvious Transfer algorithm is used in the process of request and distribution so that to avoid the collusion caused by the attribute authorities.

3) The security of algorithms generating ciphertext of keywords and trapdoors is improved though the random numbers and user key (UK), which protects the privacy of keyword in the process of searching. Finally, we prove the security of proposed scheme based on the decisional bilinear Diffie–Hellman assumption in the random oracle model and analysis the performance of the cloud storage scheme.

The rest of this paper is arranged as follows. In Section 2 are some preliminaries related to the proposed scheme. The system model and threats model are presented in Section 3. The specific algorithm of searchable CP-ABE privacy preserving scheme are all given in Section 4. The security of proposed scheme is proved in Section 5. The analysis of privacy preserving and performance are described in Sections 6. Finally, conclusion and prospect are in Section 7.

2 Preliminaries

In this section, we introduce some definitions related to our schemes.

2.1 Bilinear Map

Definition 1. *Let G_1 and G_2 be two groups of prime order p and the generator of G_1 is g. The finite field of prime order p is defined as Z_p, the set of integers $\{0, 1, \ldots, , p-1\}$. A bilinear mapping $e : G_1 \times G_1 \to G_2$ satisfies the following properties:*

- *Bilinearity: For any $u, v \in G_1, a, b \in Z_p$, it has $e(u^a, v^b) = e(u, v)^{ab}$.*
- *Non-degeneracy: there exists $u, v \in G_1$ such that $e(g, g) \neq 1$.*
- *Computability: For any $u, v \in G_1$, there is an efficient bilinear mapping computation $e(u, v)$.*

2.2 Decision Bilinear Diffie–Hellman (DBDH) Assumption

Definition 2. *DBDH problem in group G of prime order p with generator g is defined as follows: let $g^a, g^b, g^c \in G$ and $e(g, g)^{abc} = e(g, g)^z$, and then decide whether $z = abc$ or z is a random number where $a, b, c, z \in Z_p$.*

Definition 3. *The DBDH assumption is that no probabilistic polynomial-time algorithm has a non-negligible advantage in solving the DBDH problem [22].*

2.3 CP-ABE and PEKS

Ciphertext-policy attribute-based encryption (CP-ABE) is a public-key encryption mechanism proposed by Bethencourt *et al.*, which implements fine-grained access control by encrypting data with access structure. In CP-ABE, the ciphertext is related to the access structure, and the users' private keys are associated with their attribute set. Basic algorithms usually include initialization, encryption, key generation and decryption. Ciphertext-policy attribute-based encryption cloud storage model is illustrated in Figure 1.

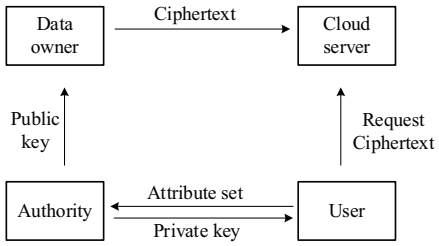


Figure 1: CP-ABE cloud storage mode

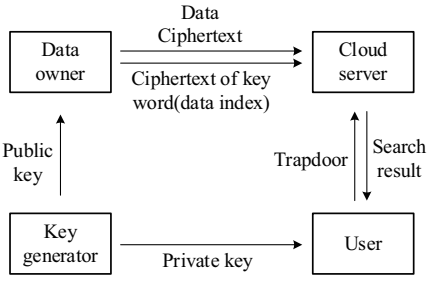


Figure 2: PEKS cloud storage model

Figure 2 shows the cloud storage model based on public key encryption with keyword search (PEKS) which is proposed by Boneh *et al.* in [1]. The keywords extracted from shared data are encrypted with the public key by the data owner, which generates the keyword ciphertext, the data index. The private key is used to encrypt the search keyword to generate the trapdoor by users. The ciphertext search operation is completed by matching the index and the trapdoor. The basic algorithm of PEKS is composed of key generation, encryption, trapdoor generation and test.

3 System Model

In this section, we will describe the basic structure as well as threats model of the proposed scheme.

3.1 Basic Structure

We propose a searchable CP-ABE privacy preserving scheme shown in Fig. 3. There are five participating entities in our scheme including data owner, users, cloud server, N attribute authorities and the central authority.

1) Data owner: In the stage of setup, data owner generates a key pair including index key and trapdoor key to encrypt the keyword. Then, data owner extracts the keywords from the shared data and encrypts the keywords as data index with the index key. The access structure is formulated for encrypting the trapdoor key and shared data. The calculated verification (VR) and ciphertext of trapdoor key are sent to central authority, while the data index and ciphertext are transmitted to the cloud server.

2) User: Users need to register themselves to get the user key (UK) and the random identity token (RID). By asking attribute authorities for private key, users can decrypt ciphertext of trapdoor key and calculate verification of user (VR') submitted to the central authority for permissions validation. And then, users encrypt the search keywords with the trapdoor key and UK, which generates trapdoor sent to the cloud server for ciphertext searching. After receiving the matched ciphertext, users can recover the shared data.

3) Cloud server: The storage and searching of data ciphertext are executed by the cloud server. Taking the UK, data index, and trapdoor as input, cloud server matches the data ciphtertext and returns the result to users. In the revocation phase, it also needs to achieve the direct revocation of users.

4) Central authority: In the register phase, UK and RID are generated by central authority (CA) which sends the trapdoor key ciphertext to users and verifies users' access permissions. If users pass the validation, UK and RID will be sent to cloud server for ciphertext searching.

5) Attribute authorities: Setup and generate the master key and public key which is sent to data owner for encrypting. The attribute authorities (AAs) respond the request of private key and generate the corresponding private key components after receiving users' attribute set.

3.2 Threats Model

In the proposed scheme, only the central authority is fully trusted. The attribute authorities will honestly generate and distribute the private key for users, but they may collude with each other to steal users' attribute information. The cloud server is honest but curious. It will abide by the protocol returning the searched ciphertext to users and revoke users' search permissions. However the semi-trusted cloud server may steal the shared data and information of users internally. Users may collude to decrypt the data that they don't have permissions.

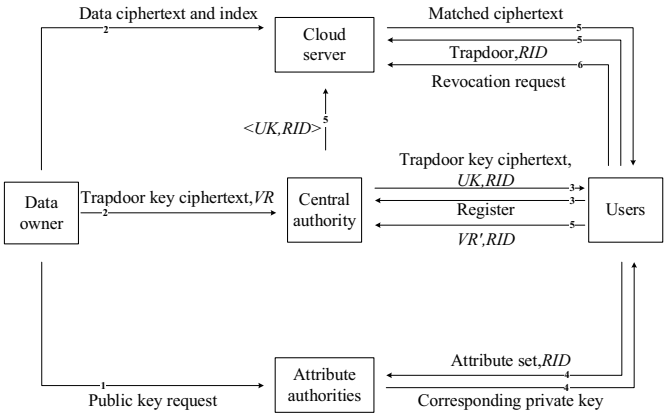


Figure 3: A searchable CP-ABE privacy preserving scheme

4 Concrete Algorithm

In this section, we will give the specific algorithm construction of searchable CP-ABE privacy protection scheme.

4.1 Setup

1) AAs setup $\alpha_i \rightarrow (PK, MK)$. Any one of attribute authorities chooses a bilinear group G of prime order p with generator g. AA_i chooses $\alpha_i \in Z_p$ and calculates $Y_i = e(g,g)^{\alpha_i}$ sent to the rest of AAs which all need to calculate $Y = \prod_{i \in AA} Y_i = e(g,g)^{\sum \alpha_i}$. The public key of system is $PK = \{G, g, Y = e(g,g)^{\sum \alpha_i}\}$. Each attribute authority randomly chooses $N-1$ integers $S_{ij} \in Z_p, (j \in \{1,2,\ldots N\} \backslash \{i\})$ and calculates $g^{S_{ij}}$ sent to all other authorities. After receiving the $N-1$ $g^{S_{ji}}$, AAs all calculate parameter x_i follows which belong to Z_p and satisfy $\prod_{i \in AA} x_i = 1 \bmod p$.

$$x_i = \left(\prod_{j \in \{1,\ldots,N\} \backslash \{i\}} g^{S_{ij}} \right) / \left(\prod_{j \in \{1,\ldots,N\} \backslash \{i\}} g^{S_{ji}} \right) = g^{\left(\sum_{j \in \{1,\ldots,N\} \backslash \{i\}} g^{S_{ij}} - \sum_{j \in \{1,\ldots,N\} \backslash \{i\}} g^{S_{ji}} \right)} \quad (1)$$

In the initialization phase, the generated master key is $MK = \{\alpha_i, x_i\}$. Each AA_i chooses random number $\gamma_i \in Z_p$ and calculates $x_i \cdot g^{\gamma_i}$ using for generating private key. The $x_i \cdot g^{\alpha_i} \cdot g^{\gamma_i}$ generated by AA_i is shared with the other attribute authorities and any

one of them calculates $Y' = \prod x_i g^{\alpha_i} g^{\gamma_i} = g^{\sum \alpha_i + \sum \gamma_i}$ sent to data owner.

2) Data owner setup. Choose a bilinear group G_1 of prime order p with generator g_1. Let $H_1 : \{0,1\}* \to G_1$ be the hash function. With choosing the random number η, μ, data owner calculates $PK = \{g_1, g_1^{\eta}\}$ and $SK = \eta$, which called index key (IK) and trapdoor key (TK) respectively in the scheme.

4.2 IndexGen $(W, IK) \to I_W$

Data owner extracts the keyword from shared data and encrypts the key words with index key and random numbers τ, μ. The index of shared data is calculated as follows:

$$I_W = (I_1, I_2) = (g_1^{\mu\tau}, e(H_1(W)^{\mu}, g_1^{\eta\tau})) \tag{2}$$

4.3 Encrypt $(M, TK, T_p, PK) \to C_M, C_{TK}, VR$

First, the algorithm choose a polynomial q_x for each node x in $\{T_p\}_{p\in\{0,\ldots,r-1\}}$. The degree d_x of polynomial q_x should less than the threshold value k_x. Starting from the root node R_p, the algorithm randomly picks $S_0 \in Z_p$ and sets $q_{R_p}(0) = S_0$ and the other coefficients of q_{R_p} are picked randomly. The attribute set in access tree is defined as A_{T_p}. With picking a random element $h \in Z_p$, the ciphertext is created as:

$$\begin{aligned} C_{TK} = \big\langle &\{T_p\}_{p\in\{0,\ldots,r-1\}}, E_0 = TK \cdot Y^{S_0}, C = g^{hS_0}, \\ &\hat{C} = (Y')^{h^{-1}}, \{C_i = g^{q_i(0)}, C_i' = H(att(i))^{q_i(0)}\} \\ &i \in A_{T_p}, p \in \{0,\ldots,r-1\}\big\rangle \\ C_M = \langle &E_1 = M \cdot Y^{S_0}, C = g^{hS_0}, \hat{C} = (Y')^{h^{-1}}\rangle \end{aligned} \tag{3}$$

And the verification used to verify the privilege of users is computed as $VR = \{Y^{S_0}\}_{p\in\{1,\ldots,r-1\}}$.

4.4 Enroll $\zeta_i \to UK, RID$

This algorithm enrolls the users who want to join the system and picks the user key (UK) randomly and generates a random sequence as the random identity of users (RID).

4.5 TrapdoorGen $(W', TK, UK, \lambda) \to T_{W'}$

In this algorithm, the random number λ is picked and the calculation of trapdoor is

$$T_{W'} = (T_1, T_2) = (\lambda \cdot UK, H_1(W')^{\lambda \cdot TK}) \tag{4}$$

4.6 Test $(RID, I_W, T_{W'}) \to \{C_M\}$

According to the RID submitted by users and the corresponding UK this algorithm performs the matching calculation $e(T_2, I_1^{UK}) = I_2^{T_1}$ like follows:

$$e(H_1(W)^{\lambda \cdot TK}, g_1^{\mu \cdot \tau \cdot UK}) = e(H_1(W')^{\mu}, g_1^{\eta \cdot \tau})^{UK \cdot \lambda} \tag{5}$$

If users' search keywords are same to those included in the index, the equation will be established. The cloud server sets $result = \{C_M\}$ and returns result to users. If not, the cloud server sets $result = \varnothing$ returned to users.

4.7 KeyGen $\{RID, PK, MK, A_u\} \to SK_{RID}$

For any attribute $k \in A_u$ every AA_i picks a random number $\beta_{RID,k} \in Z_p$ and calculates the private key components $H(att(k))^{\beta_{RID,k}}, D_k' = g^{\beta_{RID,k}}$ sent to user with $x_i \cdot g^{\gamma_i}$ where A_u is the attribute set of user. User calculates as:

$$\begin{aligned} D_k &= H(att(k))^{\beta_{RID,k}} \cdot \prod (x_i \cdot g^{\gamma_i}) \\ &= H(att(k))^{\beta_{RID,k}} \cdot g^{\sum \gamma_i} \end{aligned} \tag{6}$$

By combining the private key components, users can get the private key as $SK_{RID} = \{D_k, D_k' = g^{\beta_{RID,k}}\}$.

4.8 Decrypt $(C_M, C_{TK}, SK) \to (M, TK, VR')$

By calling this algorithm recursively, the TK and validation of user (VR') can be calculated.

1) If the node x is a leaf node and its attribute is i, the algorithm defined as follows.

If $i \in A_u$:

$$\begin{aligned} &DecryptNode(CT, SK, x) \\ &= \frac{e(D_k, C_x)}{e(D_k', C_x')} \\ &= \frac{e(H(att(i))^{\beta_{RID,k}} \cdot g^{\sum \gamma_i}, g^{q_x(0)})}{e(g^{\beta_{RID,k}}, H(att(i))^{q_x(0)})} \\ &= e(g,g)^{(\sum \gamma_i) \cdot q_x(0)} \end{aligned} \tag{7}$$

If $i \notin A_u$, the algorithm return $\varnothing$.

$$\begin{aligned} F_x &= \prod F_z^{\Delta_{index(z), S_x'}(0)} \\ &= \prod (e(g,g)^{(\sum r_i) \cdot q_z(0)})^{\Delta_{index(z), S_x'}(0)} \\ &= \prod (e(g,g)^{(\sum r_i) \cdot q_{parrent(z)}(index(z))})^{\Delta_{index(z), S_x'}(0)} \\ &= \prod (e(g,g)^{(\sum r_i) \cdot q_x(index(z))})^{\Delta_{index(z), S_x'}(0)} \\ &= e(g,g)^{(\sum \gamma_i) \cdot q_x(0)} \end{aligned} \tag{8}$$

2) If x is not a leaf node, the nodes z, children nodes of x, call $DecryptNode(CT, SK, z)$ and write the outputs as F_z. Let S_x be an arbitrary k_x -sized set of child nodes z with the index S_x'. By using polynomial interpolation the calculation is as follows.

