IJNS
International Journal of Network Security

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

Vol. 22, No. 4 (July 2020)

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

Volume: 22, No: 4 (July 1, 2020)

International Journal of Network Security

Sharing a Secret Image in the Cloud Using Two Shadows

Yu Chen[1], Jiang-Yi Lin[2,3], Chin-Chen Chang[3] and Yu-Chen Hu[4]
(Corresponding author: Chin-Chen Chang)

School of Information Science and Engineering, Fujian University of Technology[1]
33 Xuefu South Road, Fuzhou 350118, China
Department of Computer Science, Xiamen University of Technology[2]
600 Ligong Road, Xiamen 361024, China
Department of Information Engineering and Computer Science, Feng Chia University[3]
100 Wenhua Road, Taichung 40724, Taiwan
Department of Computer Science and Information Management, Providence University[4]
200, Section 7, Taiwan Boulevard, Shalu District, Taichung 43301, Taiwan
(Email: alan3c@gmail.com)
(Received Mar. 20, 2018; Revised and Accepted Sept. 7, 2018; First Online Feb. 20, 2020)

Abstract

In this paper, we present a novel (2, 2) reversible secret image sharing scheme. Our scheme permits secret messages to be shared with two participants by splitting the marked encrypted image into two shadows. The secret messages can be reconstructed if two participants collaborate with each other. The proposed scheme chooses suitable binary blocks of a cover image in which to embed the secret message and divides those blocks into two shadow blocks by executing a logical operation with all of the other binary blocks, thereby producing two shadows. In the data extraction procedure, the secret messages and the cover image can be reconstructed by the logical operation of the corresponding binary blocks of the two shadows. A practical application is demonstrated by modeling our scheme as a reversible watermarking scheme in the Cloud. The experimental results indicated that the proposed method is reversible and that it can restore the image and watermark properly.

Keywords: Data Hiding; Reversible Watermarking; Secret Image Sharing

1 Introduction

It is very important to secure information [1,12,14,15,18] in today's information age. Secret sharing is an effective approach to protect the security of information by sharing parts of the data with different holders to avoid leaking useful information. In 1979, Shamir and Blakley independently introduced the concept of secret sharing and proposed two (t, m) threshold schemes [1, 15]. In Shamir's scheme [15], a dealer divides a secret, D, into m pieces, which are kept by a group of m users; Then, t or more users can collaborate to recover D, and less than t users cannot restore D. Inspired by these (t, m) threshold schemes [3, 5, 10, 13, 16], many researchers focused on the study of secret sharing. In 1995, Naor and Shamir extended the (t, m) threshold scheme to secret image sharing and proposed the concept of visual secret sharing (VSS) [13]. Although their scheme was a novel method for sharing secrets, it applies only to binary images and incurs the pixel expansion problem.

In the past decades, as an important direction for secret message delivery applications, some schemes [2,4,6–9,11] have been proposed. Chang *et al.* introduced a secret image sharing scheme [2] in 2008. In their scheme, a magic matrix was used to modify the cover image in order to embed the secret digits. Their scheme can completely restore a cover image after the secret digits are extracted. In 2009, Lee *et al.* proposed a reversible data hiding scheme [9] in which two steganographic images were used. According to their scheme [9], a cover pixel pair is changed at most by one when two secret bits are embedded. Therefore, a high steganographic image quality is provided. In 2015, Lu *et al.* proposed a dual-image-based data hiding method [11]. In their study, the center folding strategy was used to reduce the value of the secret symbols to obtain the folded secret information that was embedded in the two images. Recently, Chang and Liu presented a novel (2, 2) secret sharing method [7]. In their research, the balance between the quality and the payload of the shadows can be achieved easily by adjusting a control parameter. Compared with the methods of Lee *et al.* [9] and Lu *et al.* [11], Chang and Liu's scheme [7] achieved the highest payload and best flexibility.

Inspired by the above schemes, we present a novel (2, 2) reversible secret sharing scheme. The proposed scheme uses the logical operation on bit-planes [17] to embed secret data, such as the watermark, and then splits the data into two shares. Thus, our scheme generates two meaningful images, called shadows. Furthermore, we applied the proposed secret image sharing scheme to a reversible watermarking scheme in the Cloud to demonstrate its validity.

The rest of this paper is organized as follows. Section 2 reviews Chang and Liu's scheme. Section 3 presents the proposed reversible secret sharing scheme. Section 4 provides our experimental results, and Section 5 gives our conclusions.

2 Brief Introduction of Chang and Liu's Scheme

Chang and Liu's scheme [7] is a reversible (2, 2) secret image sharing method. Their scheme consists of two procedures, *i.e.*, the secret sharing procedure and the secret and image reconstructing procedure. Here, the secret image and the cover image are both grayscale images. The two procedures are described in the following subsections.

2.1 Secret Sharing Procedure

In this process, a grayscale secret image is embedded into a grayscale cover image P. Assume that P with size of $H \times W$ is expressed as $P = \{p_i | i = 1,\ 2,\ \ldots,\ (H \times W)\}$, where p_i is the i-th pixel of P and $p_i \in [0, 255]$. Let G be a binary stream representing a secret image, with the length of $|G|$. G is split into k segments, and each segment, c_q, is the size of ω bits whose value is in the range of 0 to $2^\omega - 1$. Here ω is considered as a control parameter. Thus, G can be represented as $G = \{c_q | q = 1,\ 2,\ \ldots,\ k\}$, where $k = \lceil |G| / \omega \rceil$. Two shadow images, P_1 and P_2, which are the same size as P, are generated after G is embedded. The detailed steps are as follows:

Step 1: Set a value of ω.

Step 2: Sequentially read each segment c_q of ω bits from G.

Step 3: Embed c_q into p_q to produce two shadow pixels p_{q1} and p_{q2}. Let $\omega^* = 2^\omega/2$, and then perform different embedding processes for $c_q \le \omega^*$ or $c_q > \omega^*$.

Step 3.1: Embed c_q for $c_q \le \omega^*$. There are three cases to deal with.

Case 1: If $0 \le p_q \pm c_q \le 255$, p_{q1} and p_{q2} are computed as

$$p_{q1} = p_q - c_q, \quad (1)$$
$$p_{q2} = p_q + c_q. \quad (2)$$

Case 2: If $p_q + c_q > 255$, p_{q1} and p_{q2} are computed as

$$p_{q2} = p_q, \quad (3)$$
$$p_{q1} = p_q - (2e - 1), \quad (4)$$

where $e = p_q + c_q - 255$.

Case 3: If $p_q - c_q < 0$, p_{q1} and p_{q2} are computed as

$$p_{q1} = p_q, \quad (5)$$
$$p_{q2} = p_q + (2f - 1), \quad (6)$$

where $f = 0 - (p_q - c_q)$.

Step 3.2: Embed c_q for $c_q > \omega^*$. Firstly, set $c_q = c_q - \omega^*$ to satisfy $c_q < \omega^*$. Secondly, perform Step 3.1 to obtain p_{q1} and p_{q2}. Finally, swap p_{q1} and p_{q2}.

Step 4: Repeat Step 2 and Step 3 until G is fully processed and P_1 and P_2 are generated.

Step 5: Give P_1 to one receiver and give P_2 to another receiver.

In the above steps, ($2e$-1) and ($2f$-1), which appear in Cases 2 and 3, respectively, are set to be odd in order to identify the overflow condition, which will be described in detail in the subsequent extraction process. Setting $c_q = c_q - \omega^*$ in Step 3.2 is to decrease distortion. Then, p_{q1} and p_{q2} are swapped to satisfy $p_{q1} \ge p_{q2}$.

Example 1. *Assume that a binary stream* S= *'00111101' is part of the secret image* G *and that* $p_1 = 80$ *and* $p_2 = 253$ *are two pixels of the cover image* P. *Next is how to embed* S *into* p_1 *and* p_2. *First, we set* $\omega = 4$ *and* $\omega^* = 2^\omega/2 = 8$. *So,* S *is separated into two segments* c_1 *and* c_2, *both of which are 4 bits,* i.e., $c_1 = (0011)_2 = 3$ *and* $c_2 = (1101)_2 = 13$.

1) Embed $c_1 = 3$ into $p_1 = 80$: Since $c_1 < \omega^*$ and $0 \le p_1 \pm c_1 \le 255$, it is Case 1. And we apply Equations (1) and (2) to produce two shadow pixels $p_{11} = p_1 - c_1 = 80 - 3 = 77$ and $p_{12} = p_1 + c_1 = 80 + 3 = 83$.

2) Embed $c_2 = 13$ into $p_2 = 253$: Since $c_2 > \omega^*$, the process turns to Step 3.2 to set $c_2 = c_2 - \omega^* = 13 - 8 = 5$. Then, the process turns to Step 3.1. Since $p_2 + c_2 = 253 + 5 > 255$, it is Case 2, and Equations (3) and (4) are applied to compute p_{21} and p_{22} where $p_{22} = p_2$ and $p_{21} = p_2 - (2e-1) = p_2 - (2 \cdot (p_2 + c_2 - 255) - 1) = 248$. Finally, p_{22} and p_{21} are swapped.