After getting the ciphertext of trapdoor key, users call the decryption algorithm recursively starting from the root node R_p and calculate the verification of user (VR') as follows.

$$DecryptNode(C_{TK}, SK, R_p) = e(g,g)^{S_0 \sum \gamma_i} \tag{9}$$

If users' attribute set meet the access tree, they can decrypt the ciphertext of trapdoor key as:

$$\frac{E_0}{\frac{e(C,\hat{C})}{e(g,g)^{S_0\sum\gamma_i}}} = \frac{TK\cdot Y^{S_0}}{\frac{e(g^{hS_0},(g^{\sum\alpha_i+\sum\gamma_i})^{h^{-1}})}{e(g,g)^{S_0\sum\gamma_i}}} = \frac{TK\cdot e(g,g)^{(\sum\alpha_i)\cdot S_0}}{e(g,g)^{S_0\sum\alpha_i}} = TK \quad (10)$$

The shared data can be recovered as

$$M{=}E_1/\frac{e(C,\hat{C})}{VR'} \quad (11)$$

4.9 Revoke

By generating the list of *UK* and *RID*, the direct revocation to users' search permission can be achieved. In the phase of revocation, users submit their *RID* and then the cloud server remove the corresponding item of *UK* and *RID* from the list. If cloud server can't find the corresponding *UK* in the process of ciphertext matching, the ciphertext search operation is terminated and then the cloud server returns the information that authentication fails.

5 Security Proof

In this section, the security of proposed scheme is proved in the random oracle model.

Lemma 5.1. *Based on DBDH assumption, if the scheme in [15] is secure against chosen plaintext attacks (CPA) in the random oracle model, our scheme is secure against CPA.*

Proof. Suppose there exists a probabilistic polynomial time adversary A can attack our scheme with advantage ε. We prove that the following DBDH game can be solved by the challenger B with advantage $\frac{\varepsilon}{2}$.

Let $e: G\times G\to G_0$ be a bilinear map where G is a cyclic group of prime order p with generator g. First, the challenger B randomly picks $a,b,c,z\in Z_p, \theta\in\{0,1\}$ and sets tuple $(g,A,B,C,Z)=(g,g^a,g^b,g^c,e(g,g)^{abc})$ if θ=0. Otherwise if θ=1, the tuple is set to $\left(g,g^a,g^b,g^c,e(g,g)^z\right)$.

Initialization: The adversary controls part of attribute authorities where at least two authorities are not controlled by the adversary and remaining authorities are controlled by challenger B. The adversary A declares the challenged access tree T_0' of which some attributes are managed by the simulator's authorities.

Setup: The challenger sets $a=\sum\gamma_i, b=\frac{\sum\alpha_i}{\sum\gamma_i}, c=s_0$ where $\gamma_i,\alpha_i,s_0\in Z_p$ are randomly picked and gives Y and Y' to the adversary.

Query Phase 1: The adversary queries for the private keys according to attribute set and none of the attribute set satisfy the access tree. After receiving the private key queries from A with *RID*, the challenger randomly picks $\beta_{RID,k}\in Z_p$ and calculates private key components for every attribute $k\in A_u$ as follows: $D_k=H(att(k))^{\beta_{RID,k}}\cdot g^{\sum\gamma_i}, D_k'=g^{\beta_{RID,k}}$.

Query Phase 2: Repeat Phase 1 adaptively.

Guess: The adversary A submits the guess θ' of θ. When $\theta=\theta'$, the simulator represented challenger B outputs $\left(g,g^a,g^b,g^c,e(g,g)^{abc}\right)$ if θ=0, otherwise it outputs a DBDH tuple $\left(g,g^a,g^b,g^c,e(g,g)^z\right)$ composed by five random elements.

When θ= 1, the adversary A can't get any useful information and the advantage is $\Pr=\frac{1}{2}$. And the advantage is $\Pr=\frac{1}{2}+\varepsilon$ when θ=0. Therefore, the advantage of probabilistic polynomial time adversary in the DBDH game is $\Pr(\theta'=\theta)-\frac{1}{2}=\frac{1}{2}\left(\frac{1}{2}+\varepsilon\right)+\frac{1}{2}\cdot\frac{1}{2}-\frac{1}{2}=\frac{\varepsilon}{2}$. To conclude, if the adversary has non-negligible advantage in the constructed game, he can solve the DBDH problem with the non-negligible advantage $\frac{\varepsilon}{2}$. Based on the DBDH assumption, there is no adversary has significant advantage in our security game and our scheme is secure. □

Lemma 5.2. *If DBDH assumption is tenable and the scheme in [10] is semantically secure in the random oracle model, our scheme is semantically secure in the random oracle model.*

Proof. Assume the probabilistic polynomial time (PPT) adversary A can break our scheme with non-negligible advantage and then we construct a PPT algorithm B simulating the oracle to solve the DBDH problem. The adversary issues at most q_H, q_T, q_C to the hash oracle, trapdoor oracle, ciphertext oracle respectively.

Hash Oracle: The adversary queries the hash H_1 and gives a keyword W_i. The simulator B randomly picks $a_i\in Z_p$ and flip a random coin c_i such that $\Pr[c_i=0]=\delta$. If $c_i=0$, the simulator B calculates $g_1^{\frac{z}{\mu}}\cdot g_1^{a_i}=h_i$, and sets $g_1^{a_i}=h_i$ if $c_i=1$. Then, B adds tuple $[W_i,h_i,a_i,c_i]$ to the list L_{H_1} and sets $H_1(W_i)=h_i$ as the hash value of the keyword W_i. The hash value h_i is returned to the adversary A.

Trapdoor Oracle: Given a keyword W_i, the simulator retrieves tuple $[W_i,h_i,a_i,c_i]$ in the list L_{H_1}. if $c_i=0$, the simulator B aborts and out puts the guess b' of b. If $c_i=1$, B randomly chooses $\beta_i,\rho\in Z_p$ and calculates the true trapdoor as $T_i=(T_1,T_2)=(\beta_i\cdot\rho, H_1(W_i)^{\beta_i\cdot\rho})$. And then B return the trapdoor to adversary A.

Ciphertext Oracle: Given a keyword W_i, the simulator retrieves tuple $[W_i,h_i,a_i,c_i]$ in the list L_{H_1}. If $c_i=0$, the simulator B aborts and outputs the guess b' of b. If $c_i=1$, B randomly chooses $\eta,\tau,\mu\in Z_p$

and calculates the true ciphertext $C_i = (C_1, C_2) = (g_1^{\mu\tau}, e(H_1(W_i)^{\mu}, g_1^{\eta\tau}))$ returned to the adversary A.

Challenge: The adversary chooses the keyword W_0W_1 that he wants to challenge. B performs the above algorithm and retrieves the tuples $[W_0, h_0, a_0, c_0]$ and $[W_1, h_1, a_1, c_1]$. If $c_0 = 1$ and $c_1 = 1$, the simulator B aborts and outputs the guess b' of b. If $c_0 = 0$ or $c_1 = 0$, let $\hat{b}$ be the bit such that $c_{\hat{b}} = 0$ and we have $h_{\hat{b}} = g_1^{\frac{z}{\mu}} \cdot g_1^{a_{\hat{b}}}$.

The simulator B calculates $C_2 = Z \cdot e(g_1^{\eta}, g_1^{\tau})^{\mu a_{\hat{b}}}$. If $Z = e(g_1, g_1)^{\eta\tau z}$, then $C_2 = e(g_1, g_1)^{\eta\tau(\mu a_{\hat{b}} + z)} = e(h_{\hat{b}'}^{\mu}, g^{\eta\tau}), C_1 = g_1^{\mu\tau}$. B returns $C = (C_1, C_2)$ to A. The adversary continues to query for W_i where the only restriction is $W_i \neq W_0$,W_1. At last, the adversary submits a guess $\hat{b}'$ of $\hat{b}$. If $\hat{b}' = \hat{b}$, the simulator B outputs $b' = 0$. And it outputs $b' = 1$ if $\hat{b}' \neq \hat{b}$.

The probability that the simulator B doesn't aborts is $\Pr[B] = (1-\delta)^{q_T+q_C}(1-(1-\delta)^2)$. It's no-negligible because it approximately equals to $\frac{2}{(q_T+q_C)e}$. If the adversary can break the algorithm of our scheme, the simulator B can succeed in distinguish that Z is equal to $e(g_1, g_1)^{\eta\tau z}$ or a random element. The probability that simulator B succeeds in guessing b' of b is $\Pr[b' = b] = \frac{1}{2} + \varepsilon \cdot \Pr[B]$. If ε is no-negligible, so is $Pr[b' = b] - \frac{1}{2}$, the advantages of solving the DBDH problems by simulator B. based on the DBDH assumption, there is no adversary can break our algorithm with no-negligible advantage and our scheme is safe. □

Theorem 5.3. *If DBDH assumption is tenable, our scheme is safe in the random oracle model.*

Proof. Directly derived from Lemma 5.1 and Lemma 5.2. □

6 Analysis and Comparison

6.1 Privacy Preserving Analysis

6.1.1 Content Privacy

This paper adopts the CP-ABE algorithm, a public-key encryption mechanism, to encrypt the shared data, which is safer than the symmetrical encryption. By encrypting the shared data with access tree, we ensure that the safety of content privacy of data owner. Besides, the direct revocation mechanism solves the privacy disclosure problems caused by private key mismanagement. Furthermore, the random number $\beta_{RID,k}$ is introduced into the process of private key generating. The components of private key are related to *RID* which is a random sequence, an interactive identity of user. Even the different users collude with each other they can't get the private key that they don't have the permissions. Thus, the illegal user can't search and get the shared data though the collusion.

6.1.2 Identity privacy

The central authority is introduced to the multi-authority CP-ABE scheme, but the central authority in this paper doesn't participate in the process related to attributes. On the one hand, the central authority stores the ciphertext of trapdoor key so that the data owner doesn't need to be always online. On the other hand, CA registers users and randomly generates the user key (UK) and the random identity (RID) for each user. The random sequence RID replaces the user's identity in course of the interaction, which protects the identity privacy of users. Therefore, this mechanism realizes the bidirectional anonymous interaction.

6.1.3 Search privacy

The search mechanism of our scheme can against multiple attacks. By encrypting the hash value of keyword with random number μ maintained only by data owner in the process of index generating, the cloud server can't make the keyword guessing attack internally by matching the candidate keyword with trapdoor. In the stage of trapdoor generating, we hide the search keyword with the random number, which prevents the keyword replay attack executed by malicious attacker after intercepting the trapdoor. Hence, the semi-trusted cloud server and attacker can't obtain any useful information of the keyword and our scheme achieves the privacy preserving for the keyword without reducing security of previous algorithm.

6.1.4 Attribute privacy

Data owner: The fine-grained access control is achieved by the central authority. Users' search privilege is verified by the central authority though the validation (VR), which avoiding the risk brought by submitting access structure to the semi-trusted cloud server. This mechanism protects the attribute of access tree created by the data owner.

Users: Our scheme solves the privacy disclosure problems caused by the collusion of attribute authorities in the multi-authority schemes. The anonymous transfer algorithm is adopted in the interactive process of private key generating as is shown in Fig 4. We assign the attribute to AAs by category, so each attribute authority only manages one kind of attribute. Each user has one value of the attributes controlled by each attribute authority. After receiving private key request of each user, all the attribute authorities compute components of private key for every attribute value. With the anonymous transfer algorithm, attribute authority can't know the components that users choose so that the attributes of users won't be leaked, which protects users' attribute information.

Table 1: The comparison with the classic schemes

scheme	access control	ciphertext search	index security	trapdoor security	access structure security	multi-authority	against AAs' collusion	against users' collusion
[19]	√	—	—	—	×	√	×	√
[15]	√	—	—	—	×	√	√	×
[10]	—	√	√	×	—	—	—	—
[26]	√	√	×	√	×	×	—	×
[21]	√	√	×	×	×	×	—	×
our	√	√	√	√	√	√	√	√

Table 2: The comparison of performance

scheme	Setup	Encrypt	IndexGen	TrapdoorGen	Test	KeyGen	Decrypt
[26]	$O(1)$	$O(X)$	$3E+H+P$	$E+H$	$2E+P$	$O(1)$	$O(X)$
[21]	$O(1)$	$O(XI)$	$2E+H+P$	$2E+P+H$	P	$O(K)$	$O(1)$
our	$O(N^2+1)$	$O(2XI)$	$3E+H+P$	$E+H$	$2E+P$	$O(K)$	$O(X)$

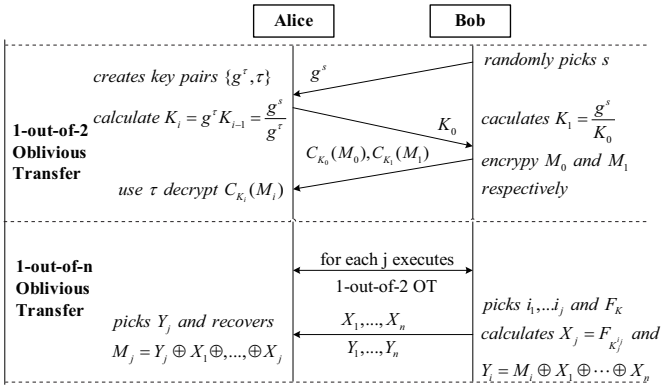


Figure 4: Anonymous transfer algorithm

6.1.5 Scheme Comparison

We compared the proposed scheme with some of classic cloud storage schemes, as is shown in TABLE I. Among the schemes, the literature [15,19] are the single attribute-based encryption cloud storage scheme of which the ciphertext isn't searchable. The shared data is encrypted by symmetric encryption algorithm in [15,21] and the security is insufficient. The scheme in [10] is based on the single public key encryption with keyword search, which lacks of the fine-grained access control. In [21, 26], the multiple encryption algorithms are adopted to achieve access control and ciphertext search function, but the single authority increases the potential vulnerability to the system. Some privacy disclosure problems still exist in the current cloud storage schemes. It can be seen from Table 1 that the scheme proposed in this paper not only realizes the access control and ciphertext search operation of the cloud storage system, but also establishes the relatively perfect privacy preserving mechanism for the hybrid cloud storage system. Our scheme protects the privacy of the keyword in the search process, and solves the collusion problem of the multi-authority mechanism. The mechanism that access control accomplished by the trapdoor key and CA also ensures the security of attribute privacy in the access structure.

6.2 Performance Analysis

In this section, we analyze the performance of the proposed scheme. We denote X as the number of nodes in the access tree and I as the average threshold value. The size of attribute set of users is denoted by K and the number of attribute authority is denoted by N. In the initialization phase of this scheme, the time complexity of the algorithm performed by each attribute authority is $O(1)$. The time complexity of the setup computation is $O(N^2+1)$. There are X nodes in access tree and the average threshold value is I, the complexity of the encryption is $O(2XI)$. In the stage of key generating phase the complexity of N attribute authorities is $O(N^2+N\cdot K)$. Users' private key is composed by K components and the complexity is $O(K)$. Because of the 1-out-of-n transfer algorithm, the communication overhand is increased to $O(K)$. The algorithm of decryption is recursive, which executed at all nodes of access tree, so the complexity is $O(X)$. The computational overhead of search mechanism in this paper is denoted by exponentiation (E), hash function (H) and bilinear pairing (P). In the index generation phase, the overhead is $3E+H+P$ and the cost of trapdoor calculating is $E+H$. Finally, the computational expense of test algorithm is $2E+P$. It can be seen from Table 2 that although we enhance the safety of the system by improving the system structure and algorithm, but the complexity of calculation and communication is increased. In addition, the computing efficiency of search mechanism isn't improved despite the improvement of keyword privacy preserving mechanism in this paper.

7 Conclusion

With the rapid development of Internet technology, cloud storage system centered on data management and sharing has received more and more attention. The cloud storage schemes based on CP-ABE can be used in many files like electronic healthcare [3], Internet of Things [27], and so on. But this model that the data stored by third-party brings the new security risks. The shared content, identity, attributes and other privacy information of users may be disclosed in the use of cloud storage system. Establishing a complete privacy protection mechanism has become an important factor in the development and promotion of cloud storage systems.

Aiming at the privacy disclosure problems caused by submitting the access structure and identities, collusion, and the attacks about keywords, this paper proposes a searchable CP-ABE privacy preserving scheme. The scheme can accomplish the access control and ciphertext search at once and establish a relatively complete privacy protection mechanism for the cloud storage system with hybrid encryption. We introduce the central authority to achieve the access control of users, which protects the attributes in access tree. The problems of collusion and keyword leakage are solved by introducing the anonymous transfer algorithm and improving the original algorithms. The scheme is proved based on the DBDH assumption. Analysis and comparison show that the proposed scheme is more secure and practical.

By analyzing the efficiency of the system, it can be found that that however the privacy protection mechanism adopted in this paper improves the security of the system, but the overhead of computation and communication is still very large. The complexity of the encryption and decryption algorithm increases with the increasing number of attributes and the efficiency of search mechanism needs to be improved. The establishment of safe and efficient Cloud storage system is the key point of our further research. In addition, the revocation in this paper is coarse-grained and it's necessary to achieve the attribute-level user revocation for the hybrid encryption cloud storage scheme in the future.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (No.61462060, No.61562059), Regional Science Foundation Project (No. 61762060), Youth Science and Technology Fund Program of Gansu (No. 1610RJYA008). The authors gratefully acknowledge the anonymous reviewers for their valuable comments.

References

[1] D. Boneh, G. D. Crescenzo, R. Ostrovsky, and G. Persiano, "Public key encryption with keyword search," in *International Conference on the Theory and Applications of Cryptographic Techniques*, pp. 506–522, 2004.

[2] P. S. Chung, C. W. Liu, and M. S. Hwang, "A study of attribute-based proxy re-encryption scheme in cloud environments", *International Journal of Network Security*, vol. 16, no. 1, pp. 1-13, 2014.

[3] S. Divyashikha, S. Huzur, and G. Daya, "CP-ABE for selective access with scalable revocation: A case study for mobile-based healthfolder," *International Journal of Network Security*, vol. 20, no. 44, pp. 689-701, 2018.

[4] T. Feng and J. Guo, "A new access control system based on CP-ABE in named data networkin," *International Journal of Network Security*, vol. 20, pp. 710–720, 2018.

[5] T. Feng and X. Yin, "Research on privacy preserving mechanism of attribute-based encryption cloud storage," *Chinese Journal of Network and Information Security*, vol. 2, no. 7, pp. 8–17, 2016.

[6] Z. T. Guan, T. T. Yang, R. Z. Xu, and Z. X. Wang, "Multi-authority attribute-based encryption access control model for cloud storage," *Journal on Communications*, vol. 36, no. 6, pp. 116–126, 2015.

[7] H. Hong and Z. Sun, "High efficient key-insulated attribute based encryption scheme without bilinear pairing operations," *SpringerPlus*, vol. 5, no. 1, pp. 1–12, 2016.

[8] W. 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.

[9] S. T. Hsu, C. C. Yang, and M. S. Hwang, "A study of public key encryption with keyword search", *International Journal of Network Security*, vol. 15, no. 2, pp. 71–79, Mar. 2013.

[10] Q. Huang and H. Li, "An efficient public-key searchable encryption scheme secure against inside keyword guessing attacks," *Information Sciences*, vol. 403, pp. 1–14, 2017.

[11] M. S. Hwang, S. T. Hsu, C. C. Lee, "A new public key encryption with conjunctive field keyword search scheme," *Information Technology and Control*, vol. 43, no. 3, pp. 277–288, Sep. 2014.

[12] M. S. Hwang, C. C. Lee, T. H. Sun, "Data error locations reported by public auditing in cloud storage service," *Automated Software Engineering*, vol. 21, no. 3, pp. 373–390, Sep. 2014.

[13] M. S. Hwang, T. H. Sun, C. C. Lee, "Achieving dynamic data guarantee and data confidentiality of public auditing in cloud storage service," *Journal of Circuits, Systems, and Computers*, vol. 26, no. 5, 2017.

[14] T. Jung, X. Y. Li, Z. Wan, and M. Wan, "Privacy preserving cloud data access with multi-authorities," in *Proceedings IEEE INFOCOM*, pp. 2625–2633, 2013.

[15] T. Jung, X. Y. Li, Z. Wan, and M. Wan, "Control cloud data access privilege and anonymity with fully anonymous attribute-based encryption," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 1, pp. 190–199, 2015.

[16] C. C. Lee, S. T. Hsu, M. S. Hwang, "A study of conjunctive keyword searchable schemes," *International Journal of Network Security*, vol. 15, no. 5, pp. 311–320, 2013.

[17] J. Li, Y. Shi, and Y. Zhang, "Searchable ciphertext-policy attribute-based encryption with revocation in cloud storage," *International Journal of Communication Systems*, vol. 30, no. 1, 2017.

[18] J. W. Li, C. F. Jia, Z. L. Liu, J. Li, and M. Li, "Survey on the searchable encryption," *Journal of Software*, vol. 26, no. 1, pp. 109–128, 2015.

[19] Q. Li, J. Ma, R. Li, X. Liu, J. Xiong, and D. Chen, "Secure, efficient and revocable multi-authority access control system in cloud storage," *Computers & Security*, vol. 59, pp. 45–59, 2016.

[20] 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.

[21] Q. Wang, Y. Zhu, and X. Luo, "Multi-user searchable encryption with coarser-grained access control without key sharing," in *International Conference on Cloud Computing and Big Data (CCBD'14)*, pp. 119–125, 2014.

[22] S. Wang, J. Zhou, J. K. Liu, J. Yu, J. Chen, and W. Xie, "An efficient file hierarchy attribute-based encryption scheme in cloud computing," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 6, pp. 1265–1277, 2016.

[23] W. Wang, P. Xu, H. Li, and L. T. Yang, "Secure hybrid-indexed search for high efficiency over keyword searchable ciphertexts," *Future Generation Computer Systems*, vol. 55, pp. 353–361, 2016.

[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] P. Xu, H. Jin, Q. Wu, and W. Wang, "Public-key encryption with fuzzy keyword search: A provably secure scheme under keyword guessing attack," *IEEE Transactions on Computers*, vol. 62, no. 11, pp. 2266–2277, 2013.

[26] Y. Yang and B. G. Lin, "Secure hidden keyword searchable encryption scheme with fine-grained and flexible access control," *Journal on Communications*, vol. 34, no. Z1, pp. 92–100, 2017.

[27] X. Yao, Z. Chen, and Y. Tian, "A lightweight attribute-based encryption scheme for the internet of things," *Future Generation Computer Systems*, vol. 49, pp. 104–112, 2015.

[28] Z. b. Ying, J. F. Ma, and J. T. Cui, "Partially policy hidden CP-ABE supporting dynamic policy updating," *Journal on Communications*, vol. 36, no. 12, pp. 178–189, 2015.

[29] R. Zhang, H. Ma, and Y. Lu, "Fine-grained access control system based on fully outsourced attribute-based encryption," *Journal of Systems and Software*, vol. 125, pp. 344–353, 2017.

Biography

Tao Feng is researcher and doctoral supervisor, CCF senior member, IEEE and ACM member. He received the Ph.D. degrees in Xidian University and then worked at school of computer and communication in Lanzhou University of Technology. His research interests include Cryptography and information security.

Xiaoyu Yin is a graduate student at School of Computer and Communication, Lanzhou University of Technology. Her research interest is Network and information security.

Ye Lu is a doctoral student at college of Electrical and Information Engineering, LanZhou University of Technology. His research interests include information security and industrial control system.

Junli Fang received her Master's degree in Communication and Information System from Beijing JiaoTong University, Beijing, China in 2009. She is a lecturer in the School of Computer and Communication, Lanzhou University of Technology, China. Her research interests include network and information security.

Fenghua Li is a PhD supervisor worked in The State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences. His current research interests include Network Security, System Security & Evaluation and Trusted Computation.

Implementation, Performance and Security Analysis for CryptoBin Algorithm

Ahmed H. Eltengy[1], Samaa M. Shohieb[2], Ali E. Takieldeen[3], and Mohamed S. Ksasy[4]
(Corresponding author: Ahmed H. Eltengy)

Faculty of Engineering, Alexandria University[1]
22 El-Guish Road, El-Shatby, Alexandria 21526, Egypt
(Email: tengy_fox@yahoo.com)
Faculty of Computers and Information, Mansoura University[2]
Mansoura, Egypt
Faculty of Engineering, Delta University, Mansoura, Egypt[3]
Faculty of Engineering, Mansoura University, Mansoura, Egypt[4]
Mansoura, Egypt
(Received Aug. 23, 2018; Revised and Accepted Nov. 13, 2018; First Online Jan. 22, 2019)

Abstract

With the fast evolution of digital data exchange, security is the main concern in today's world. It is important to secure data from uncertified access, security information becomes essential in information and data storage, and transmission over open networks such as the Internet. The traditional algorithms face some drawbacks of little key-space and poor security. This paper proposes a new way to encrypt data on the basis of the binary form which is considered the simplest form of data that is consisted of zero and one. This new system converts the target message into zero and one and then swap the bit value from one to zero and from zero to one by using mathematical equations built on the truth table in which the secret key and the target message are the main elements. This algorithm is characterized by a secret key that has an unlimited length and a sub-secret key added to the system. The use of the sub-secret key helps to generate a different encrypted message every time even if the same secret key, the sub-secret key, and the same plain-text are used, which increases the confidentiality and strength of the system. This system provides all the demands of secrecy and strength to confront the intruders with high efficiency and has high-security analysis such as key space analysis, statistical analysis. For example, if the secret key is chosen as 1 MB length that means a number of trials equal to 28388608 to estimate it, which is considered very large to be adequate to safeguard information and data that is encrypted by the proposed encryption system against any attacks. Therefore, the proposed system can be used to secure any software applications.

Keywords: Encryption; Decryption; Information Security; cryptography

1 Introduction

The science of cryptography is the science of coverage and verification of information. Often referred to as "the study of secret" when data exchanged over the Internet, networks or other media. It's the technique of protecting data and information from non-authorized access [7] by transforming it into a non-readable format, called ciphertext.

Only those who have the secret key of the encryption system can decrypt the encrypted message and return it to a readable format. It includes algorithms, protocols, and methodologies to secure prevent or delay unapproved access to sensitive information and to enable verifiability of every component in the communication. A cryptographic algorithm, which is also known as a cipher, could be the mathematical function or equation used for encryption and decryption [3].

Generally, data decryption process is similar to the data encryption process, but in a reversed way. Encryption/Decryption protects data and information from being hacked by the hacker [4]. Encryption/Decryption is a security system where cipher or encryption algorithms are executed together with a secret key to encrypt/decrypt data so that they are unreadable in the event that they are intercepted [6].

With a dramatic increase in the number of Internet users around the world, the need to protect data, information, and multimedia on the Internet has become a high priority. Most operations in governments, military installations, financial institutions, hospitals, and private companies deal heavily with data that is in the form of an image or multiple media, most encryption algorithms today are based on text-only data [19]. Encryption of Digital Image is a beanch of software encryption and has

become very important to prevent and thwart any attack on them to obtain information without prior authorization. Min-Shiang Hwang [11] proposed a new secure cryptographic system built on the Merkle-Hellman public key cryptographic system (knapsack public-key). This method proposes a new Permutation Combination Algorithm.

Gilhorta and Singh [18] proposed the plaintext is converted to a floating number in a range from 0 to 1 and then this floating number converted to binary code and by using a secret key it is converted to encrypted binary code. Animesh Hazra et al. [8] present a brief review of using DNA as a method of cryptography in real time implementation. Li-Chin Huang and Min-Shiang Hwang [9] proposed a study of data hiding in medical images.

Mohamed Rasslan et al. [15] presented a public model to execute any cryptographic algorithm by way of a parallel- pipelined design. Ali E. Takieldeen et al. [20] suggested a method of cryptography which uses the image as a public key and random integers as a private key which is used to permutate the image. Lihua Liu et al. [13] designed a cryptographic system of private broadcast encryption to encrypt a plaintext or a message for multi recipients and hide the recipients identities. Cheng-Chi Lee and Min-Shiang Hwang [12] designed a new convertible authenticated encryption scheme built on the ElGamal cryptosystem. Said Bouchkaren and Saiida Lazaar [2] proposed some tests concentrate on the randomness of tests and on differential cryptanalysis Managed on the CAES (Cellular automata Encryption System).

In this paper, a modified cryptographic algorithm system based on binary codes (0,1) is designed by using mathematical equations [5]. The main idea of this algorithm is converting 0-bit value to 1-bit value and 1-bit value to 0-bit value by using a mathematical equation depends on the bit values of secret key and target message by using logic functions. The target message is divided into bytes each of which is composed of 8 bits. The secret key length is modified to be equal to the target message length. Changing the value of the bits depends on a truth table in which the secret key and target message act as main elements. Each person who receives the message has their own sub-secret key. This key is composed of two parts; the first part is a value that points to where the first place of a dummy bit is added to each byte in the encrypted message, and the second part is also a value that points to where the second place of a dummy bit is added to each byte in the encrypted message.

The values of these two dummy bits are generated randomly by the system. Finally, the system generates a different encrypted message every time even if the same secret key, the sub-secret key, and the same plaintext are used several times because of each byte contains two bits have random values. The decryption procedure is similar to the encryption procedure in processing but in reverse order starts by removing the previously generated random dummy bits and then decrypting the message. The proposed algorithm system can regenerate the original binary data byte with no loss or lack of data during and after the encryption or decryption process. By using unlimited secret keys length, and a sub-secret key is owned by each person who receives the encrypted message, the algorithm is more secure and it's hard to guess the key value or be attacked.

The proposed algorithm has been tested and compared with other recent algorithm and it was fast, simple and flexible enough. Validation of the new algorithm security requirements have been applied and it has been suitable for using it in many software applications.
The second section demonstrates the proposed algorithm (CryptoBin), the third section discusses the architecture of the algorithm system, the proves of the strength, the performance and security analysis for CryptoBin based system and its results. Finally, the conclusion will be introduced.

2 Proposed Work

The cryptographic algorithm system for binary codes which discussed and published after many tests and trials to attack it found that it has some drawbacks and weakness in the secret key system, and should be improved. This article will perform a study of the CryptoBin algorithm (ours) and try to explain its strength and resistance to attacks. So that a new method for secure communication of information and multimedia encryption proposed here. This technique contains advantages of both multimedia (Audio, Image, and video) cryptography and normal encryption data. This article is used to achieve and solve the problem of the weakness and drawbacks of the former system that mentioned before and it strengthens the secret key to be difficult to break. This cryptographic algorithm system is called CryptoBin which deals with Binary codes (0,1) bits.

The proposed secret key is a binary number which characterized by an unlimited size of bits, the bits can be less, equal or greater than the target message (plain text). The algorithm system compares the bit value of the secret key with the bit value of the target message and generates a new encrypted message that has an equal length of the target message. The algorithm system compares the bit value of the secret key and the target message using logical equations based on a given truth table, resulting in the encrypted message. For example, if the bit value of the secret key is equal to "1", the bit value of the target message will change from "1" to "0" or "0" to "1", else if the bit value of the secret key is "0", the bit value of the target message will not change and be as it "0" is "0" and "1" is "1".

2.1 Architecture of The Algorithm System

The CryptoBin algorithm consists of a secret key, plain message, and a truth table. The secret key and the plain

message are binary numbers, and the truth table controls the output bit's value (an encrypted bit).