2.2 Secret and Image Reconstructing Procedure

Let p_{j1} and p_{j2} be the corresponding pixels of the two shadow images P_1 and P_2, respectively. The secret image and the cover image can be reconstructed by the following steps:

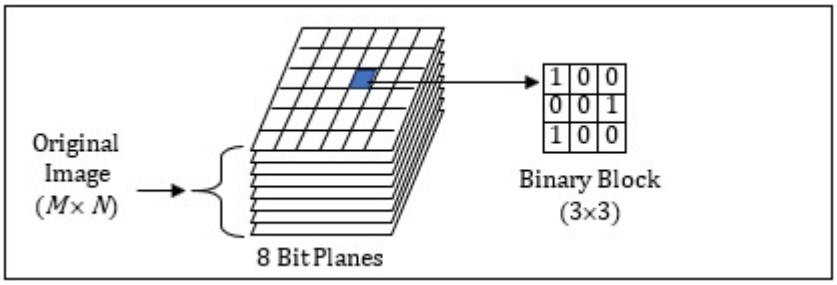


Figure 1: Schematics of bit-planes and a binary block

Step 1: Swap p_{j1} and p_{j2} if $p_{j1} > p_{j2}$.

Step 2: Reconstruct secret segment c_q and cover pixel p_q. The reconstruction process consists of the following three cases:

Case 1: If $(p_{q1} + p_{q2}) \bmod 2 = 0$,

$$p_q = \frac{p_{q1} + p_{q2}}{2}, \tag{7}$$

$$c_q = p_q - p_{q1} \text{ or } c_q = p_{q2} - p_q. \tag{8}$$

Case 2: If $(p_{q1}+p_{q2}) \bmod 2 \neq 0$ and $p_{q2}+\omega^* > 255$,

$$p_q = p_{q2}. \tag{9}$$

According to Equations (3) and (4), c_q can be calculated as

$$c_q = 255 + \frac{p_{q2} - p_{q1} + 1}{2} - p_{q2}. \tag{10}$$

Case 3: If $(p_{q1} + p_{q2}) \bmod 2 \neq 0$ and $p_{q1} - \omega^* < 0$,

$$p_q = p_{q1}.$$

According to Equations (5) and (6), c_q can be calculated as

$$c_q = \frac{p_q + p_{q2} + 1}{2}.$$

Step 3: Set $c_q = c_q + \omega^*$ if Step 1 is executed.

Step 4: Repeat Steps 1-3 until all the pixels of the shadow images are processed.

In this scheme, the parity of a number is used to distinguish whether or not it is overflow. The order of the values of two shadow pixels is employed to determine whether a normal value or a processed value was embedded.

Example 2. *Use the two pairs of shadow pixels generated in Example 1,* $(p_{11} = 77, p_{12} = 83)$ *and* $(p_{21} = 253, p_{22} = 248)$*, to demonstrate the reconstructing procedure for the secret and the cover images.*

1) Use p_{11} and p_{12} to restore c_1 and p_1. Notice that $p_{11} < p_{12}$, we can get $c_1 \leq \omega^*$ based on the method used in the secret sharing procedure. Since $(p_{11} + p_{12}) \bmod 2 = 0$, Equations (7) and (8) are applied to reconstruct p_1 as $p_1 = (p_{11} + p_{12})/2 = (77 + 83)/2 = 80$ and c_1 as $c_1 = p_1 - p_{11} = 80 - 77 = 3 = (0011)_2$.

2) Use p_{21} and p_{22} to restore c_2 and p_2. Here $p_{21} > p_{22}$ means that $c_2 > \omega^*$, and the process turns to Step 1. Also, because $(p_{21} + p_{22}) \bmod 2 \neq 0$ and $p_{22} + \omega^* > 255$, that is Case 2, and Equations (9) and (10) are used to calculate p_2 and c_2. Here $p_2 = p_{22} = 253$ and $c_2 = 255 + \frac{p_{22} - p_{21} + 1}{2} - p_{22} = 255 + \frac{253 - 248 + 1}{2} - 253 = 5$. Finally, adjust c_2 to $c_2 + \omega^* = 5 + 8 = 13 = (1101)_2$. Accordingly, we obtain the binary secret stream S = '00111101' and the cover image pixels are $p_1 = 80$ and $p_2 = 253$.

3 Proposed Secret Image Sharing Scheme

Inspired by Chang and Liu's scheme [7], we propose a novel (2, 2) secret image sharing scheme using a model in which an image is watermarked and shared by the Cloud. The bit-planes-based technique is used for image processing in the proposed scheme, and it is introduced as follows. An original grayscale image of 8-bit resolution can be decomposed into eight bit-planes in such a way that we can perform image processing at the bit-level. Yi *et al.* [17] embedded the bits in the lower bit-planes of an original image into the higher bit-planes, allowing the lower bit-planes to be reserved for hiding secret data later.

The logical operation is performed on the binary blocks in the bit-planes and the binary blocks with certain characteristics are selected to embed the watermark. Then, the embedded bit-planes are combined to form a stego-image. A schematic of the bit-planes of an image and a binary block is shown in Figure 1.

After receiving an image, the Cloud embeds a watermark in it, produces two shadows, S_1 and S_2, of the marked image, and then distributes them to the two recipients. By incorporating S_1 and S_2, we can extract the watermark and recover the original image. The procedures in our proposed scheme are shown schematically in Figure 2.

3.1 Watermark Embedding and Shadows Generating

To prevent revealing any information of an original image I, the proposed method encrypts I by processing the

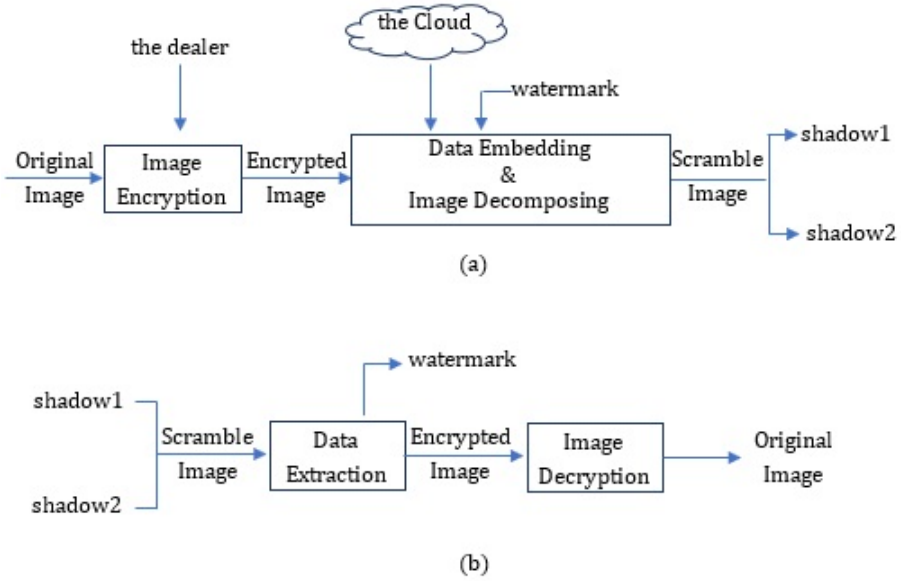


Figure 2: Procedures of (a) Secure image sharing; (b) Secret data retrieving and image restoring

encryption procedure, and only the encrypted image will be sent to the Cloud. Let E be an encrypted image. $E = \{r_l | l = 1, 2, \ldots, (M \times N)\}$, where $r_l \in [0, 255]$. Now we decompose E into eight bit-planes, and every bit-plane then is divided into a set of non-overlapping 3×3 binary-blocks. For each block, let b be the central bit of the block. Let n and n' be the total numbers of the remaining bits in this block excluding b that are equal to b or not, respectively. Based on the result of the comparison of n and n', all the blocks are classified into two categories, *i.e.*, the Nice block where $n > n'$ and the Bad block where $n \leq n'$.

The Nice blocks are selected and the central bits of them will be modified to embed the watermark. Since the Cloud will produce two shadows of E, *i.e.*, S_1 and S_2, for the two receivers, the proposed scheme utilizes different combinations of the values of the central bits of the blocks in the corresponding S_1 and S_2 to denote that 0 or 1 has been embedded. Next, we give an example of the Nice block and the Bad block and the structure of the shadow block in Figure 3.

The Nice blocks will be chosen to embed the watermark. In the proposed scheme, we generate two corresponding shadow blocks for one block of the cover image. The central point bits of a block and its corresponding shadow blocks are designed and shown in Table 1. The column L_1 indicates which number in a block is more. The column L_2 represents the combination of b_1 and b_2, the central bits of the two corresponding shadow blocks. According to the columns L_1 and L_2, the proposed scheme gives four scenarios, respectively, against $b = 0$ or $b = 1$.