Encryption System.

For example:

plain text = "Hello World"

binary input = "010010000110010101101100011011000111100100000010101110110111011100100110110001100100" secret key = "723" in decimal form

secret key = "1011010011" in binary form

To obtain a modified secret key to be equal to the number of bits of the target message it should be repeated.

Modifiedsecretkey="1011010011101101001110110100111011010011101101001110110100111011010011101101001110100"

sub-secret key = "010110"

By using the sub-secret key, the system will generate random values of two dummy bits and add them to the encrypted message. The first part of the sub-secret key points to the position of first dummy bit will be added to each byte of the encrypted message, the second part of the sub-secret key points to the position of the second dummy bit will be added to each byte of the encrypted message. In this case, the first dummy bit position will be the third bit, the second dummy bit position is the seventh bit with random values.

Truth Table =

Key	Msg	Enc Msg
0	0	0
0	1	1
1	0	1
1	1	0

The common idea of Truth Table that if the key's bit value is "0" then the encrypted bit value will be unchanged, else if the key's bit value is "1" as shown in Figure 1 then the encrypted bit value will be changed from "1" to "0" or "0" to "1".

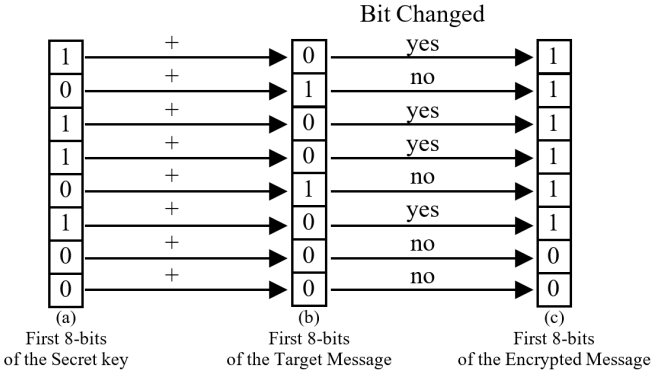


Figure 1: One-byte encryption

The Plain Message

Bit no.	1	2	3	4	5	6	7	8	9	10	11	12	
	0	1	0	0	1	0	0	0	0	1	1	0	
13	0	1	0	1	0	1	1	0	1	1	0	0	24
25	0	1	1	0	1	1	0	0	0	1	1	0	36
37	1	1	1	1	0	0	1	0	0	0	0	0	48
49	0	1	0	1	0	1	1	1	0	1	1	0	60
61	1	1	1	1	0	1	1	1	0	0	1	0	72
73	0	1	1	0	1	1	0	0	0	1	1	0	84
85	0	1	0	0	88								

The Secret Key

Bit no.	1	2	3	4	5	6	7	8	9	10	11	12	
	1	0	1	1	0	1	0	0	1	1	1	0	
13	1	1	0	1	0	0	1	1	1	0	1	1	24
25	0	1	0	0	1	1	1	0	1	1	0	1	36
37	0	0	1	1	1	0	1	1	0	1	0	0	48
49	1	1	1	0	1	1	0	1	0	0	1	1	60
61	1	0	1	1	0	1	0	0	1	1	1	0	72
73	1	1	0	1	0	0	1	1	1	0	1	1	84
85	0	1	0	0	88								

The Encrypted Message

Bit no.	1	2	3	4	5	6	7	8	9	10	11	12	
	1	1	1	1	1	1	0	0	1	0	0	0	
13	1	0	0	0	0	1	0	1	0	1	1	1	24
25	0	0	1	0	0	0	1	0	1	0	1	1	36
37	1	1	0	0	1	0	0	1	0	1	0	0	48
49	1	0	1	1	1	0	1	0	0	1	0	1	60
61	0	1	0	0	0	0	1	1	1	1	0	0	72
73	1	0	1	1	1	1	1	1	1	1	0	1	84
85	0	0	0	0	88								

The sub-secret key is "010110"

Encrypted Message After adding Dummy Bits

Bit no.	1	2	3	4	5	6	7	8	9	10	11	12	
	1	1	1	1	1	1	0	1	0	0	0	1	
13	0	0	1	0	1	0	0	0	0	0	1	1	24
25	0	1	1	0	1	1	0	1	0	0	1	1	36
37	0	0	0	0	1	0	0	1	0	1	0	1	48
49	1	1	1	0	0	1	1	0	0	1	0	0	60
61	1	0	0	0	1	0	1	1	1	1	1	0	72
73	1	0	1	0	1	0	0	1	0	1	0	0	84
85	0	0	0	0	1	1	0	1	1	0	0	0	96
97	1	0	0	1	1	1	1	1	1	1	1	1	108
109	1	0	1	1	0	0	1	0	0	117			

EncryptedBinary="1111110100010010100000110110110100110000101010111100110010100101011111010101001010000011011001001011111111101100100"

Decryption System.

For Decrypting the same example:

EncryptedBinary="11111100100010000101011100100010101111001001010010111010010101000011110010111111111010000"
secret key = "723" in decimal form
Secret key = "1011010011" in binary form
sub-secret key = "010110"

By using the sub-secret key, the system will search for the two dummy bits that have already been added to the encrypted message, then remove them from that message. The first part of the sub-secret key points to the position of the first dummy, the second part of the sub-secret key points to the position of the second dummy bit that has been added to each byte of the encrypted message. In this case, the first dummy bit position will be the third bit, the second dummy bit position is the seventh bit.

To obtain a modified secret key to be equal to the number of bits of the target message it should be repeated.

Modifiedsecretkey="10110100111011010011101101001110111010011101101001110111010011101101001110111010011101101001110110100"

By using the same truth table in the decryption process the bit value of 0 changed to 1 and 1 to 0 according to the rules which added in the truth table and Figure 2 shows a sample of a ciphered 8-bits changed to decrypted 8-bits.

Truth Table =

Key	Msg	Enc Msg
0	0	0
0	1	1
1	0	1
1	1	0

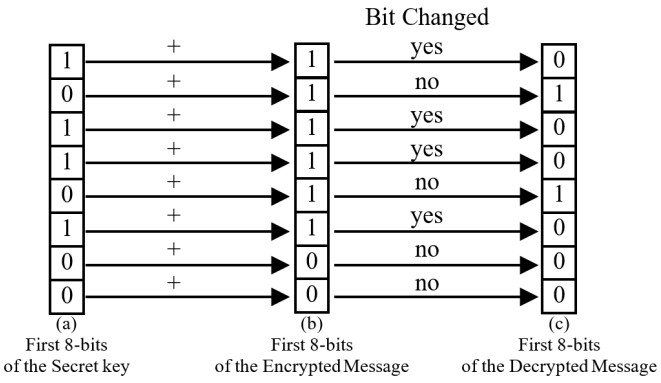


Figure 2: One-byte decryption

The Encrypted Message Including Dummy Bits

Bit no.	1	2	3	4	5	6	7	8	9	10	11	12	
	1	1	1	1	1	1	0	1	0	0	0	1	
13	0	0	1	0	1	0	0	0	0	0	1	1	24
25	0	1	1	0	1	1	0	1	0	0	1	1	36
37	0	0	0	0	1	0	0	1	0	1	0	1	48
49	1	1	1	0	0	1	1	0	0	1	0	0	60
61	1	0	0	0	1	0	1	1	1	1	1	0	72
73	1	0	1	0	1	0	0	1	0	1	0	0	84
85	0	0	0	0	1	1	0	1	1	0	0	0	96
97	1	0	0	1	1	1	1	1	1	1	1	1	108
109	1	0	1	1	0	0	1	0	0	117			

The sub-secret key is "010110"
The Encrypted Message After Removing Dummy Bits

Bit no.	1	2	3	4	5	6	7	8	9	10	11	12	
	1	1	1	1	1	1	0	0	1	0	0	0	
13	1	0	0	0	0	1	0	1	0	1	1	1	24
25	0	0	1	0	0	0	1	0	1	0	1	1	36
37	1	1	0	0	1	0	0	1	0	1	0	0	48
49	1	0	1	1	1	0	1	0	0	1	0	1	60
61	0	1	0	0	0	0	1	1	1	1	0	0	72
73	1	0	1	1	1	1	1	1	1	1	0	1	84
85	0	0	0	0	88								

The Secret Key

Bit no.	1	2	3	4	5	6	7	8	9	10	11	12	
	1	0	1	1	0	1	0	0	1	1	1	0	
13	1	1	0	1	0	0	1	1	1	0	1	1	24
25	0	1	0	0	1	1	1	0	1	1	0	1	36
37	0	0	1	1	1	0	1	1	0	1	0	0	48
49	1	1	1	0	1	1	0	1	0	0	1	1	60
61	1	0	1	1	0	1	0	0	1	1	1	0	72
73	1	1	0	1	0	0	1	1	1	0	1	1	84
85	0	1	0	0	88								

The Decrypted message

Bit no.	1	2	3	4	5	6	7	8	9	10	11	12	
	0	1	0	0	1	0	0	0	0	1	1	0	
13	0	1	0	1	0	1	1	0	1	1	0	0	24
25	0	1	1	0	1	1	0	0	0	1	1	0	36
37	1	1	1	1	0	0	1	0	0	0	0	0	48
49	0	1	0	1	0	1	1	1	0	1	1	0	60
61	1	1	1	1	0	1	1	1	0	0	1	0	72
73	0	1	1	0	1	1	0	0	0	1	1	0	84
85	0	1	0	0	88								

Decryptedbinaryoutput="01001000011001010110110001101100011011110010000001010111011011110111001001100110110001100100"
Decrypted text = "Hello World"

2.2 CryptoBin Implementation

Now we propose the CryptoBin algorithm encryption by using a simple coding language such as VB.NET. For simplicity, we use a simple series of binary codes for plaintext and secret key as shown in Algorithm 1, and 2.

Algorithm 1 CryptoBin Algorithm (Encryption)

```
'Encryption Process
'plain text = "Hello World"
Diminputstr="01001000011001010110110001101100011011110010000001010111011011110111001001101100011001100"
Dim key = "1011010011" '723 in decimal form
Dim keybit = ""
Dim txtbit = ""
Dim resbit = ""
Dim result = ""
Dim keylength
For x = 0 To inputstr.Length - 1 Step key.Length
keylength = key.Length
'if no. of bits of plaintext length < no. of bits of key length
If (inputstr.Length) - x < key.Length Then keylength = (inputstr.Length) - x
For n = 1 To keylength
txtbit = Mid(inputstr, x + n, 1)
keybit = Mid(key, n, 1)
If keybit = "0" And txtbit = "0" Then
resbit = "0"
ElseIf keybit = "0" And txtbit = "1" Then
resbit = "1"
ElseIf keybit = "1" And txtbit = "0" Then
resbit = "1"
ElseIf keybit = "1" And txtbit = "1" Then
resbit = "0"
End If
result = result + resbit
Next n
Next x
```

Encryptedresult="11111100100010000101011100100010101111001001010010111010010101000011110010111111111010000"

Decryptedresult="01001000011001010110110001101100011011110010000001010111011011110111001001101100011001100"

We can use this code for Image Encryption also. For example; we use an image file named "m.png" and the encrypted file named "m-Encrypt.png", Figure 3 shows the image before and after encryption.

Algorithm 2 CryptoBin Algorithm (Decryption)

```
'Decryption Process
'Encryptedtext="11111100100010000101011100100010101111001001010010111010010101000011110010111111111010000"
Diminputstr="11111100100010000101011100100010101111001001010010111010010101000011110010111111111010000"
Dim key = "1011010011" '723 in decimal form
Dim keybit = ""
Dim txtbit = ""
Dim resbit = ""
Dim result = ""
Dim keylength
For x = 0 To inputstr.Length - 1 Step key.Length
keylength = key.Length
'if no. of bits of plaintext length < no. of bits of key length
If (inputstr.Length) - x < key.Length Then
keylength = (inputstr.Length) - x
For n = 1 To keylength
txtbit = Mid(inputstr, x + n, 1)
keybit = Mid(key, n, 1)
If keybit = "0" And txtbit = "0" Then
resbit = "0"
ElseIf keybit = "0" And txtbit = "1" Then
resbit = "1"
ElseIf keybit = "1" And txtbit = "0" Then
resbit = "1"
ElseIf keybit = "1" And txtbit = "1" Then
resbit = "0"
End If
result = result + resbit
Next n
Next x
```

(a) Original image (b) Encrypted image

Figure 3: Image before and after encryption

3 Performance and Security Analysis

For designing a very good encryption system, it should be resisting all kinds of common attacks such as brute-force attacks, the man in middle attack, dictionary at-

tack, side channel attack, cipher-text attack, and various attacks. Some of the security analysis techniques can perform on the CryptoBin encrypting system while the statistical analysis and key space are included.

The security analysis of the proposed CryptoBin encryption for image encryption will be discussed in this section, such as Histogram Analysis, Correlation between plain images and cipher images, Information Entropy, and Key Space Analysis to prove that the proposed encryption system is effective, safe and more secure against all common attacks. Experiments are executed by using the "Matlab" software. The key parameter for example; $\text{key})_{\text{Decimal}} = 723$ or $\text{key})_{\text{Binary}} = 1011010011$. This parameter must be kept secret. The same key is used to decrypt the cipher-images.

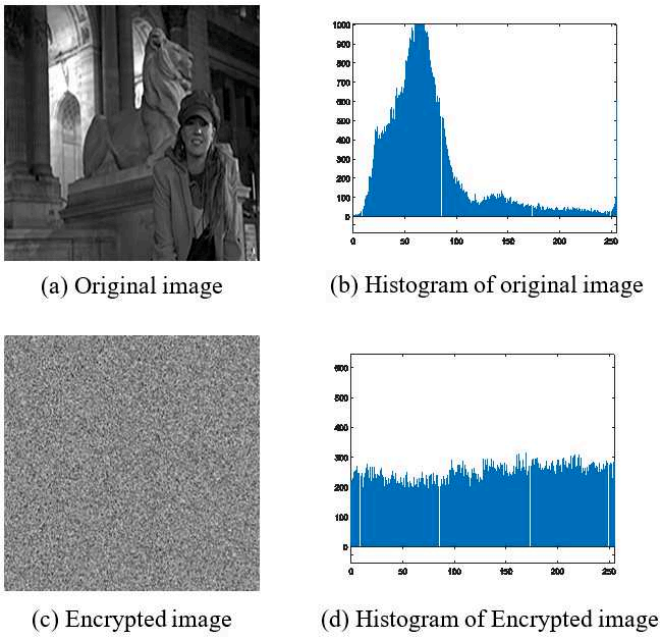
3.1 Statistical Analysis

To demonstrate the strength of the proposed encryption system, a statistical analysis was performed showing superior confusion characteristics and also diffusion characteristics in the nature of strong resistance against all kinds of statistical attacks. This is done by the study of Histogram Analysis, Correlation, Key Space Analysis, and Information Entropy between the plain images and ciphered images [17]. Applying the statistical analysis on the CryptoBin system demonstrated the properties of diffusion and the superior confusion of the system that effectively protect from statistical attacks. these results will be shown by the histogram tests on the plain and the ciphered images.

3.1.1 Histogram Analysis

Two techniques of confusion and diffusion may be used, as Shannon pointed out, to defeat any strong attacks depending on the statistical analysis. Histogram test is one of Shannon methods and it is applied to ciphered images. We have a grey-scale image (256X256) has different contents, and we calculate its histogram which shows the distribution of pixel intensities of the image. The attacker uses frequency analysis to obtain the secret key or the plain-pixels. This attack type is called a statistical attack. To prevent that statistical attack, the histogram of the original image and histogram of the encrypted image shouldn't have a statistical similarity. Therefore, the histogram of the encrypted image should be relatively flat or with a uniform statistical distribution, indicating the strength and quality of the encryption system [10].

Figure 4 show histograms of image 'm.png' before and after encryption. Histogram of the encrypted image looks relatively flat and with a uniform statistical distribution and distinctive from the histogram of the original image. Based on the experiment results above, the encryption process turned out to return a noisy image, and also the histograms of the previously encrypted image are very similar to the uniform distribution, distinctive from the original image and no statistical similarity to the original



(a) Original image (b) Histogram of original image

(c) Encrypted image (d) Histogram of Encrypted image

Figure 4: Histograms of the plain image and ciphered image

one is contained. The flat histogram in encrypted images can make an attacker's task very difficult to infer pixel values or secret keys using a statistical attack. This corresponds to the ideal security set by Shannon, and the encryption system resists against the known attacks [1].

3.1.2 Correlation between Plain and Cipher Images

Correlation is some of a wide class of statistical relationships involving dependence, though in keeping usage it usually identifies how close two variables are to presenting a linear relationship together. We have analyzed the correlation between horizontally, vertically, and diagonally adjacent pixels in a wide range of normal images as well as their encrypted images. The correlation coefficient analysis indicates the partnership among pixels in the cipher image [14]. In the newest scheme, the correlation among adjacent pixels is less than that of the original image. This low correlation value between the original images and their encryption indicates less resemblance between them, which supplies more resistant to attacks. The Statistical correlation is a measure that states the effectiveness of the linear relationship between two random variables. Let 'a' and 'b' are two random variables, each consisting of n elements, the correlation coefficient of both random variables is calculated by the Equations (1,2,3, and 4):

$$r_{\text{ab}} = \frac{cov(a,b)}{\sqrt{D(a)D(b)}} \quad (1)$$

$$D(a) = \frac{1}{n}\sum_{i=1}^{n}[a_{\text{i}} - E(a)]^2 \quad (2)$$

$$cov(a,b) = \frac{1}{n}\sum_{i=1}^{n}[a_i - E(a)][b_i - E(b)] \quad (3)$$

$$cov(a,b) = \frac{1}{n}\sum_{i=1}^{n} a_i. \quad (4)$$

Table 1: Correlation between both adjacent pixels in the plain and ciphered images

	Plain image	Ciphered image
Diagonal	0.9358	-0.4020
Horizontal	0.9427	0.0082
Vertical	0.9858	-0.0005

From Table 1 results and the correlation charts, we noticed that there is a negligible correlation between both adjacent pixels in the ciphered image. But both adjacent pixels in the original image are extremely correlated. Correlation in the encrypted images is exceptionally little or insignificant while the suggested encrypting scheme is utilized. Therefore, the suggested encryption system has a great change and substitution properties.

3.1.3 Information Entropy

The information entropy is simply the average (expected) amount of the information from the event or how much information there's in an event. In general, the more uncertain or random the event is, the more information it will contain. It was founded in 1949 by Claude E. Shannon [16]. Entropy test is the other one of Shannon methods and it is applied to ciphered images, the indicator of randomness is the information entropy that can be calculated from the following Equation (5).

$$H(x) = -\sum_{i=1}^{2^N-1} P(x_i) log_2[P(x_i)] \quad (5)$$

The entropy amounts the random value or average uncertainty in xi where P(xi) is how much information from one instance of the random variable xi. If all symbols have the same probability then the information entropy will be H(x) = 8, while x = (x_0,x_1,x_2, ..., x_{28} -1) and P(x_i) = 1/28 (i=0, 1, ..., 255), that matches the ideal case. Basically, the scrambled images information entropies are less set alongside to the perfect case. The expected entropy of the scrambled image is close to the perfect case in order to create a great image encryption scheme. We may consider the image to be more random somehow if the information entropy is closer to 8. Table 2 shows the plain image entropy value and its equivalent ciphered image entropy value.

3.2 Keyspace Analysis

For the encryption scheme to be so effective, it must be sensitive to the secret keys. The key space size has to be big enough to prevent and stop the brutal attacks [21]. In this case, the size of the key space is unlimited. The results of the experiments showed that CryptoBin is quite sensitive to the secret key. Table 3 shows the CryptoBin is sensitive to the secret keys. As visible once the secret key is changed a little the correlation coefficients become absolutely different.

Table 2: Entropy values for original and encrypted images

Image	Entropy value
Original Image(plain Image)	7.1200
Encrypted Image (Cipher Image)	7.9919

Table 3: Correlation values for the image by using different secret keys

Correlation	Vertical	Horizontal
Original Image	0.9858	0.9427
Encrypted Image (key1)	0.0160	-0.0174
Encrypted Image (key2)	0.0785	0.0711
Encrypted Image (key3)	0.0068	0.0078

4 Conclusion

An advanced approach for a cryptographic system using binary codes based on (0,1) called CryptoBin is proposed. This new algorithm depends on converting the target message into zero and one and then swap the bit value from one to zero and from zero to one by using mathematical equations. This new system has a secret key, this secret key has an unlimited length and a sub-secret key added to the system. The sub-secret key is used to generate a different encrypted message every time even if the same secret key, the sub-secret key, and the same plaintext are used several times, that increased the confidentiality and strength of the system. To prove the effectiveness of the proposed encryption system Histogram Analysis, Correlation between plain images and cipher images, Information Entropy, and Key Space Analysis has been tested. The demands of secrecy and strength to confront the intruders with high efficiency have been achieved and the new system introduced high-security analysis. The data was encrypted by the proposed encryption system against any attacks. The proposed system can be used to secure any software applications.

References

[1] L. Bi, S. Dai, and B. Hu, "Normalized unconditional e-security of private-key encryption," *Entropy*, vol. 19, no. 3, p. 100, 2017.

[2] S. Bouchkaren and S. Lazaar, "Caes cryptosystem: Advanced security tests and results," *International Journal of Network Security*, vol. 20, no. 1, pp. 177–183, 2018.

[3] S. Dey and R. Ghosh, "A review of cryptographic properties of s-boxes with generation and analysis of crypto secure s-boxes," *International Journal of Electronics and Information Engineering*, vol. 8, no. 1, pp. 49–73, 2018.

[4] N. K. El Abbadi, S. T. Abaas, and A. A. Alaziz, "New image encryption algorithm based on diffie-hellman and singular value decomposition," *matrix*, vol. 55, no. 89, p. 144, 2016.

[5] A. H. Eltengy, S. M. Shohieb, M. S. Ksasy, and A. E. TakielDeen, "A new advanced cryptographic algorithm system for binary codes by means of mathematical equation," *ICIC Express Letters*, vol. 12, no. 2, pp. 300–308, 2018.

[6] A. H. Eltengy, A. E. Takieldeen, and H. M. Elbakry, "Implementation of a hybrid encryption scheme for sms/multimedia messages on android," *International Journal of Computer Applications*, vol. 85, no. 2, pp. 300–308, 2014.

[7] A. H. Eltengy, A. E. TakielDeen, and H. M. Elbakry, "Implementation of an encryption scheme for voice calls," *International Journal of Computer Applications*, vol. 144, no. 2, pp. 300–308, 2016.

[8] A. Hazra, S. Ghosh, and S. Jash, "Review on dna based cryptographic techniques," *International Journal of Network Security*, vol. 20, no. 6, pp. 1093–1104, 2018.

[9] L.-C. Huang, L.-Y. Tseng, and M.-S. Hwang, "The study of 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] M.-S. Hwang, C.-C. Lee, and S.-F. Tzeng, "A new knapsack public-key cryptosystem based on permutation combination algorithm," *Information Journal of Applied Mathematics and Computer Sciences*, vol. 5, no. 1, pp. 33–38, 2009.

[12] C.-C. Lee, M.-S. Hwang, and S.-F. Tzeng, "A new convertible authenticated encryption scheme based on the elgamal cryptosystem," *International Journal of Foundations of Computer Science*, vol. 20, no. 02, pp. 351–359, 2009.

[13] L. Liu, Y. Li, Z. Cao, and Z. Chen, "One private broadcast encryption scheme revisited," *International Journal of Electronics and Information Engineering*, vol. 7, no. 2, pp. 88–95, 2017.

[14] N. K. Pareek, "Design and analysis of a novel digital image encryption scheme," *arXiv preprint arXiv:1204.1603*, pp. 300–308, 2012.

[15] M. Rasslan, G. Elkabbany, and H. Aslan, "New generic design to expedite asymmetric cryptosystems using three-levels of parallelism," *International Journal of Network Security*, vol. 20, no. 2, pp. 371–380, 2018.

[16] R. A. Rodríguez, A. M. Herrera, Á. Quirós, M. J. Fernández-Rodríguez, J. D. Delgado, A. Jiménez-Rodríguez, J. M. Fernández-Palacios, R. Otto, C. G. Escudero, T. C. Luhrs *et al.*, "Exploring the spontaneous contribution of claude e. shannon to eco-evolutionary theory," *Ecological modelling*, vol. 327, pp. 57–64, 2016.

[17] T. Shah, I. Hussain, M. A. Gondal, and H. Mahmood, "Statistical analysis of s-box in image encryption applications based on majority logic criterion," *International Journal of Physical Sciences*, vol. 6, no. 16, pp. 4110–4127, 2011.

[18] A. Singh and R. Gilhotra, "Data security using private key encryption system based on arithmetic coding," *International Journal of Network Security & Its Applications*, vol. 3, no. 3, pp. 58–67, 2011.

[19] A. E. Takieldeen, E. El-Badawy, S. Gobran *et al.*, "Digital image encryption based on RSA algorithm," *IOSR Journal of Electronics and Communication Engineering*, vol. 9, no. 1, pp. 69–73, 2014.

[20] A. E. Takieldeen, M. A. Shawky, H. M. Elkamchouchi, I. M. Fouda, and M. M. Khalil, "A new image encryption algorithm combining the meaning of location with output feedback mode," in *13th IEEE APCA International Conference on Control and Soft Computing (CONTROLO'18)*, pp. 521–525, 2018.

[21] H. Zhu and R. Wang, "A survey to design privacy preserving protocol using chaos cryptography," *International Journal of Network Security*, vol. 20, no. 2, pp. 313–322, 2018.

Biography

Ahmed H. Eltengy is currently a PhD candidate at Mansoura University, Computer Science Department. He received M. D. in Computer Sciences (2014). He has a lot of publications in various international journals (i.e. "International Journal of Scientific & Engineering Research", "International Journal of Computer Applications", and "ICIC Express Letters"). His interests of research include Software and Hardware Security Programming, Microcontroller and Embedded Systems.

Dr. Samaa M. Shohieb is a professor assistant in computer information systems department, faculty of Computers and Information, Mansoura University, Egypt. She's interested in Human-computer Interaction integrated with E-society and designing creative ICT solutions for diverse users with specified capabilities. She is an editorial board of many international Journals including inderscience, and Elsevier.

Dr. Ali E. Takieldeen (IEEE Senior Member) received the PhD degree in Electronics and Communications Engineering in "Encryption and Data Security in Digital Communication Systems". He has a lot of publications in var-

ious international journals and conferences. His current research interests are in multimedia processing, wireless communication systems, Microcontroller and Field Programmable Gate Array (FPGA) applications.

Prof. Mohamed S. Ksasy holds Ph. D. in Electronics and Communications (1992). He received M. Sc. In Electronics and Communications (1985). He is a Member of the Institute of Electrical & Electronics Engineers (IEEE), Member of the International Journal of Computers Applications (IJCA), and Member of The Arab Control Systems Association (ACSA).

Network Security Situation Assessment Based on Text SimHash in Big Data Environment

Pengwen Lin and Yonghong Chen
(Corresponding author: Yonghong Chen)

College of Computer Science, Technology, Huaqiao University
No.668 Jimei Avenue, Xiamen, Fujian 361021, China
(Email: djandcyh@163.com)

(Received Feb. 1, 2018; Revised and Accepted June 21, 2018; First Online Dec. 10, 2018)

Abstract

The existing methods of network security situation assessment have high complexity and not effectively in the big data environment. This paper proposes an assessment model based on SimHash in the big data environment. First, a large-scale network is divided into multiple modules. Then get secure data of the internal nodes of modules. Based on the SimHash algorithm, in turn quantifies the node security situation, module security situation, network security situation. Finally, the experiment is designed to verify the model. Results show that the model can effectively adapt to a large-scale network and have high accuracy.

Keywords: Big Data; Complex Network; Network Security Situation Assessment; Text SimHash Algorithm

1 Introduction

With the continuous development and popularization of network technology, the amount of data is growing at an unimaginable speed in recent years. More and more people use the term "big data" to describe and define the generation of massive data during the information explosion era. Today, both industry and academia have generated a great deal of interest in the field of big data. The value inherent in the big data has become the driving force behind the storage and processing of big data [19]. Many studies, including [26] and [17], have made cloud storage widely used, which provides the basis for big data analysis. [14] pointed out that because of the concept of data processing changes in the big data environment, many scholars have already begun to study the big data analysis technology deeply. For instance, [24] elaborates on the techniques of big data analysis and the challenges it faces.

On the other hand, the Internet has become a critical infrastructure and Internet security has a direct bearing on the fundamental interests of the public [11]. Today the scale of the network is getting bigger and bigger, the topological structure and environment of the network are more and more complicated, cybersecurity incidents have risen dramatically, and the issues of cybersecurity have become increasingly prominent [1]. In order to address these challenges, Intrusion Detection System, Firewall, security-audit and other security protection and management system have been widely used. However, these products all consider network security from a single aspect. The lack of a synergistic mechanism between each of these products can only be used by themselves and form isolated islands of information.

Network security situation assessment has been proposed by many scholars under such a background. Strengthening the assessment of the security situation of information systems is a necessary management measure to protect the core information infrastructure [8]. Network security situation assessment means that the security-related elements are perceived and acquired from the perspective of time and space through technological means, and the network security status is judged through the integrated analysis of data information [12, 23].

Endsley in [7] put forward and defines the concept of situation awareness for the first time in 1988. Endsley believed that situation awareness was the perception of the elements of the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status into the near future. However, this concept was mainly applied to the aviation field. Bass [2] first introduced the concept of situational awareness into the field of network security in 1999 and proposed an evaluation framework based on intrusion detection system, but it did not implement it. Gorodetsky *et al.* [9] proposed an evaluation method of network security situation based on asynchronous data flow, which used multi-agent anomaly detection network data flow to analyze various security events to get the security situation. But usually, data flow cannot represent all the basic security information in networks. XiuZhen-Chen *et al.* [4] proposed a quantitative hierarchical threat evaluation model for network security. The evaluation policy used in the model is from bottom to top and from local to the whole. The

threat metrics of services, hosts, and local networks are calculated by weighting the importance of services and hosts based on attack frequency, severity and network bandwidth consumption, and then evaluates the status of the security threat. [25] proposed a situational awareness model, which analyzes the current situation in the network environment and generates corresponding measures. The situation is influenced 00by the measures and then a new trend is formed. [13] references the mechanism of body temperature change caused by biological immune system imbalance, analyzes antibody concentrations change caused by the change process of various types of detectors in computer immune system, and proposes a quantitative risk evaluation model for network security based on body temperature. [27] classifies the security events in the network to identify attackers, then casually correlates each attack scenario, identifies the corresponding track and phase of an attack, and finally establishes the situation quantitative criteria, combining the attack phase and its threat index evaluation of the cybersecurity situation. However, all of the models mentioned above are difficult to adapt to the situation of a large quantity of data and fast generating of data in the big data environment. Secure data in [29] includes intrusion detection log, firewall log, virus log, network scanning, illegal external links, and running state of equipment and real-time alarm. Then combined with the PSR method, the fuzzy logic model and the entropy weight method in an empirical study for feasible urban public security evaluation modeling. It gives us a good reference value. However, it does not give a comprehensive measure of the value of the situation. [5] used SIEM as input data, and based on CVSS and attack models, the technologies used including a set of integrated security metrics to conduct risk assessment. [6] considered the uncertainty of the assessment data and translates it into an objective weight through uncertainty measures. Then, using D-S evidence theory and pignistic (from the Latin pignus, a bet) probability transformation, a consensus decision about the degree of network security risk is obtained.