Let B be a 3×3 Nice block divided from the bit-planes of the cover image, whose structure is shown in Figure 4(a). Let SB_1 and SB_2 be two corresponding shadow blocks to B, and distribution of their bits is shown in Figures 3 (c) and (d), respectively. The two shadow

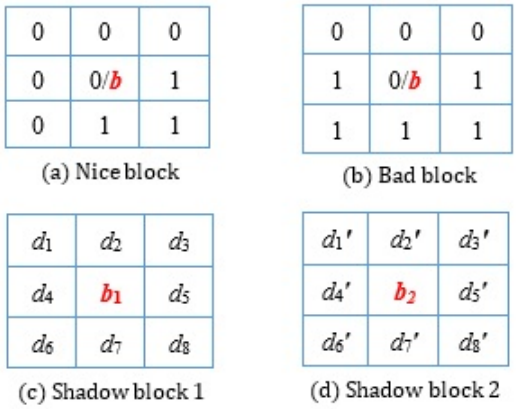


Figure 3: An example of the blocks in encrypted image and shadow image where the value of the central point, b, is 0; (a) The Nice block for $n = 5$ and $n' = 3$; (b) The Bad block with $n = 3$ and $n' = 5$; (c) and (d) The bits structure of shadow blocks corresponding to (a).

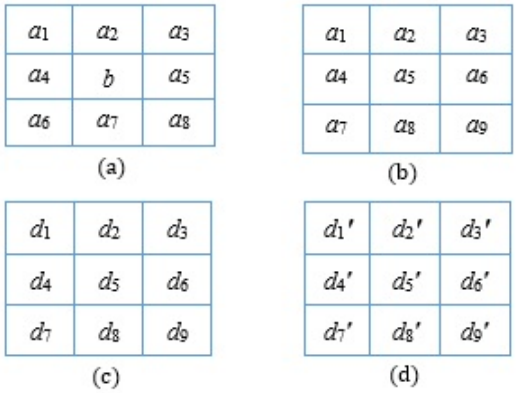


Figure 4: Examples of the structure of the blocks

Table 1: Central point bits of the binary block and its corresponding shadow blocks

$b=0$			$b=1$		
Majority bit L_1	b_1b_2 L_2	Description L_3	Majority bit L_1	b_1b_2 L_2	Description L_3
0	00	embedding 0	0	11	bad block
0	01	embedding 1	1	01	no embedding
0	10	no embedding	1	10	embedding 0
1	00	bad block	1	11	embedding 1

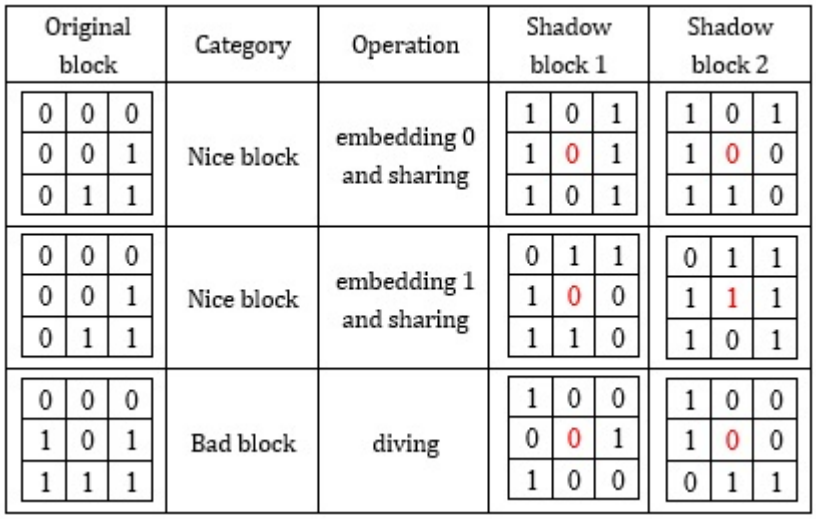


Figure 5: Examples of an original block being divided into two shadow blocks

blocks SB_1 and SB_2 must satisfy $d_i \oplus d'_i = a_i$, where $i = 1, 2, \ldots, 8$, $\oplus$ represents the XOR operation, and (b_1, b_2) must satisfy the conditions in Table 1 for embedding zero and one. Similarly, for the Bad block B is equal to Figures 4 (b). Let the two shadow blocks SB_1 and SB_2 be equal to Figures 4 (c) and (d), respectively. And SB_1 and SB_2 must satisfy $d_i \oplus d'_i = a_i$, where $i = 1, 2, \ldots, 9$.

As we can see, d_i can be replaced by 0 or 1, so can d'_i. Thus, d_i and d'_i are unfixed for a shadow block. After all the binary-blocks are processed as shown above, the Cloud produces two shadows of the encrypted image and delivers them to the two receivers. However, they only can obtain the original block of the encrypted image by the cooperation of the two shadow blocks, since neither of the single shadow blocks can give the original block. Therefore, the proposed scheme allows the secure sharing of secret images.

Example 3. *The following example illustrates how to generate two shadow blocks for embedding a single bit in a given Nice block with the situation that the central bit equal zero. Figure 5 lists three cases of the process. Similarly, for the case of the central bit being one, we can also list like this according to the conditions in Table 1.*

3.2 Watermark Extraction and Image Recovering

After embedding the watermark in the encrypted image, the Cloud delivers two shadows, S_1 and S_2 of the encrypted image to two different receivers, R_1 and R_2. After being permitted, R_1 and R_2 can extract the watermark and restore the image to the state that the Cloud accepted, and they can restore the image to original image by decryption. The process is as follows. First, two shadows will be decomposed into 8 bit-planes, and, then, each bit-plane will be divided into a set of non-overlapping 3×3 binary blocks. Second, we perform the XOR operation on a pair of blocks from corresponding bit-plane pairs from the two shadows. Finally, according to the state of the two central point values and the operating result blocks, we can obtain the original block without any embedded secret bits.

Example 4. *Use two shadow blocks to perform the restoring and extracting process. We can use the two shadow blocks generated in Example 3. Three examples of obtaining the original blocks and secret bits are shown in Figure 6.*

First, the XOR operation is performed on two shadow blocks to get a result block, where their central bits are labeled as b_1 and b_2. Second, determine the number, d,

Shadow block 1	Shadow block 2	XORing result	Majority bit	Original block	Secret bit
1 0 1 1 0 1 1 0 1	1 0 1 1 0 0 1 1 0	0 0 0 0 0 1 0 1 1	0	0 0 0 0 0 1 0 1 1	0
0 1 1 1 0 0 1 1 0	0 1 1 1 1 1 1 0 1	0 0 0 0 1 1 0 1 1	0	0 0 0 0 0 1 0 1 1	1
1 0 0 0 0 1 1 0 0	1 0 0 1 0 0 0 1 1	0 0 0 1 0 1 1 1 1	1	0 0 0 1 0 1 1 1 1	null

Figure 6: Examples of the restoration of the original block and the extraction of secret bits

that appears more often except for the central point. Finally, according to columns L_1 and L_2 in Table 1, we can get the original block and the secret bit by comparing the relationships between L_1 and d and between L_2 and the connection string of b_1 and b_2. Therefore, in the first row of Figure 6, the central bit of the original block and the secret bit are both 0 when $d = 0$ and the concatenation of b_1 and b_2 is '00'. By proceeding with all of the blocks in the two shadows in this way, the encrypted image that was sent to the Cloud is restored, and the watermark is extracted.

4 Experimental Results

In this section, some experiments were conducted on some test images to evaluate the correctness of the proposed scheme. Our experiments were conducted by using software MATLAB R2012b running on a personal computer whose operation system is Windows 10. The CPU of the computer is Intel Xeon E3-1225 v5, 3.3GHz, and the memory is 8GB. Focusing on the method related to secret sharing rather than encrypting images, the Henon map, an extensively-used, easily-implemented method, was used as the simulation to encrypt our test image. The Henon map is a two-dimensional, non-linear map, and it is defined as follows:

$$\begin{cases} x_{n+1} = 1 - ax_n^2 + y_n \\ y_{n+1} = bx_n \end{cases} \quad (11)$$

In our experiments, we set $x_1 = -0.4$, $y_1 = -0.4$, $a = 1.2$, $b = 0.3$, and these four values represent the secret key. By using Equation (11), two one-dimensional chaotic maps can be produced for use in creating a transform matrix to change the pixel locations of the original image and to obtain the shuffled image, *i.e.*, the encrypted image. Two test grayscale images with sizes of 512×512 were encrypted by using a Henon map with the above parameters to generate two encrypted images, while two grayscale watermark images were embedded into the encrypted images, respectively. The two watermark images we used are shown in Figure 7.

(a) watermark image 1 with size 57×57

(b) watermark image 2 with size 102×102

Figure 7: Two test watermark images

Figure 8 shows the test image "Lena" and its encrypted image and the corresponding two sets of shadow images. Figure 9 shows the case in which the test image was "Baboon". Then, in the reconstruction process, we used the two cooperative shadow images generated above to extract the logos and reconstruct the encrypted images successfully. We also used Equation (11) to restore the encrypted image exactly to its original state after obtaining the secret key used for encryption. We also used additional images to test the payload of the proposed scheme.

To illustrate the security of the proposed scheme, we will explain it from two different perspectives. One is the computational cost, and the other is the most common quantities, *i.e.*, the number of pixels change rate (*NPCR*)

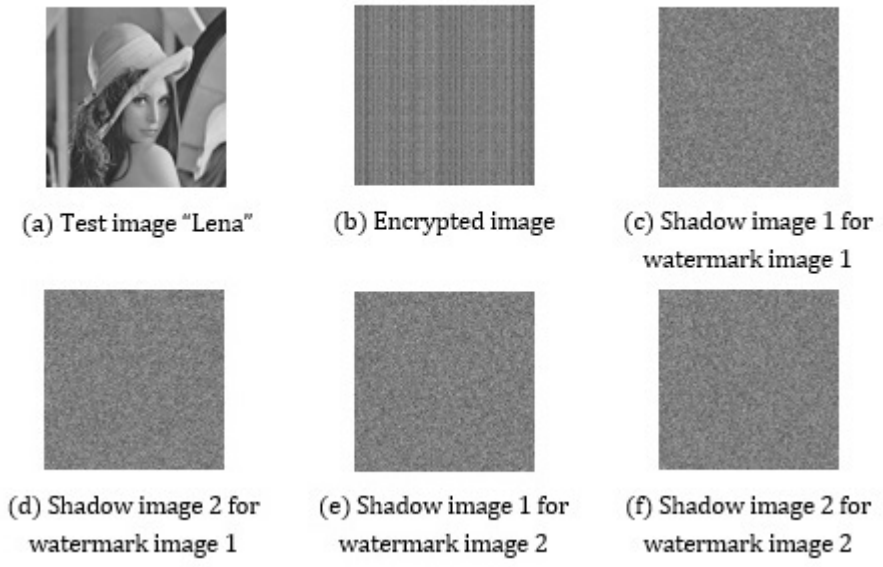


Figure 8: Example of the proposed scheme using the test image "Lena"

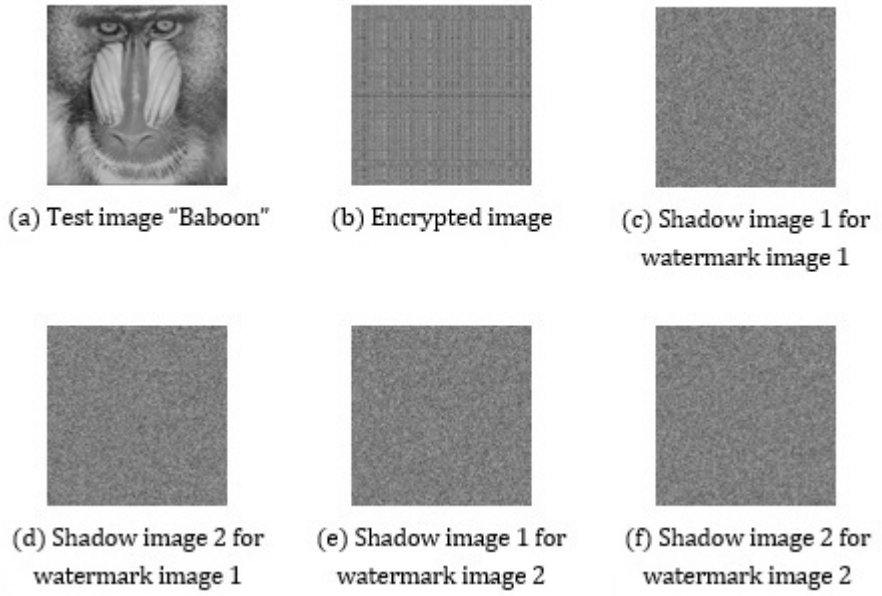


Figure 9: Example of the proposed scheme using the test image "Baboon"

and the unified average changing intensity ($UACI$). Assume that an attacker gets one shadow image, which won't leak any secret information. As described in Section 3 of the shadow images generation process, if he/she applies a brute-force attack to obtain the other shadow image, the possibility of success can be calculated as:

$$pb = \frac{1}{2^{M\times N\times 8}} = \frac{1}{256^{M\times N}}.$$

Obviously, this is very hard to accomplish. The computational cost of our method is mainly the XOR operation of the bit matrix in the image, which is linear, so the computational complexity of the proposed scheme is low. Next, we look at the second perspective, $NPCR$ and $UACI$. Generally, a high $NPCR/UACI$ score is interpreted as a stronger anti-attack performance, and they are the most common standardized tests for the security of an image. These two quantities were used in our experiment to test the two shadow images. Let us assume that we tested the two shadow images A^1 and A^2, respectively; the pixel value at corresponding positions were denoted as $A^1(i,j)$ and $A^2(i,j)$, and an array F is defined in Equation (12). Then, the $NPCR$ and $UACI$ are defined by Equations (13) and (14), respectively, where M and N are the width and height of the shadow images, L is the maximum pixel value, which is 255 for a grayscale image.

$$F(i,j) = \begin{cases} 0, \text{ if } A^1(i,j) = A^2(i,j) \\ 1, \text{ if } A^1(i,j) \neq A^2(i,j). \end{cases} \quad (12)$$

$$NPCR(A^1,A^2) = \sum\nolimits_{i,j} \frac{F(i,j)}{M\times N} \times 100\%. \quad (13)$$

$$UACI(A^1,A^2) = \sum\nolimits_{i,j} \frac{|A^1(i,j) - A^2(i,j)|}{M\times N\times L} \times 100\%. \quad (14)$$

The ranges of $NPCR$ and $UACI$ are all [0, 1]. If $NPCR(A^1,A^2) = 0$, it means that all of the pixels in A^2 have the same value as in A^1, and, if $NPCR(A^1,A^2) = 1$, then all of the corresponding pixels have different values in A^1 and A^2. Obviously, the ideal value for $NPCR$ is close to 1, but, for $UACI$, it is not obvious that the better value also is close to 1. However, it generally is believed that the expected $UACI$ value of a grayscale image is about 33%. Based on Eqs. 12- 14, we calculate the $NPCR$ and $UACI$ scores of the four sets of shadow images in the previous examples, and the data are shown in Table 2. It can be observed that the $NPCR$ scores of the shadow images are greater than 99.8%. It indicates that the two shadow images in each pair differ greatly. In addition, $UACI$ scores of the shadow images are higher than 34%. It is obvious that the average changing intensity of the corresponding pixels in the shadow images is very strong.

In terms of image embedding capacity, in addition to the previous two test images, we also tested several other images in our experiment. Figure 10 shows the other four test images and their encrypted images. Table 3 shows the payloads of all six encrypted images, from which it can be seen that the payload of the encrypted aircraft image reaches the maximum of 117805 bits and the encrypted baboon image has the smallest payload of 87272 bits. Thus, the experimental results show that proposed scheme can successfully share a secret image. It is a good method for the watermarking in the Cloud and achieving the secure sharing of secret information.

5 Conclusions

In this paper, a novel (2, 2) reversible secret sharing scheme is proposed, and the scheme was demonstrated by an application model of reversible watermarking in the Cloud. In this scheme, first, the original image is encrypted before uploading to the Cloud so that it cannot be leaked to any third party. Second, two shadows of the encrypted image with watermark information are generated. The test results in Table 2 indicate that $NPCR$ score and $UACI$ score are ideal. And from the experimental results in Table 3, we find that the average payload of these six encrypted images is 99185 bits. Finally, only through the collaboration of the two recipients can the exact watermark be determined and the encrypted image be obtained. The original image can be recovered without any damage, if desired.

References

[1] G. R. Blakley, "Safeguarding cryptographic keys," in *Proceedings of National Computer Conference, American Federation of Information Processing Societies*, pp. 313–317, June 1979.

[2] C. C. Chang, Y. C. Chou, and T. D. Kieu, "An information hiding scheme using sudoku," in *Proceedings of The Third International Conference on Innovative Computing Information and Control (ICICIC'08)*, pp. 17–21, June 2008.

[3] C. C. Chang, Y. P. Hsieh, and C. H. Lin, "Sharing secrets in stego images with authentication," *Pattern Recognition*, vol. 41, no. 10, pp. 3130–3137, 2008.

[4] C. C. Chang, T. D. Kieu, and Y. C. Chou, "Reversible data hiding scheme using two steganographic images," in *Proceedings of IEEE (TENCON'07)*, pp. 1–4, Oct. 2007.

[5] C. C. Chang, C. C. Lin, and N. T. Huynh, "Safeguarding visual information using (t, n) verifiable secret shares," *Journal of Computers*, vol. 22, no. 2, pp. 72–88, 2011.

[6] C. C. Chang, C. C. Lin, T. H. N. Le, and H. B. Le, "Sharing a verifiable secret image using two shadows," *Pattern Recognition*, vol. 42, no. 11, pp. 3097–3114, 2009.

[7] C. C. Chang, Y. J. Liu, and H. L. Wu, "Distortion-free secret image sharing method with two meaningful shadows," *IET Image Processing*, vol. 10, no. 8, pp. 590–597, 2016.