The above methods provide a feasible solution for researching network security situation assessment. In the meanwhile, there are some common defects. For instance, the existing methods are hard to adapt to the big data environment because of the high complexity of evaluation models and algorithms, which leads to the deviation of the quantitative results of the network security situation, and the feedback is not timely enough. To address these problems, this paper presents a network security situation evaluation model based on the SimHash algorithm to adapt to the big data environment. First, the method of complex networks is used to divide a large-scale network. Fusion of multi-source heterogeneous data on each node in the local network to obtain the secure data, and then use SimHash to assess the security situation of nodes quickly and efficiently. Finally, integrated by the weights of nodes and modules to quantitative the status of network security.

The limitation of the model in this paper is that in a large-scale network, the topology is dynamically changing. However, the division of modules in our model is completed before the assessment of the network security situation and does not change in real time following the change of the topology.

The rest of this paper is arranged as follows: In Section 2 we introduce the framework of our assessment model. Section 3 gives the key algorithms and related theoretical basis. Section 4 presents the experimental results and discussion. The paper is concluded in Section 5.

2 Network Security Situation Assessment Model

A large-scale network usually contains a great number of hosts, network devices, and various detection systems, and these detection systems monitor the network from different perspectives and generate logs and alerts. Traditional network security situation assessment usually uses only a single alert or log detected, and the single data source also directly leads to the deviation of the assessment result from the actual situation. And the method of evaluation often adopted a relatively complicated algorithm, which directly affected the timeliness of the assessment, and delays the best time for the network administrator to take measures. Aiming at these problems, the network security situation assessment model based on the SimHash algorithm in the big data environment is put forward.

Firstly terms used in the assessment model are explained.

Topology (T). It's graph structure which used to represent the information about nodes and their connection in a large-scale network environment.

Service (S). It refers to the services provided by the node to determine the weight of the node.

Log (L). It contains information such as system log, security log, application log, and alert log generated during network operation. The information of every log can be characterized by a sextuple $(id_l, time_l, type, info_l, id_{st}, id_{dt})$, where (id_l) is the unique identification of log, $time_l$ is the time when log generates, $type$ is the type of log, $info_l$ is the description of log, id_{st} is the identification of the node which generates log, and id_{dt} is the identification of the node which is the target of the security event.

Vulnerability (V). It refers to the vulnerability of the node and determines the success probability of an attack when it occurs. Every vulnerability information could be characterized by a quadruple $(id_v, time_v, pro_v, impact_v, info_v)$, where id_v is the unique identification of vulnerability, $time_v$ is the time when vulnerability scans, pro_v is the probability

of successful exploitation, and $info_v$ is the description of the vulnerability.

Attack (A). It represents the attack on the node. The information of every attack can be characterized by a sextuple $(id_a, time_a, st, dt, info_a, id_v)$, where id_a is the unique identification of attack, $time_a$ is the time when attack occurs, st and dt represented the source and destination of attack respectively, $info_a$ is the description of attack, and id_v is the identification of vulnerability which used by attack.

Node situation awareness (NSA). It's the value of the security situation of the node and consists of topology, vulnerability and attack and denoted by $NSA = (T, V, A)$.

Module situation awareness (MSA). It's the value of the security situation of the module and consists of NSA and the weight of the nodes in the module and denoted by $MSA = (NSA, \omega_{node})$.

Network situation awareness (SA). It's the value of the security situation of a large-scale network and consists of MSA and the weight of the modules and denoted by $SA = (MSA, \omega_{module})$.

Then the framework of the model is shown in Figure 1.

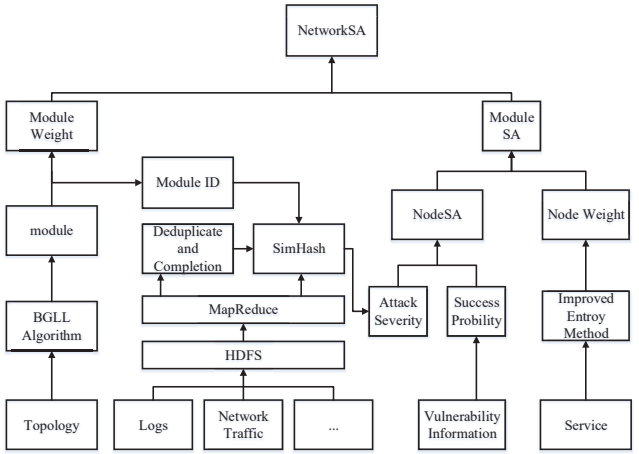


Figure 1: Network security situation assessment model

The calculation steps is shown in Table 1.

3 Network Security Situation Assessment Based on Big Data Analytics and SimHash

In this section, first, we divide the network into multiple modules and preprocess the data on the Hadoop platform. Then we introduce SimHash algorithm and use it to evaluate the security of the nodes. Finally, we determine the weight of the nodes and calculate the security situation of module and network.

3.1 Data Preprocessing

Due to the data obtained from various data sources such as logs and network traffic, their format is different, their generation is fast and the data contains dirty data. All of these leads to data exchange and sharing cannot be performed with each other efficiently. At the same time, [21] pointed out today's sophisticated network-attacks that occur across multiple dimensions and stages, traditional platforms will have no chance to defend a network.

To address these problems, the idea of the module is brought from the complex network into network security situation assessment.

The complex network generally consists of a mass of nodes and the connections between nodes are seriously complex. A complex network is widely used in various scientific fields to model and analyzes complex systems. Many networks have a community structure. Community structure is there are many associations in the network, and the connection among these associations is relatively sparse and the connection within the associations is relatively dense. Community discovery is using information contained in the topological structure from the complex network to resolve its modular community structure.

Community module index Q [20] is usually used to characterize the strength of community characteristics. Defined as in Equation (1):

$$Q = \frac{1}{2m} \sum_{ij} \left(A_{ij} - \frac{k_i k_j}{2m} \right) \delta(C_i, C_j). \tag{1}$$

Where k_i and k_j are the degrees of nodes, A_{ij} are the weight of the edge between node i and node j, C_i is the community of node i, m is the total number of network edges. $\delta(C_i, C_j) = 1$ when $C_i = C_j$, otherwise is 0. The value of Q is between 0 and 1, generally $Q = 0.3$ as the lower bound of the social structure of the network.

In this paper, we use BGLL algorithm [3] to classify a large-scale network. The algorithm uses the positive or negative of ΔQ to determine whether the ith node should join the module the jth node belongs to. ΔQ is defined in Equation (2):

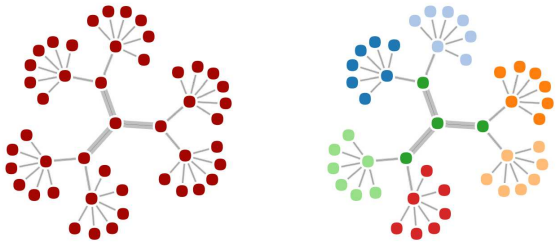
$$\Delta Q = \left[\frac{\sum_{in} + 2k_{i,in}}{2m} - \left(\frac{\sum_{tot} + k_i}{2m} \right)^2 \right] - \left[\frac{\sum_{in}}{2m} - \left(\frac{\sum_{tot}}{2m} \right)^2 - \left(\frac{k_i}{2m} \right)^2 \right]. \tag{2}$$

Where $\sum_{in}$ is the sum of the weights of all the edges inside the community; $\sum_{tot}$ is the sum of the weights of all the edges associated with the nodes inside the community; $k_{i,in}$ is the sum of weight of all edges connected to the community C. An example of dividing the network is shown in Figure 2.

Next, multi-source heterogeneous data, which is generating by the nodes inside each partitioned module is integrated. The purpose of data integration is to organize data in various independent systems into a whole

Table 1: Assessment steps

Step 1	Modules and its weight are obtained by using algorithms for detecting community structure in a complex network to divide a large-scale network.
Step 2	Collecting network security situation elements and then upload to the distributed file system to storage.
Step 3	The type of attack and the number of the attack in a period is obtained by preprocessing elements of network security situation.
Step 4	Inside the module, scan the vulnerabilities of nodes and calculate the success probability for each type of attack.
Step 5	For each node in the module, an algorithm based on SimHash is used to calculate the severity of the attack by attack type and a number of attacks.
Step 6	According to the severity of the attacks and the probability of successful attacks, node security situation is calculated.
Step 7	The weights of the nodes are calculated by using the services provided by the nodes, and then module security situation is obtained according to the security situation of the nodes.
Step 8	Using the module weight, combined with the security situation of modules, get the network security situation.



(a) predetect (b) detect

Figure 2: An example of dividing

according to certain rules by some technical means, so that other systems or users can access data effectively. First, collect multi-source heterogeneous data, and then upload them to the distributed file system for storage. Because of the correlation between the data generated by various detection systems, on the one hand, there is a large amount of redundant information in these data and cannot be directly used in the network security situation evaluation. On the other hand, the detection system also has omissions and false positives, and the data of multiple detection systems will be merged to complement each other. On the basis of the distributed file system, it is possible to unify the format of multi-source heterogeneous data, excluding a lot of noise data which are not related to network security situation assessment, and merge duplicate attribute data. Finally, these pre-processed data are stored in the database as the data that can be used directly by the network security situation assessment.

3.2 Node Security Situation Assessment Based on Simhash

3.2.1 SimHash

SimHash algorithm [10] is an efficient algorithm uses to find similar texts. It avoids the complicated way of comparing texts with each other, which greatly improves the efficiency compared with algorithms such as cosine similarity, Euclidean distance, Jaccard similarity coefficient.

SimHash algorithm is a type of dimension reduction method essentially, which maps high-dimensional vectors into smaller-sized signatures to represent the features of the original vectors. The main character is the Hamming distance between two signatures is positively correlated with cosine similarity between the corresponding feature vectors. [18] and [22] improve SimHash algorithm and apply the improved algorithm to different fields. This brings great inspiration to this research.

The SimHash algorithm is described in (Algorithm 1).

Algorithm 1 SimHash

Require: Text T, the length of hash b.
Ensure: Array $W[\ldots]$.

```
Begin
F(t) ← feature vector in T
W ← array of b zeros
for i ∈ F(t) do
    φ_i ← TraditionalHash(i)
    for j = 1 to b do
        if φ_ij = 1 then
            W[j] ← W[j] + ω_i
        else
            W[j] ← W[j] − ω_i
        end if
    end for
end for
for j = 1 to b do
    if W[j] ≥ 0 then
        W[j] ← 1
    else
        W[j] ← 0
    end if
end for
return W
End
```

3.2.2 Text Processing

Before using the SimHash algorithm for node security situation assessment, we need to construct the text as the input of the algorithm.

Firstly, a text is randomly generated, and the words that make up the text are not repeated with each other, and the number of words that make up the text is related to the total number of attacks during that time.

Then assign different words for different types of attacks. These words do not overlap with the words in the text.

Finally, extract the attack information within a period of time, and calculate the attack number for different types of attack. If there are n types of attacks, n copies of the original text are generated. For each type of attack, according to the number of attacks, replace some words in the copy with the assigned words.

According to the above description, if there are several attacks over a period of time, several texts which modified will eventually be obtained. Utilizing the SimHash algorithm to get several hash values corresponding to these texts which modified, and compare the Hamming distance between these hash values and the hash value generated by the original text.

Hamming distance is the number of different bits between the hash values of two b-bits that can be used to estimate the similarity between two vectors. The greater the Hamming distance, the less similarity between the two vectors is. This feature can be used to quantify the severity of a certain attack on a node over a period of time.

For the existing vulnerabilities, different types of attacks have different successful probability. If there are numerous attacks over a period of time, we will get several Hamming distances. We need to combine these Hamming distances with the corresponding attack success probability, and then reduce the result.

3.2.3 Assessment Algorithm Based on SimHash

Traditionally, SimHash algorithm is used for web page deduplication and document similarity detection. Due to the huge amount of data volume is generated by various security devices in a large-scale network, an efficient network security situation evaluation algorithm is urgently needed to enable network administrators to quickly understand the current security status of the network. However, most existing evaluation algorithms have a disadvantage in the computation time because of its complexity. So it's difficult to apply to a large-scale network environment.

In order to solve the problems, we introduce SimHash algorithm to the network security situation awareness. Based on SimHash, we propose our node security situation assessment algorithm. First, we use text processing which described above to generate pre-attack text and post-attack text. And then use these texts to quantify the severity of the attack and finally quantify the security situation of the node. The algorithm is shown in (Algorithm 2).

Algorithm 2 Node security situation assessment algorithm

Require: Attack information A, vulnerability information V.
Ensure: Node security situation NSA.

Begin
$b \leftarrow$ the length of hash
$d \leftarrow 0$
$T_1 \leftarrow$ randomly generate n words and each word isn't repeating
$F_1(t) \leftarrow$ feature vector on T_1
$H_1 \leftarrow \text{SimHash}(T_1, b)$
for $a \in A$ and $i = 0$ *to* n **do**
  $word \leftarrow$ randomly generate a word
  $T_2 \leftarrow \text{replace}(T_1, i, word)$
end for
$F_2(t) \leftarrow$ feature vector on T_2
$H_2 \leftarrow \text{SimHash}(T_2, b)$
for $i = 1$ *to* b **do**
  if $H_1[i] = H_2[i]$ **then**
    $d \leftarrow d + 1$
  end if
end for
for $a \in A$ **do**
  if id_v *in* V **then**
    $NSA \leftarrow NSA + d \cdot impact_v \cdot pro_v$
  else
    $NSA \leftarrow NSA + 0$
  end if
end for
return NSA
End

Based on vulnerability information about a node, the success probability of the attack is obtained and the security situation of the node is calculated using Equation (3).

$$NSA = \sum_{i=1}^{n} (svy_i \times p_i). \tag{3}$$

Where svy_i is the severity of the ith attack, a node may be attacked by many types of attack, p_i is the success probability of the attack based on vulnerability information.

3.3 Determine the Weight of Node

The improved entropy method [16] is used to determine the node weight. In information theory, entropy reflects the degree of disorders of information and is a measure of uncertainty. The smaller the entropy of an index, the smaller the uncertainty and the greater the amount of information carried, the greater the impact of this index on the comprehensive evaluation. The main steps of calculation are as follows:

Step 1: Depending to the service status of the host node, constructs a judgment matrix **R**:

$$\mathbf{R} = \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ r_{21} & r_{22} & \dots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m1} & r_{m2} & \dots & r_{mn} \end{pmatrix} \quad (4)$$

Where r_{ij} is the quantized value of the j service for the i nodes, 0 if it contains. n is the total number of nodes in the network and m is the total number of services in the network.