Table 2: Quantitative data for each pair of shadow images that were tested

Factor	Lenna with watermark image 1	Lenna with watermark image 2	Baboon with watermark image 1	Baboon with watermark image 2
$NPCR(\%)$	99.9722	99.8539	99.9626	99.8169
$UACI(\%)$	34.1083	34.1208	34.1164	34.1255

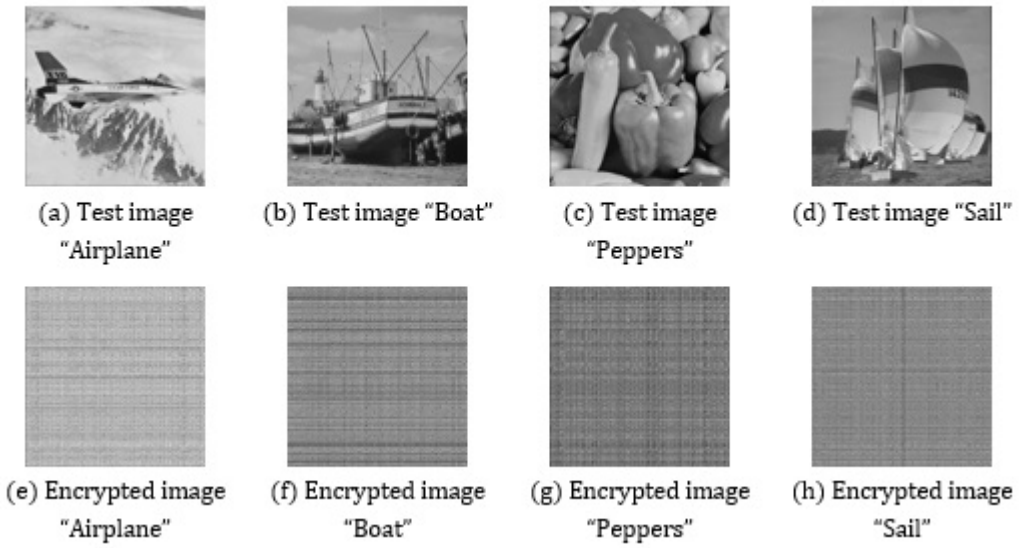


Figure 10: Examples of the test images and the corresponding encrypted images

Table 3: Performances of the test images

Factor	Encrypted Airplane	Encrypted Baboon	Encrypted Boat	Encrypted Lena	Encrypted Peppers	Encrypted Sail
Payloads(bits)	117805	87272	109139	90758	89269	100863

[8] C. F. Lee and Y. L. Huang, "Reversible data hiding scheme based on dual stegano-images using orientation combinations," *Telecommunications Systems*, vol. 52, no. 4, pp. 2237–2247, 2013.

[9] C. F. Lee, K. H. Wang, C. C. Chang, and Y. L. Huang, "A reversible data hiding scheme based on dual steganographic images," in *Proceedings of The Third International Conference on Ubiquitous Information Management and Communication*, pp. 228–237, Jan. 2009.

[10] C. C. Lin and W. H. Tsai, "Secret image sharing with steganography and authentication," *The Journal of Systems and Software*, vol. 73, no. 3, pp. 405–414, 2004.

[11] T. C. Lu, J. H. Wu, and C. C. Huang, "Dual-image-based reversible data hiding method using center folding strategy," *Signal Processing*, vol. 115, pp. 195–213, 2015.

[12] K. Ma, W. Zhang, X. Zhao, N. Yu, and F. Li, "Reversible data hiding in encrypted images by reserving room before encryption," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 3, pp. 553–562, 2013.

[13] M. Naor and A. Shamir, "Visual cryptography," *Lecture Notes in Computer Science*, vol. 950, pp. 1–12, 1995.

[14] Z. Qian and X. Zhang, "Reversible data hiding in encrypted images with distributed source encoding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 26, no. 4, pp. 636–646, 2016.

[15] A. Shamir, "How to share a secret," *Communications of the Association for Computing Machinery*, vol. 22, no. 11, pp. 612–613, 1979.

[16] C. C. Thien and J. C. Lin, "An image-sharing method with user-friendly shadow images," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13 no. 12, pp. 1161–1169, 2003.

[17] S. Yi and Y. Zhou, "Binary-block embedding for reversible data hiding in encrypted images," *Signal Processing*, vol. 133, pp. 40–51, 2017.

[18] X. Zhang, "Reversible data hiding in encrypted image," *IEEE Signal Processing Letters*, vol. 18, no. 4, pp. 255–258, 2011.

Biography

Yu Chen received the B.S. degree in Computer and Application from Hunan University, Hunan, China in 1993, and M.S. degree in Software Engineering from Fuzhou University, Fujian, China, in 2006. Currently, he is an associate professor in the School of Information Science and Engineering, Fujian University of Technology(FJUT), China. His current research interests include information retrieval, data mining, and digital image processing.

Jiang-Yi Lin received the B.S. and M.S. degrees in Computer science and Technology from FuZhou Uniersity, FuJian, China, in 2005 and 2008, repectively. He is currently pursuing the Ph.D degree with the Multimedia and Secure Networking Laboratory (MSN lab), the Department of Information Engineering and Computer Science of Feng Chia University, Taichung, Taiwan. His research interests include image processing, secret sharing and steganography.

Chin-Chen Chang received his Ph.D. degree in computer engineering from National Chiao Tung University. His current title is Chair Professor in Department of Information Engineering and Computer Science, Feng Chia University, from February 2005. He is currently a Fellow of IEEE and a Fellow of IEE, UK. And, since his early years of career development, he consecutively won Outstanding Talent in Information Sciences of the R. O. C., AceR Dragon Award of the Ten Most Outstanding Talents, Outstanding Scholar Award of the R. O. C., Outstanding Engineering Professor Award of the R. O. C., Distinguished Research Awards of National Science Council of the R. O. C., Top Fifteen Scholars in Systems and Software Engineering of the Journal of Systems and Software, and so on. On numerous occasions, he was invited to serve as Visiting Professor, Chair Professor, Honorary Professor, Honorary Director, Honorary Chairman, Distinguished Alumnus, Distinguished Researcher, Research Fellow by universities and research institutes. His current research interests include database design, computer cryptography, image compression, and data structures.

Yu-Chen Hu received his PhD. degree in computer science and information engineering from the Department of Computer Science and Information Engineering, National Chung Cheng University, Chiayi, Taiwan in 1999. Currently, Dr. Hu is a professor in the Department of Computer Science and Information Management, Providence University, Sha-Lu, Taiwan. He is a senior member of IEEE. He is also a member of Computer Vision, Graphics, and Image Processing (CVGIP), Chinese Cryptology and Information Security Association (CCISA), Computer Science and Information Management (CSIM) and Phi Tau Phi Society of the Republic of China. He servers as the Editor-in-Chief of International Journal of Image Processing from June 2009 to May 2015. In addition, he is the managing editor of Journal of Information Assurance & Security since March 2009. He is the associated editor of Human-centric Computing and Information Sciences since Feb. 2011. He joints the editorial boards of several other journals. His research interests include digital forensics, information hiding, image and signal processing, data compression, information security, and data engineering.

An Identity Based Proxy Signcryption Scheme without Pairings

Hui Guo and Lunzhi Deng
(Corresponding author: Lunzhi Deng)
School of Mathematical Sciences, Guizhou Normal University
Guiyang 550001, China
(Email: denglunzhi@163.com)
(Received Dec. 25, 2018; Revised and Accepted Aug. 3, 2019; First Online Feb. 9, 2020)

Abstract

The identity-based cryptography avoids the storage problem of public key certificate of public key infrastructure. The signcryption mechanism completes both authentication and encryption functions with lower communication cost. The proxy signature allows the proxy signer to sign a message on the behalf of the original signer. In this paper, a new identity based proxy signcryption (IBPS) scheme without pairings is proposed, and it is proved to be secure in the random oracle model. To the best of our knowledge, our scheme is more efficient than previous ones in computation.

Keywords: Identity Based Cryptography; Proxy Signcryption; Random Oracle Model

1 Introduction

Traditional public key cryptography [11] needs a trusted certification authority (CA) to issue a certificate which links the identity and the public key of the user. Hence, the problem of certificate management arises. To solve the problem, the notion of the identity-based public key cryptography was introduced by Shamir [20] in 1984. In this cryptography, a user's public key can be arbitrary string that can identify the user, such as the e-mail address or telephone number and so on.

In the areas of computer communications and electronic transactions, one of the essential topics is how to send data in confidential and authentication way. In 1997, Zheng [28] proposed a novel cryptographic primitive, called signcryption [21] that satisfies both the functionality of digital signature and encryption in a single logical step.

The proxy signature [9, 26] is a useful tool in real life. For example, if a document is to be signed by a CEO (original signer) of the company while he/she is absent, then the document can be signed by a manager (proxy signer) designated by the CEO (original signer) [12, 17]. The proxy signature was firstly introduced by Mambo *et al.* [19] in 1996. It allows the proxy signer to sign a message on the behalf of the original signer. On the basis of the deledation type, the proxy signature is calssified into three types: Full delegation, partial delegation and delegation by warrant. Because the first two types have some drawbacks [3], most proxy signature schemes has focused on the type of the delegation with warrant.

To delegate the signcryption righs to a trusted agent, Gamage *et al.* [4] proposed a new ideal of proxy signcryption by combining the notions of proxy signature and signcryption in 1999. But their scheme does not support provable security [22]. In 2004, Li and Chen [13] proposed the first identity-based proxy signcryption scheme using bilinear pairings.

1.1 Related Work

Many researchers have been proposed variations of signcryption schemes. Arijit Karati *et al.* [10] designed a practical identity based signcryption scheme from bilinear pairing, which is based on CDH assumption and proved to be secure under standard security model. An identity-based signcryption scheme that is forward secure in a stronger sense was proposed by Madeline González Muñiz *et al.* [18].