Step 2: Using traditional concept of entropy to calculate the entropy of the jth services(H_j):

$$H_j = -\Big(\sum_{i=1}^{n} f_{ij} \ln f_{ij}\Big) / \ln n \quad (5)$$

$$f_{ij} = r_{ij} / \sum_{i=1}^{n} r_{ij} \quad (6)$$

Where f_{ij} is the proportion of the ith nodes under the jth service in the service. Obviously, if $f_{ij} = 0$ that $\ln f_{ij}$ is meaningless, so the calculation of f_{ij} is modified to be:

$$f_{ij} = (1 + r_{ij}) / \sum_{i=1}^{n} (1 + r_{ij}). \quad (7)$$

Step 3: Calculating the difference coefficient for the jth service g_j:

$$g_j = (1 - H_j)/(m - E_c) \quad (8)$$

$$E_c = \sum_{j=1}^{m} H_j; 0 \le g_i \le 1, \sum_{j=1}^{m} g_i = 1 \quad (9)$$

For the jth service, the smaller the entropy, the greater the difference coefficient, the greater the impact on the node is.

Step 4: Calculating the objective weight (w_i) of each node in the network:

$$w_i = \sum_{j=1}^{m} (g_j / \sum_{j=1}^{m} g_j) \cdot f_{ij} \quad (10)$$

3.4 Calculate the Value of Network Security Situation

The previous section calculates the node's security situation NSA and the weight of the node. Then use Equation (11) to quantify network security situation of the module:

$$MSA = \sum_{j=0}^{m} (NSA \times \omega_{node}). \quad (11)$$

Where ω_{node} is the weight of the node.

Finally, use Equation (12) to quantify the network security situation:

$$SA = \sum_{i=0}^{n} (MSA \times \omega_{module}). \quad (12)$$

Where ω_{module} is the weight of the module.

4 Experiment and Analysis

To verify the applicability of the proposed model, we selected the 2000 DARPA assessment dataset [15] provided by MIT Lincoln Lab datasets as experimental data. This dataset provides two attack scenarios LLDOS1.0 and LLDOS2.0.2 and contains network traffic and host audit logs, which can be used as a data source for the proposed model. In this paper, we will conduct an assessment of network security situation against these attack scenarios.

4.1 Experiment Environment

Due to there are many network nodes involved in the dataset, it is not clear enough if the complete network topology is drawn. Therefore, the network topology contains the key nodes only.

(a) predetect (b) detect

Figure 3: Network division

As shown in Figure 3, the figure on the left shows the network topology that contains the key nodes. We use the BGLL algorithm to divide the network, and the figure on the right shows the result. As you can see from the figure on the right, we divide the network into three modules, which are consistent with the dataset in which all the nodes are distributed in three regions: inside, outside, and in the DMZ. Besides these key nodes, we still need to consider the weight of those nodes that are not attacked in the network security situation assessment process.

Based on the information provided in the dataset, vulnerability information and the probability of success is shown in Table 2.

The service information of the key nodes in the network is shown in Table 3.

4.2 Node Security Situation Assessment Based on SimHash

Algorithm 2 is used to evaluate the security situation of the node. First, we write detection rules of the IDS to

Table 2: Network host vulnerability information

Vulnerability information	Mill	Locke	Pascal	Hume	Robin	af.mil	pro	impact
ICMP incorrectly configured	√	√	√	√	√	×	1.0	0.1
SunRPC incorrectly configured	√	√	√	×	×	×	0.8	0.2
Sadmind buffer overflow	√	√	√	×	×	×	0.8	0.8
RCP incorrectly configured	√	√	√	×	×	×	1.0	0.2
HINFO query incorrectly configured	√	×	×	×	×	×	0.8	0.6
SYN Flood	×	×	×	×	×	√	0.7	1.0

Table 3: Network hosting service information

Service information	Mill	Locke	Pascal	Hume	Robin	af.mil
HTTP	×	√	√	×	√	√
FTP	√	√	√	×	×	×
TELNET	√	√	√	×	×	×
DNS	√	×	×	×	×	×
SMTP	×	×	×	√	×	×
POP3	×	×	×	√	×	×

analyze the network traffic and get the alerts. Then upload these alerts and system logs to the distributed file system (HDFS) to storage. Because of the intrusion detection system alerts, some are found in the logs. Then we design MapReduce program to analyze the data on the HDFS to exclude duplicate Data, and finally get the attack information A, according to Table 2 to construct vulnerability information V. Taking A and V as the input data of the Algorithm 2, the security situation values of each node are generated. In order to ensure that the chart clearly, we only plot the security situation of the three nodes and shows in Figure 4.

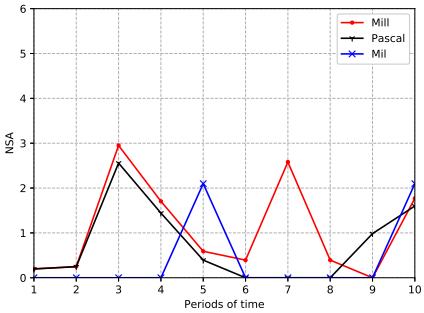


Figure 4: Node security situation

It can be seen from Figure 4 that in the periods of 1, 2 and 6, the nodes are scanned and the impact of scanning on the nodes is minimal, so we also calculated the NSA to be lower. In the period of 3, Mill and Pascal suffered a buffer overflow attack, so the value of NSA we calculated is higher. In the periods of 3 and 10, the attacker controls Mill and Pascal to initiate a DDoS attack on Mil node. Therefore, the NSA of the three nodes in both periods is greater. It can be seen that the algorithm we use to evaluate the node security situation is accurate, which can reflect the severity of each node being attacked.

4.3 Module Security Situation Assessment

We use the improved entropy method, according to the services provided by each node, get the weight of each node in the module, and the security situation of the module is calculated by Equation (11). There are 3 modules in our experiment: INSIDE, OUTSIDE, and DMZ. The result of the security situation of INSIDE module is shown in Figure 5.

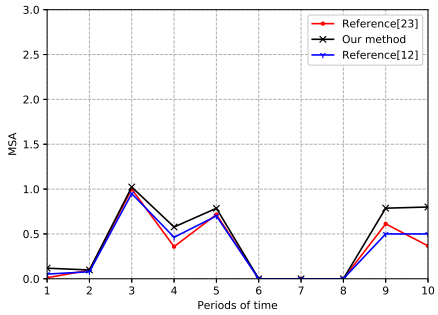


Figure 5: Module security situation assessment

There is no concept of module in [16] and [28]. Therefore, the results of these two methods are obtained by calculating the sum of the security situation of nodes which in INSIDE module. As can be seen from Figure 5, the trends of three methods are consistent. However, in some key stages, our method gets a higher value of the situation. For example, in the periods of 5 and 10, the attacker will have Pascal nodes. The attacker performs a DDoS attack on the Mil node through Pascal and mill node, and Pascal node is in the INSIDE module. Therefore, we hold

the view that the value of the INSIDE module is higher in these two periods, but in [28] it is lower in the period of 10 than before.

4.4 Network Security Situation Assessment

In the base of MSA, the security situation of the whole network is calculated using Equation (12), where the weight of the module is obtained by dividing the whole network using the BGLL algorithm. The network security situation obtained is shown in Figure 6, where the larger the SA is, the more insecure the network is.

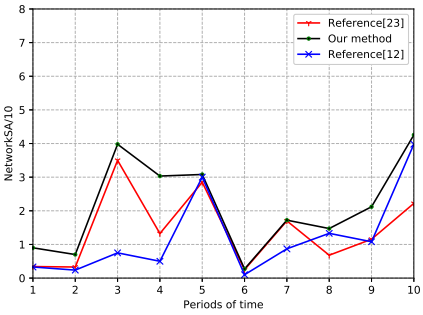


Figure 6: Network security situation assessment

It can be seen from the Figure 6 that our model can better reflect the security state of the network than [16] and [28]. As in the periods of 5 and 10, the network was attacked by a large number of distributed denial-of-service attacks, so our assessment results were relatively high for both periods. In the periods of 3 and 8, key nodes are compromised and the root access is taken by the attacker. At this point, the subsequent series of attacks are all based on root access, so both of these evaluations higher. In the end, network administrators can decide whether to take action or not based on the network security situation.

4.5 Performance Analysis

On the storage, we only need to store the hash values corresponding to the initial text and the modified text respectively. This saves a lot of space compared to storing network traffic and logs directly. And we can store these hashes in HDFS in the big data environment. Using the LZO compression algorithm to compress the data, which can save the disk space occupied by the data further and speed up the data transmission in the disk and the network, so as to improve the processing speed of the system. LZO compression algorithm allows us to split the compressed algorithm allows us to split the compressed file processing, file segmentation in the big data processing is very important. It will affect the number of parallel execution of the job, thus affecting the efficiency of the implementation of the job. Table 4 shows the comparison of several compression algorithms.

In the big data environment, the traditional situation assessment algorithm is more complex. When the data of the network node increases sharply, network status cannot be feedback to the network administrator timely and effectively. To test the efficiency of our algorithm, we randomly generate two different types of attacks in every period, and the number of each type also generate randomly. The length of hash value which is calculated by the SimHash algorithm is 64-bits. The result is shown in Figure 7. It can be seen that the time complexity of our algorithm is closed to $\mathcal{O}(n^2)$. Through the theoretical analysis of the algorithm, it can be seen that although there are several loops, only one is a two-layer loop, so the time complexity is $\mathcal{O}(n^2)$ is correct. The algorithm calculates the security situation of four million nodes takes only about 70 seconds. Moreover, this is just a node's computing power, we can dynamically increase the number of computing nodes if it is needed in the big data environment.

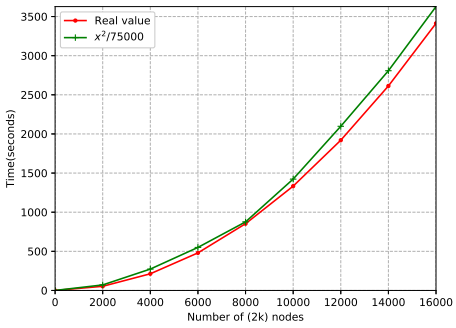


Figure 7: Calculating speed

Our model can be applied to a large-scale network, because using BGLL algorithm divides a large-scale network into multiple modules, as long as there is a topology, the topology can be determined when the network is generated. Alerts, a variety of logs and other network security-related data can be collected and uploaded to HDFS to storage and analysis as they are generated, so this process is generally done in parallel with data generation.

5 Conclusions

This paper analyzes and compares the existing evaluation methods of network security situation. To address the problem that these methods are difficult to adopt in a large-scale network environment, this paper proposes a network security situation assessment model based on text SimHash algorithm in the big data environment. The model divides a large-scale network into multiple modules by using the method of dividing the network structure of complex networks, and then analyzes the nodes in each module and quantifies the node security situation, the module security situation, and the network security situation gradually. Administrators know the status of network security at any time. And the experimental analysis

Table 4: Compression algorithm comparison

Compression algorithm	Initial file size	Compressed file size	Compression speed	Decompression speed	Separability
LZO	8.0GB	2.0GB	148.95MB/s	234.06MB/s	Yes
GZIP	8.0GB	1.3GB	33.99MB/s	113.78MB/s	No
BZIP2	8.0GB	1.06GB	6.13MB/s	24.5MB/s	Yes

verifies the applicability and characteristics of the evaluation model we proposed.

In the future, we will improve a large-scale network security situation assessment model and the quantitative assessment method, and on this basis, make a prediction of a large-scale network security situation. And designing a multidimensional visualization system to help administrators seize the status of network security more accurately.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (No.61370007), Program for New Century Excellent Talents of Fujian Provincial (No. 2014FJ-NCET-ZR06), and the Postgraduate Scientific Research Innovation Ability Training Plan Funding Projects of Huaqiao University (1611314012).

References

[1] A. A. Al-khatib, W. A. Hammood, "Mobile malware and defending systems: Comparison study," *International Journal of Electronics and Information Engineering*, vol. 6, no. 2, pp. 116–123, 2017.

[2] T. Bass, "Intrusion detection systems and multisensor data fusion," *Communications of the ACM*, vol. 43, no. 4, pp. 99–105, 2000.

[3] V. D. Blondel, J. L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2008, no. 10, pp. P10008, 2008.

[4] X. Z. Chen, Q. H. Zheng, X. H. Guan, and C. G. Lin, "Quantitative hierarchical threat evaluation model for network security," *Journal of Software*, vol. 17, no. 4, pp. 885–897, 2006.

[5] E. Doynikova and I. Kotenko, "Cvss-based probabilistic risk assessment for cyber situational awareness and countermeasure selection," in *25th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP'17)*, pp. 346–353, 2017.

[6] Y. Duan, Y. Cai, Z. Wang, and X Deng, "A novel network security risk assessment approach by combining subjective and objective weights under uncertainty," *Applied Sciences*, vol. 8, no. 3, pp. 428, 2018.

[7] M. R. Endsley, "Design and evaluation for situation awareness enhancement," *Proceedings of the Human Factors Society Annual Meeting*, vol. 32, no. 2, pp. 97–101, 1988.

[8] U. Franke and J. Brynielsson, "Cyber situational awareness–a systematic review of the literature," *Computers & Security*, vol. 46, pp. 18–31, 2014.

[9] V. Gorodetsky, O. Karsaev, and V. Samoilov, "Online update of situation assessment based on asynchronous data streams," in *Proceedings of the Knowledge Based Intelligent Information and Engineering Systems*, pp. 1136–1142, 2004.

[10] M. Henzinger, "Finding near-duplicate web pages," *Proceedings of the 29th annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'06)*, pp. 284-291, 2006.

[11] M. S. Hwang, C. T. Li, J. J. Shen, Y. P. Chu, "Challenges in e-government and security of information", *Information & Security: An International Journal*, vol. 15, no. 1, pp. 9–20, 2004.

[12] S. Islam, H. Ali, A. Habib, N. Nobi, M. Alam, and D. Hossain, "Threat minimization by design and deployment of secured networking model," *International Journal of Electronics and Information Engineering*, vol. 8, no. 2, pp. 135–144, 2018.

[13] Y. P. Jiang, C. C. Cao, X. Mei, and H. Guo, "A quantitative risk evaluation model for network security based on body temperature," *Journal of Computer Networks and Communications*, vol. 2016, pp. 3, 2016.

[14] S. J. Walker, "Big data: A revolution that will transform how we live, work, and think," *International Journal of Advertising*, vol. 33, no. 1, pp. 181–183, 2014.

[15] MIT Lincoln Lab, *2000 Darpa Intrusion Detection Scenario Specific Datasets*, 2000. (`https://www.ll.mit.edu/ideval/data/2000data.html`)

[16] C. Li and X. L. Shen, "Network security situation awareness model based on multi-period assessment," *Applied Mechanics and Materials*, vol. 411-414, pp. 613–618, 2013.

[17] 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 Network Security*, vol. 18, no. 4, pp. 650–666, 2016.

[18] Jie Liu, Ting Jin, Kejia Pan, Yi Yang, Yan Wu, and Xin Wang, "An improved knn text classification algorithm based on simhash," in *IEEE 16th Interna-*

tional Conference on Cognitive Informatics & Cognitive Computing (ICCI'17), pp. 92–95, 2017.

[19] L. Liu, Z. Cao, C. Mao, "A note on one outsourcing scheme for big data access control in cloud," *International Journal of Electronics and Information Engineering*, vol. 9, no. 1, pp. 29–35, 2018.