Deng *et al.* [3] proposed an identity based proxy signature from RSA without pairings in the randon oracle model that admits formal proofs for unforgeability of proxy signature. He *et al.* [7] introduced an ID-based proxy signature schemes without bilinear pairings, which is secure aginst adaptive chosen message and ID attack. In 2016, Hu *et al.* [5] presented a proxy signature scheme with a formal security proof based on the CDH and BDH assumption.

Since identity-based proxy signcryption (IBPS) plays an important role in practical applications such as mobile communication and e-commerce and so on, it has attracted great attention when it was proposed, and has been studied by many scholars at home and abroad. Wu Jian [27] proposed an identity-based proxy signcryption schemes. Li and Chen [13] proposed an identity based

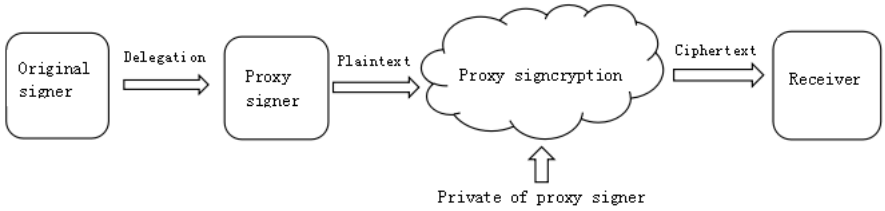


Figure 1: Process of a IBPS scheme

proxy signcryotion scheme which is based on the Libert and Quisquater's [14] identity based signcryption scheme. But Wang *et al.* [25] point that the scheme does not satisfy the strong unforgeability security in the strict sense. Saraswat [22] proposed a secure proxy signcryption scheme which provides anonymity to the proxy signer from the receiver.

Swapna *et al.* [23] introduced an efficient ID-based proxy signcryption scheme, which offers both public verifiability and forward security. Lin *et al.* [15] introduced an efficient proxy signcryption with provable CCA and CMA security. Unfortunately, Lo and Tsai [16] pointed that the scheme is not secure against the chosen warrant attack. Other schemes proposed including proxy blind signcryption [24], generalized proxy signcryption [29]− [30], certificateless proxy signcryption [2], *etc.*

1.2 Our Contributions

In this paper, we propose a new identity based proxy signcryption scheme. The main contributions of this paper are as follows:

1) The proposed scheme is proved to be secure in the random oracle model.

2) The proposed scheme does not use pairing operation, which is more efficient than that of previous schemes [13, 16, 23, 25, 27] in computation.

2 Preliminaries

Definition 1. *Given a generator P of group G with prime order q, and a tuple $(P, aP, bP, X \in G)$ for unknown $a, b \in z_q^*$, the Decisional Diffie-Hellman problem (DDH) is to decide whether $X = abP$.*

Definition 2. *Given a generator P of group G with prime order q, and a tuple (P, aP), the Discrete Logarithm problem (DLP) is to compute a.*

2.1 Model of Identity based Proxy Signcryption

An identity based proxy signcryption scheme is composed of six polynomial time algorithms, it is defined as follows:

- Setup: Input a security parameter k, private key generator (PKG) outputs the system parameters $params$ and a master secret key msk.
- Private-Key-Extract: Input the system parameters $params$, the master secret key msk and the identity $ID_i \in \{0,1\}^*$ of a user, PKG returns a private key s_i to the user ID_i via a secure channel, and the user publish its public key R_i.
- Delegation Generate: Input the system parameters $params$, the private key s_A of original signer ID_A and a warrant w, this algorithm outputs a delegation π and sends π to the proxy signer ID_B.
- Delegation Verify: This algorithm takes as input the system parameters $params$, delegation π, and verifies whether π is a valid delegation from the original signer ID_A.
- Proxy Signcryption: Input the private key s_B of proxy signer ID_B, the receiver identity ID_C, a message m and a delegation π, this algorithm outputs a proxy signcryption ciphertext σ on behalf of the original signer ID_A.
- Proxy Unsigncryption: After receiving the ciphertext σ, the receiver ID_C decrypts the ciphertext and obtains the message m or the symbol $\perp$ if σ is a invalid ciphertext.

Definition 3. *An identity based proxy signcryption scheme is said to be indistinguishable under adaptive chosen ciphertext attacks if the polynomially bounded adversary with a negligible advantage in the following game.*

Game I. A challenger $\mathscr{C}$ and a adversary $\mathscr{A}$ play the following game.

Initialization. $\mathscr{C}$ runs the setup algorithm to generate a master secret key msk and the public system parameters $params$. $\mathscr{C}$ sends $params$ to $\mathscr{A}$ and keeps msk secret.

Phase 1. $\mathscr{A}$ makes a polynomially bounded number of adaptive queries to $\mathscr{C}$.

- Hash functions query: $\mathscr{A}$ can ask for the values of any hash functions.
- private key query: $\mathscr{A}$ chooses an identity ID_i, $\mathscr{C}$ runs the private key extraction algorithm to generate private key s_i, and sends to $\mathscr{A}$.

- Delegation query: When $\mathscr{A}$ submits the identity of original signer ID_A and a warrant w to the challenger $\mathscr{C}$, $\mathscr{C}$ responds the corresponding delegetion π to $\mathscr{A}$.
- Proxy Signcryption query: $\mathscr{A}$ chooses a message m, a receiver ID_C and the private key s_B of proxy signer ID_B, a delegetion π, and sends to $\mathscr{C}$. $\mathscr{C}$ returns the proxy signcryption ciphtext σ to $\mathscr{A}$.
- Proxy Unsigncryption query: When $\mathscr{A}$ chooses a ciphertext σ, a receiver's identity ID_C and a proxy signer ID_B, $\mathscr{C}$ outputs plaintext m generated by the proxy unsigncryption algorithm. Or $\mathscr{C}$ returns the the symbol $\perp$, if σ is an invalid proxy unsigncryption ciphertext.

Challenge. $\mathscr{A}$ sends following information to the challenger: two equal length messages m_0, m_1, a specified receiver ID_C and proxy signer ID_B, $\mathscr{C}$ takes randomly a bit $\mu \in \{0,1\}$ and computes the ciphertext σ^* on the message m_μ.

($\mathscr{A}$ should not have requested the private key for ID_C in Phase 1.)

Phase 2. $\mathscr{A}$ performs a polynomially bounded number of queries just like in phase 1, and fulfills the following restrictions:

1) $\mathscr{A}$ should not have requested the private key for ID_C.
2) $\mathscr{A}$ can not have made the proxy unsigncryption query for the ciphertext σ^*.

Response. $\mathscr{A}$ produces a bit μ' and wins the game if $\mu' = \mu$. The advantage of $\mathscr{A}$ is defined as: $Adv_{\mathscr{A}}^{IND-CLRSC}(\nu) = |2\Pr[\mu' = \mu] - 1|$.

Definition 4. *An identity based proxy signcryption scheme is said to be unforgeable under adaptive chosen message attacks if the polynomially bounded adversary with a negligible advantage in the following game.*

Game II. A challenger $\mathscr{C}$ and a adversary $\mathscr{A}$ play the following game:

Initialization, Query. Same as that in the Game I.

Forge. $\mathscr{A}$ produces a tuple $\{ID_A, ID_B, \pi\}$ or $(\sigma, w, ID_A, ID_B, ID_C)$. When one of the following conditions hold, $\mathscr{A}$ wins the game.

Case 1: The final output is $\{ID_A, ID_B, \pi\}$ and it fulfills:

1) π is a valid delegation.
2) $\mathscr{A}$ should have not queried the private key of original signer ID_A.
3) π is not obtained by the delegation query.

Case 2: The final output is $(\sigma, w, ID_A, ID_B, ID_C)$ and it fulfills:

1) σ is a proxy signcryption.
2) $\mathscr{A}$ should have not queried the private key of original signer ID_A
3) The tuple $(\pi, ID_A, ID_B)$is not appear in delegation query.
4) σ is not obtained by the proxy signcryption query.

Case 3: The final output is $(\sigma, w, ID_A, ID_B, ID_C)$ and it fulfills:

1) σ is a proxy signcryption.
2) The private key of proxy signer ID_B has not been queried.
3) σ is not obtained by the proxy signcryption query.

The advantage of $\mathscr{A}$ is defined as: $Adv_{\mathscr{A}}^{UNF-IBPS}$ =Pr[$\mathscr{A}win$].

3 Proposed Scheme

- Setup: Given the security parameter of the system k and l, PKG chooses an additive cyclic group $G = \langle P \rangle$ of prime order $q > 2^k$. Then PKG chooses four hash functions $H_1 : \{0,1\}^* \times G \to Z_q^*$, $H_2 : \{0,1\}^* \times G \times G \times G \times \{0,1\}^* \times \{0,1\}^* \to Z_q^*$, $H_3 : \{0,1\}^* \to \{0,1\}^l, H_4 : \{0,1\}^* \to Z_q^*$. The PKG randomly chooses its master secret key $x \in Z_q^*$ and computes the public key $P_{pub} = xP$. The message space is $M = \{0,1\}^l$. The PKG publishes the set of public system parameters: $params = \{G, q, P, P_{pub} = xP, H_1, H_2, H_3, H_4\}$ and keep the master key x secret.