[20] M. E. J. Newman and M. Girvan, "Finding and evaluating community structure in networks," *Physical Review E*, vol. 69, no. 2, 2004.

[21] E. U. Opara and 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] Y. Qiao, X. Yun, and Y. Zhang, "Fast reused function retrieval method based on simhash and inverted index," in *IEEE Trustcom/BigDataSE/ISPA*, pp. 937–944, 2016.

[23] A. Tayal, N. Mishra and S. Sharma, "Active monitoring & postmortem forensic analysis of network threats: A survey," *International Journal of Electronics and Information Engineering*, vol. 6, no. 1, pp. 49–59, 2017.

[24] K. Vassakis, E. Petrakis, and I. Kopanakis, "Big data analytics: Applications, prospects and challenges," in *Mobile Big Data*, pp. 3–20, 2018.

[25] J. Webb, A. Ahmad, S. B. Maynard, and G. Shanks, "A situation awareness model for information security risk management," *Computers & security*, vol. 44, pp. 1–15, 2014.

[26] J. Wu, L. Ping, X. Ge, Y. Wang, and J. Fu, "Cloud storage as the infrastructure of cloud computing," in *International Conference on Intelligent Computing and Cognitive Informatics (ICICCI'10)*, pp. 380–383, 2010.

[27] H. P. Yang, H. Qiu, and K. Wang, "Network security situation evaluation method for multi-step attack," *Journal on Communications*, vol. 38, no. 1, pp. 187–198, 2017.

[28] W. Yong, L. Yifeng, and F. Dengguo, "A network security situational awareness model based on information fusion," *Journal of Computer Research and Development*, vol. 3, pp. 000, 2009.

[29] Qingyuan Zhou and Jianjian Luo, "The study on evaluation method of urban network security in the big data era," *Intelligent Automation & Soft Computing*, pp. 1–6, 2017.

Biography

Pengwen Lin graduate student. His main research direction is network security situation awareness.

Yonghong Chen prefessor. He mainly engaged in computer network and information security research, including Internet of things and security, cloud computing and security, intrusion detection, digital watermarking, big data security.

Comment on "Improved Secure RSA Cryptosystem (ISRSAC) for Data Confidentiality in Cloud"

Chenglian Liu[1,2] and Chieh-Wen Hsu[3]

(Corresponding author: Chenglian Liu)

Department of Computer Engineering, Dongguan Polytechnic Institute[1]
Dongguan 523808, China
(Email: chenglian.liu@gmail.com)
School of Computing, Neusoft Institute of Guangdong[2]
Foshan 528225, China
Economics and Management College, Zhaoqing University[3]
Zhaoqing 526061, China

(Received Apr. 15, 2018; Revised and Accepted Aug. 16, 2018; First Online Mar. 17, 2019)

Abstract

Data storage in the cloud is a very popular storage method because of the cost savings resulting from the user not needing hardware, software or space. However, data storage in the cloud has unique requirements because the current technical and environment. On March 2018, Thangavel and Varalakshmi proposed an "Improved Secure RSA cryptosystem (ISRSAC) for Data Confidentiality in Cloud". They modified the RSA algorithm into another one called MRSAC. In this paper, the author will examine the flaw from that scheme.

Keywords: Flaw; MRSAC; RSA

1 Introduction

Thangavel *et al.* [4] proposed the ESRKGS scheme in 2005, which stemmed from their RSA modification. The $\phi(n)$ of RSA [2] used two prime numbers, and Thangavel used four primes to generate its $\phi(n)$. Lüy *et al.* [1] discussed the vulnerability of Thangavel's scheme, and also gave an example. Thangavel and Varalakshmi [3] modified the ESRKGS algorithm to ISRSAC scheme. Unfortunately, there is a flaw in ISRSAC scheme, we would describe this situation on next section.

2 ISRASC Algorithm

RSA cryptosystem consists of three phases: Prime key generation, encryption and decryption. The problem is that RSA is not secure against a brute force attack. The security of RSA cryptosystem depends on the large prime number because it is difficult to break. Hence a modified version of RSA for secure key generation is used to generate the public and private keys. The resulting algorithm is known as ISRSAC. To start the algorithm, we randomly choose two large primes p and q where where $p \neq q, p > 3, q > 3$, and find

$$\begin{aligned} n &= p \cdot q \cdot (p-1) \cdot (q-1) \\ m &= p \cdot q. \end{aligned}$$

An integer r is randomly selected where $p > 2^r > q$, which generates $\alpha(n)$

$$\alpha(n) = \frac{(p-1)(q-1)(p-2^r)(q-2^r)}{2^r}$$

The public key e is satisfying $1 < e < \alpha(n)$ where $\gcd(e, \alpha(n)) = 1$

$$e \cdot d \equiv 1 \pmod{\alpha(n)}.$$

The public key pair is (e, n), and private key pair is (d, m).

Encryption: Suppose the M is plaintext or digitize message. We get ciphertext C by

$$C \equiv M^e \pmod{n}.$$

Decryption: We recovery message by

$$M \equiv C^d \pmod{m}.$$

3 Our Comment

3.1 The Parameter 2^r Problem

As known 2^r where

$$p > 2^r < q.$$

There are four types as follow:

Type 1: $p = 4n + 1$ and $q = 4m + 1$ forms.

Type 2: $p = 4n + 3$ and $q = 4m + 1$ forms.

Type 3: $p = 4n + 1$ and $q = 4m + 3$ forms.

Type 4: $p = 4n + 3$ and $q = 4m + 3$ forms.

Proof.

The Type 1. Since

$$4m + 1 < 2^r < 4n + 1,$$

we get

$$\begin{aligned} 4m + 1 < 2^r \quad &\Rightarrow \quad m < \frac{2_r - 1}{4} \\ 4n + 1 > 2^r \quad &\Rightarrow \quad n > \frac{2_r - 1}{4} \end{aligned}$$

The Type 2. Since

$$4m + 1 < 2^r < 4n + 3,$$

we get

$$\begin{aligned} 4m + 1 < 2^r \quad &\Rightarrow \quad m < \frac{2_r - 1}{4} \\ 4n + 3 > 2^r \quad &\Rightarrow \quad n > \frac{2_r - 3}{4} \end{aligned}$$

The Type 3. Since

$$4m + 3 < 2^r < 4n + 1,$$

we get

$$\begin{aligned} 4m + 3 < 2^r \quad &\Rightarrow \quad m < \frac{2_r - 3}{4} \\ 4n + 1 > 2^r \quad &\Rightarrow \quad n > \frac{2_r - 1}{4} \end{aligned}$$

The Type 4. Since

$$4m + 3 < 2^r < 4n + 3,$$

we get

$$\begin{aligned} 4m + 3 < 2^r \quad &\Rightarrow \quad m < \frac{2_r - 3}{4} \\ 4n + 3 > 2^r \quad &\Rightarrow \quad n > \frac{2_r - 3}{4} \end{aligned}$$

By Type 1 and Type 2, we get

$$m < \frac{2^r - 1}{4} \tag{1}$$

By Type 3 and Type 4, we get

$$m < \frac{2^r - 3}{4} \tag{2}$$

Suppose Equation (1) ∩ Equation (2), we know

$$m < \frac{2^r - 3}{4}$$

By Type 1 and Type 3, we get

$$n > \frac{2^r - 1}{4} \tag{3}$$

By Type 2 and Type 4, we get

$$n > \frac{2^r - 3}{4} \tag{4}$$

Suppose Equation (3) ∩ Equation (4), we obtain

$$n > \frac{2^r - 1}{4}$$

When $n > \frac{2^r-1}{4}$ and $m < \frac{2^r-3}{4}$ where $r \in \mathbb{Z}^+$, we get

$$q < 2^r < p.$$

□

3.2 The Core Algorithm

As known

$$\alpha(n) = \frac{(p-1)(q-1)(p-2^r)(q-2^r)}{2^r}.$$

p, q are both primes. If $r > 2, r \in \mathbb{Z}^+$, $\alpha(n)$ is not an integer.

Proposition 1. *From $\alpha(n)$ above, p, q are both primes. p or q can be written three forms as follows:*

Case 1. *$p = 4n + 3$ and $q = 4n + 3$ forms.*

Case 2. *$p = 4n + 1$ and $q = 4n + 3$ forms.*

Case 3. *$p = 4n + 1$ and $q = 4n + 1$ forms.*

We first consider Case 1.

$$\begin{aligned} (p-1)(q-1) &= (4n+2)(4m+2) \\ &= 4[4mn + 2(m+n) + 1]. \\ (p-2^r)(q-2^r) &= pq - 2^r p - 2^r q + 2^{2r}. \end{aligned}$$

$$\begin{aligned} \alpha(n) &= 4[4mn + 2(m+n) + 1](\frac{p \cdot q}{2^r} - p - q + 2^r) \\ &= \frac{pq4[4mn + 2(m+n) + 1]}{2^{r-2}} \\ &\quad -4(p + q - 2^r)[4mn + 2(m+n) + 1] \end{aligned}$$

since p, q and $[4mn+2(m+n)+1]$ are odd, therefore the product of $p \cdot q \cdot (4mn+2(m+n)+1)$ is also odd. This equation is not divisible by 2^{r-2} where 2^{r-2} is even. And

$$(p + q) - 2^r(4mn + 2(m+n) + 1)$$

is an integer. Thus, while r greater than $2, r \in \mathbb{Z}^+$, the $\alpha(n)$ is not an integer (a solution).

We then discuss Case 2, namely $p = 4n + 1$ and $q = 4m + 3$. Since

$$(p-1)(q-1) = 16mn + 8n = 8n(2m+1),$$

then

$$(p-2^r)(q-2^r) = pq - 2^r p - 2^r q + 2^{2r}.$$

We get

$$\begin{aligned}\alpha(n) &= 8n(2m+1)(\frac{pq}{2^r} - p - q + 2^r) \\ &= \frac{pqn(2m+1)}{2^{r-3}} - 8n(2m+1)(p+q-2^r)\end{aligned}$$

1) If n is odd since p, q, n and $(2m+1)$ are odd, the product of $p \cdot q \cdot n(2m+1)$ is also odd. We obtain

$$8n(2m+1)(p+1)(p+q-2^r)$$

is not divisible by 2^{r-3} if $r > 3$. Therefore, $\alpha(n)$ is not possible an integer.

2) If n is even where $n = a \cdot 2^i, i \in \mathbb{Z}^+$, a is odd. We get

$$\alpha(n) = \frac{p \cdot q \cdot a(2m+1)}{2^{r-i-3}} - 8n(2m+1)(p+q-2^r).$$

Since p, q, a and $(2m+1)$ are odd, the product of $p \cdot q \cdot a \cdot (2m+1)$ is also odd and is not divisible by 2^{r-i-3} when $r > i + 3, i \in \mathbb{Z}^+$.

By above Items 1) and 2), we know the $\alpha(n)$ is not an integer when $r > i + 3, i \in \mathbb{Z}_0^+$.

We keep discussing Case 3, namely $p = 4n + 1$ and $q = 4m + 1$.

$$\begin{aligned}(p-1)(q-1) &= 16mn \\ (p-2^r)(q-2^r) &= pq - 2^r p - 2^r q + 2^{2r} \\ \alpha(n) &= 16mn(\frac{p \cdot q}{2^r}) - p - q + 2^r \\ &= \frac{p \cdot q \cdot m \cdot n}{2^{r-4}} \\ &\quad -16mn(p+q-2^r).\end{aligned}$$

1) If m, n are odd, and p, q are odd, the product of $m \cdot n \cdot p \cdot q$ is odd and is not divisible by 2^{r-4}.

2) We start with $m \cdot n$ is even this is true under following condition m is odd and n is even, or m is even and n is odd.

Suppose

$$m \cdot n = b \cdot 2^j,$$

where b is odd, $j \in \mathbb{Z}^+$. We get

$$\alpha(n) = \frac{b \cdot p \cdot q}{2^{r-j-4}} - 16mn(p+q-2^r),$$

since b, p, q are odd, the product of $b \cdot p \cdot q$ is odd. Therefore it will not be divisible by 2^{r-j-4} when $r > 4 + j, j \in \mathbb{Z}^+$. From above Items 1) and 2), when $r > 4 + j$, $j \in \mathbb{Z}_0^+$, the $\alpha(n)$ is not an integer.

Summary of Cases 1, 2 and 3.

1) $p = 4n + 3$, $q = 4m + 3$, $\alpha(n)$ is not an integer when $r > 2, r \in \mathbb{Z}^+$.

2) $p = 4n + 1$, $q = 4m + 3$, $\alpha(n)$ is not an integer when $r > 3 + j, j \in \mathbb{Z}_0^+$.

3) $p = 4n + 1$, $q = 4m + 1$, $\alpha(n)$ is not an integer when $r > 4 + j, j \in \mathbb{Z}_0^+$.

4 Conclusion

Thangavel and Varalakshmi proposed ISRSAC scheme by RSA modification algorithm. From our comment, $\alpha(n) = \frac{(p-1)(q-1)(p-2^r)(q-2^r)}{2^r}$ where p and q are primes, and $r > 3, r \in \mathbb{Z}^+$, then the $\alpha(n)$ is not a integer. On their scheme, it is not possible to generate the public key e randomly. On other hand, it is also impossible satisfied $e \cdot d \equiv 1 \pmod{\alpha(n)}$ where $1 < e < \alpha(n)$ and $\gcd(e, \alpha(n)) = 1$. Therefore, the ISRSAC algorithm has a certain theoretical defect.

Acknowledgement

The authors would like to thank the anonymous reviewers for their useful comments. This work is partially supported from Neusoft Institute of Guangdong under the project number NUIT2018-001, and school project grant number 2018ZXB09 of Zhaoqing University. This work also partially supported by student innovation training program under the grant number PDJH2018B0579 and PDJH2019B0569.

References

[1] E. Lüy, Z. Y. Karatas, and H. Ergin, "Comment on "An enhanced and secured rsa key generation scheme (ESRKGS)"," *Journal of Information Security and Applications*, vol. 30, pp. 1–2, 2016.

[2] R. L. Rivest, A. Shamir, and L. Adleman, "A method for obtaining digital signatures and public-key cryptosystems," *Communincation of ACM*, vol. 21, pp. 120–126, Feb. 1978.

[3] M. Thangavel and Varalakshmi P., "Improved secure rsa cryptosystem (ISRSAC) for data confidentiality in cloud," *International Journal of Information Systems and Change Management*, 2018. In press.

[4] M. Thangavel, P. Varalakshmi, Mukund Murrali, and K. Nithya, "An enhanced and secured rsa key generation scheme (ESRKGS)," *Journal of Information Security and Applications*, vol. 20, pp. 3–10, 2015.

Biography

Chenglian Liu received his B.S degree in information management from National Union University in 1992

and the MSc degree in National Defense from National Defense University in 2004. He studied his doctorate course at Royal Holloway, University of London from 2006 to 2009 under the supervised by Chris Mitchell. He is with a distinguished associate professor at Huizhou University since 2014. His research interests are in Information Security, Number Theory and Goldbach's Conjecture so on.

Chieh-Wen Hsu received his B.S degree in department of mathamatics from Tamkang University in 1992 and the Master degree in National Cheng Kung University in 1994. He completed his Ph.D in 2010 in National Kaohsiung University of Science nd Technology in Taiwan. His research interests are in Probability and Statistics , Number Theory and Cryptanalysis so on.

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.