- Private-Key-Extract: Given a user's identity $ID_i \in \{0,1\}^*$, the PKG randomly selects $r_i \in Z_q^*$ and computes $R_i = r_i P$, $d_i = H_1(ID_i, R_i)$, $s_i = r_i + d_i x$ and sends (R_i, s_i) to the user via a secure channel. The user ID_i publish his/her the public key R_i.

- Delegation Generation: The original signer ID_A selects at random $t \in Z_q^*$ and computes $T = tP$, $h = H_2(w, T, R_A, R_B, ID_A, ID_B)$, $y = t + hs_A$. Then original signer ID_A sends the delegation $\pi = (T, y, w)$ to proxy signer ID_B securely. Where w is warrant, the warrant includes the property of message to be delegated, the identity information of original signer and proxy signer, the delegation relationship between them and period of delegation, *etc.*

- Delegation Verification: On receiving the delegation $\pi = (T, y, w)$, proxy signer ID_B checks the delegation as follows:

1) Computes: $h = H_2(w, T, R_A, R_B, ID_A, ID_B)$.

2) Checks if $yP = T + h(R_A + d_A P_{pub})$. If the equality holds, accepts π as a valid delegation. Otherwise, proxy signer ID_B rejects the delegation π.

- Proxy Signcryption: To signcrypt a message m on the behalf of the original signer ID_A for the receiver ID_C, the proxy signer ID_B proceeds as following:

1) Randomly selects $n_1, n_2 \in Z_q^*$, computes $N_1 = n_1 P$, $N_2 = n_2 P$, $V = n_1(R_C + d_C P_{pub})$, $C = H_3(N_1, N_2, V, R_A, R_B, R_C, ID_A, ID_B, ID_C) \oplus m$;
2) Computes: $g = H_4(m, \pi, N_1, N_2, V, R_A, R_B, R_C, ID_A, ID_B, ID_C)$, $z = y + n_2 + g s_B$;
3) Outputs the proxy signcryption: $\sigma = \{C, N_1, N_2, z, \pi\}$.

- Proxy Unsigncryption: On receiving the ciphertext $\sigma = \{C, N_1, N_2, z, \pi\}$, the receiver ID_C decrypts the ciphertext as follows:

1) Computes: $V = s_C N_1$, $m = C \oplus H_3(N_1, N_2, V, R_A, R_B, R_C, ID_A, ID_B, ID_C)$, $g = H_4(m, \pi, N_1, N_2, V, R_A, R_B, R_C, ID_A, ID_B, ID_C)$.
2) Checking whether $zP = T + N_2 + h(R_A + d_A P_{pub}) + g(R_B + d_B P_{pub})$. If the equality holds, accepts m as a valid message. Otherwise, the receiver rejects the ciphertext.

4 Analysis of Proposed Scheme

4.1 Correctness Analysis

$$\begin{aligned} V &= n_1(R_C + d_C P_{pub}) \\ &= n_1(r_C P + d_C x P) \\ &= (r_C + d_C x) n_1 P = s_C N_1; \\ yP &= (t + h s_A) P \\ &= tP + h s_A P \\ &= T + h(r_A + d_A x) P \\ &= T + h(r_A P + d_A x P) \\ &= T + h(r_A P + d_A P_{pub}) \\ &= T + h(R_A + d_A P_{pub}); \end{aligned}$$

$$\begin{aligned} zP &= (y + n_2 + g s_B) P \\ &= yP + n_2 P + g s_B P \\ &= T + h(R_A + d_A P_{pub}) + N_2 + g(r_B + d_B x) P \\ &= T + h(R_A + d_A P_{pub}) + N_2 + g(r_B P + d_B x P) \\ &= T + h(R_A + d_A P_{pub}) + N_2 + g(R_B + d_B P_{pub}). \end{aligned}$$

4.2 Security Analysis

Theorem 1. *In random oracle model, the scheme is indistinguishable against the adversary $\mathscr{A}$ if the DDH is hard.*

Proof. Assume that the challenger $\mathscr{C}$ receives a random instance (P, aP, bP, X) of the DDH, the goal of $\mathscr{C}$ is to determine whether $X = abP$ or not. $\mathscr{C}$ runs $\mathscr{A}$ as a subroutine and plays the role of the challenger in the Game I. □

Initialization. $\mathscr{C}$ runs the setup algorithm to generate system parameters. Then $\mathscr{C}$ sends the system parameters $params = \{G, q, P, P_{pub} = xP, H_1, H_2, H_3, H_4\}$ to $\mathscr{A}$.

Queries. Without losing generality, assuming that each query is different. $\mathscr{A}$ will ask for $H_1(ID_i)$ before the identity ID_i is used any other queries. $\mathscr{C}$ will maintain some lists to store the queries and answers, all of the lists are initially empty.

- H_1 queries: $\mathscr{C}$ maintains the list L_1 of tuple (ID_i, R_i, d_i). When $H_1(ID_i, R_i)$ is queried by $\mathscr{A}$, $\mathscr{C}$ selects at random $d_i \in Z_q^*$ and sets $H_1(ID_i, R_i) = d_i$, and adds (ID_i, R_i, d_i) to list L_1.

- H_2 queries: $\mathscr{C}$ maintains the list L_2 of tuple (β, h). When $H_2(\beta)$ is queried by $\mathscr{A}$, $\mathscr{C}$ selects at random $h \in Z_q^*$, sets $H_2(\beta) = h$ and adds (β, h) to list L_2.

- H_3 queries: $\mathscr{C}$ maintains the list L_3 of tuple (U, α). When $H_3(U)$ is queried by $\mathscr{A}$, $\mathscr{C}$ selects at random $\alpha \in \{0,1\}^l$, sets $H_3(U) = \alpha$ and adds (U, α) to list L_3.

- H_4 queries: $\mathscr{C}$ maintains the list L_4 of tuple (β', h'). When $H_4(\beta')$ is queried by $\mathscr{A}$, $\mathscr{C}$ selects at random $h' \in Z_q^*$, sets $H_4(\beta') = h'$ and adds (β', h') to list L_4.

- User public key queries: $\mathscr{C}$ maintains the list L_U of tuple (ID_i, R_i). When $\mathscr{A}$ makes this query, $\mathscr{C}$ answers the query as follows:

  At the j^{th} query, $\mathscr{C}$ sets $R_j = aP$. For $i \neq j$, $\mathscr{C}$ selects at random $r_i \in Z_q^*$ and sets $R_i = r_i P$, the query and the respond will be stored in the list L_U.

- private key queries: $\mathscr{C}$ maintains the list L_K of tuple (ID_i, R_i, d_i). When $\mathscr{A}$ makes this query, $\mathscr{C}$ answers the query as follows:
  If $ID_i = ID^*$, $\mathscr{C}$ fails and stops. Otherwise $\mathscr{C}$ finds the tuple (ID_i, R_i, d_i) in list L_1, responds with $s_i = r_i + x d_i$ and adds (ID_i, R_i) to list L_D.

- Proxy Delegation queries: $\mathscr{C}$ answers the query as follows:

  If $ID_A \neq ID^*$, $\mathscr{C}$ give a delegation π by calling the proxy delegation algorithm to answer $\mathscr{A}$. Otherwise, $\mathscr{C}$ does as follows.

1) Randomly chooses $y, h \in Z_q^*$, computes: $T = yP - h(R_A + d_A P_{pub})$;
2) Stores the relation: $h = H_2(w, T, R_A, R_B, ID_A, ID_B)$ and adds to the list L_1. If collision occurs, repeats the steps (1)-(2).

3) Outputs the delegation: $\pi = (T, y, w)$.

- Proxy Signcryption queries: When $\mathscr{A}$ selects a message m, proxy signer ID_B and receiver ID_C, $\mathscr{C}$ returns a proxy signcryption as follows:
If $ID_B \neq ID^*$, $\mathscr{C}$ give a proxy signcryption σ by calling the the proxy signcryption algorithm to answer $\mathscr{A}$. Otherwise, $\mathscr{C}$ does the following steps:

  1) Randomly selects $n_1, n_2, g \in Z_q*$, computes: $N_1 = n_1P$, $N_2 = n_2P - g(R_B + d_BP_{pub})$, $V = n_1(R_C + d_CP_{pub})$, $C = H_3(N_1, N_2, V, R_A, R_B, R_C, ID_A, ID_B, ID_C) \oplus m$;
  2) Computes: $z = y + n_2$;
  3) Stores the relations: g=$H_4(m, w, N_1$,$V, N_2, R_A, R_B, R_C, ID_A, ID_B, ID_C)$. If collision occurs, repeats Steps (1)-(3);
  4) Outputs the proxy signcryption:

$$\sigma^* = \{C, N_1, N_2, z, \pi\}.$$

- Proxy Unsigncryption queries: If $ID_C \neq ID^*$, $\mathscr{C}$ give a message m by calling the proxy unsigncryption algorithm. Otherwise, $\mathscr{C}$ notifies that σ is an invaild ciphertext.

Challenge. $\mathscr{A}$ chooses two equal length messages m_0, m_1, a specified receiver ID_C, and proxy signer ID_B. If $ID_C \neq ID^*$, $\mathscr{C}$ fails and stops. Otherwise, $\mathscr{C}$ picks $\mu \in \{0, 1\}$, and computes ciphertext σ^* on the message M_μ as follows:

1) Randomly selects $b, n_2 \in Z_q^*$, computes: $N_1 = bP$, $N_2 = n_2P$, $V = X + d_Cx \cdot N_1$, $C = H_3(N_1, N_2, V, R_A, R_B, R_C, ID_A, ID_B, ID_C) \oplus m$;
2) Computes: $g = H_4(m, \pi, N_1, V, N_2, R_A, R_B, R_C, ID_A, ID_B, ID_C)$, $z = y + n_2 + gs_B$;
3) Outputs the proxy signcryption ciphertext:

$$\sigma = \{C, N_1, N_2, z, \pi\}.$$

Phase 2. $\mathscr{A}$ makes a polynomially bounded number of queries just like Phase 1. (but $\mathscr{A}$ should not have queried the private key for ID_C and requested the plaintext corresponding to the ciphertext σ^*).

Response. $\mathscr{A}$ outputs $\mu' \in \{0, 1\}$. If $\mu' \doteq \mu$, $\mathscr{C}$ outputs 1. Otherwise, $\mathscr{C}$ outputs 0. If $X = abP$, σ^* is a valid ciphertext. Then $\mathscr{A}$ can distinguishes μ with the advantage ε. So $\Pr[\mathscr{C} \longrightarrow 1|X = abP] = \Pr[\mu' \doteq \mu|X = abP] = \frac{1}{2} + \varepsilon$.

If $X \neq abP$, when $\mu = 0$ or $\mu = 1$, each part of the ciphertext has the same probability distribution, so $\mathscr{A}$ has no advantage in distinguishing μ. So $\Pr[\mathscr{C} \longrightarrow 1|X \neq abP\ ] = \Pr[\mu' \doteq \mu|X \neq abP] = \frac{1}{2}$.

Probability. Let $q_{H_i}(i = 1, 2, 3, 4)$, q_U, q_K, q_D and q_S be the number of $H_i(i = 1, 2, 3, 4)$ queries, public key queries, private key queries, delegating queries and proxy signcryption queries, respectively.

We denotes some events as follows:

- π_1: $\mathscr{C}$ does not fail in private key queries;
- π_2: $\mathscr{C}$ does not fail in proxy unsigncryption queries;
- π_3: $\mathscr{C}$ does not fail in challenge stage.

It is easy to get following results:

$$\begin{aligned}
Pr[\pi_1] &= 1 - \frac{q_K}{q_U}, \\
Pr[\pi_2] &= 1 - \frac{1}{2^k}, \\
Pr[\pi_3] &= \frac{1}{q_U - q_K}. \\
\Pr[\mathscr{C}\ success] &= \Pr[\pi_1 \wedge \pi_2 \wedge \pi_3] \\
&= \Pr[\pi_1] \cdot \Pr[\pi_2] \cdot \Pr[\pi_3] \\
&= (1 - \frac{q_K}{q_U}) \cdot (1 - \frac{1}{2^k}) \cdot \frac{1}{q_U - q_K} \\
&\approx \frac{1}{q_U}
\end{aligned}$$

Therefore, if $\mathscr{A}$ can succeed with the probability ε, then $\mathscr{C}$ can solve the DDH with probability $\frac{\varepsilon}{q_U}$.

Theorem 2. *In random oracle model, the scheme is unforgeable against adversary $\mathscr{A}$ if the DLP is hard.*

Proof. Assume that the challenger $\mathscr{C}$ receives a random instance (P, aP) of the DLP. the goal of $\mathscr{C}$ is to compute the value of a. $\mathscr{C}$ will run $\mathscr{A}$ as a subroutine and play the role of challenger in the Game II. □

Initialization, Query. Same as that in the Game II.

Forge. $\mathscr{A}$ outputs a tuple $\{\pi = \{T, y, w\}, ID_A\}$ or $\{\sigma = (C, N_1, N_2, z, \pi), ID_A, ID_B, ID_C\}$. There are three situations to consider:

Case 1. The final output is $\{\pi = \{T, y, w\}, ID_A\}$ and the output fulfills the demande of Case 1 as defined in the game.

Solve DLP. Using the forking lemma for generic signature scheme [1], after replays $\mathscr{A}$ with the same random tape except the λ^{th} result returned by H_2 query of the forged message, $\mathscr{C}$ gets two valid proxy signcryptions: $\{T, y, w\}$ and $\{T, y', w\}$. Where $h = H_2(w, T, R_A, R_B, ID_A, ID_B)$, $h' = H_2'(w, T, R_A, R_B, ID_A, ID_B)$, $h \neq h'$. If $ID_A = ID^*$, $\mathscr{C}$ solves DLP by computing: $a = (h' - h)^{-1}(y' - y) - d_Ax$.

Probability. Let $q_{H_i}(i = 1, 2, 3, 4)$, q_U, q_K, q_D and q_S be the number of $H_i(i = 1, 2, 3, 4)$ queries, public key queries, private key queries, delegating queries and proxy signcryption queries, respectively.

We denote some events as follows: π_1: $\mathscr{C}$ does not fail during the queries; π_2: $\mathscr{C}$ does not fail in proxy unsigncryption queries. π_3: $ID_A = ID^*$.

It is easy to get following results:

$$\begin{aligned}
\Pr[\pi_1] &= \frac{q_U - q_K}{q_U}, \\
\Pr[\pi_2|\pi_1] &= 1 - \frac{1}{2^k}, \\
\Pr[\pi_3] &= \frac{1}{q_U - q_K}. \\
\Pr[\mathscr{C}\ success] &= \Pr[\pi_1 \wedge \pi_2 \wedge \pi_3] \\
&= \Pr[\pi_1] \cdot \Pr[\pi_2|\pi_1] \cdot \Pr[\pi_3] \\
&= \frac{q_U - q_K}{q_U} \cdot (1 - \frac{1}{2^k}) \cdot \frac{1}{q_U - q_K} \\
&\approx \frac{1}{q_U}
\end{aligned}$$

Therefore, if $\mathscr{A}$ can succeed with the probability ε, then $\mathscr{C}$ can solve DLP with the probability $\frac{\varepsilon}{q_U}$.

Case 2. The final output is $\{\sigma = (C, N_1, N_2, z, \pi), ID_A, ID_B, ID_C\}$ and the output fulfills the demand of Case 2 as defined in Game II.

Solve DLP. Using the forking lemma for generic signature Scheme [1], after replays $\mathscr{A}$ with the same random tape except the result returned by H_2 query of the forged message, $\mathscr{C}$ gets two valid proxy signcryptions: $\{C, N_1, N_2, z, \pi = (T, y, w)\}$ and $\{C, N_1, N_2, z, \pi' = (T, y, w)\}$. Where $h = H_2(w, T, R_A, R_B, ID_A, ID_B)$, $h' = H_2'(w, T, R_A, R_B, ID_A, ID_B)$, $h \neq h'$. $g = g' = H_4(m, \pi, N_1, V, N_2, R_A, R_B, R_C, ID_A, ID_B, ID_C)$. If $ID_A = ID^*$, $\mathscr{C}$ solves DLP by computing: $a = (h' - h)^{-1}(y' - y) - d_A x$.

Probability. Probability of success is same as the probability in Case 1.

Case 3. The final output is $\{\sigma = (C, N_1, N_2, z, w), ID_A, ID_B, ID_C\}$ and the output fulfills the demand of Case 3 as defined in Game II.

Solve DLP. Using the forking lemma for generic signature Scheme [1], after replays $\mathscr{A}$ with the same random tape except the result returned by H_4 query of the forged message, $\mathscr{C}$ gets two valid proxy signcryptions: $\{C, N_1, N_2, z, \pi\}$ and $\{C, N_1, N_2, z', \pi\}$. Where $g = H_4(m, \pi, N_1, V, N_2, R_A, R_B, R_C, ID_A, ID_B, ID_C)$, $g' = H_4'(m, \pi, N_1, V, N_2, R_A, R_B, R_C, ID_A, ID_B, ID_C)$, $g \neq g'$. If $ID_c = ID^*$, $\mathscr{C}$ solves DLP by computing: $a = (g' - g)^{-1}(z' - z) - d_B x$.

Probability. Probability of success is same as the probability in Case 1.

5 Efficiency and Comparison

By using a famous encryption library (MIRACL) on a mobile device (Samsung Galaxy S5 with a Quad-core 2.45G processor, 2G bytes memory and the Google Android 4.4.2 operating system), He *et al.* [7] obtained the running time for cryptographic operations. The running time are listed in Table 1.

For the IBPS scheme based on biliner pairing, to achieve the 1024 bits RSA level security, a Tate pairing $G_1 \times G_1 \longrightarrow G_2$ defined over the supersigingular elliptic curve E/F_p: $y^2 = x^3 + x$ was used, where both q and p are 160 bits and 512 bits, respectively. To achieve the same level of scurity, for the IBPS scheme based on the non-singular elliptic curve cryptography, they used an additive group with the prime order q, which is defined on a non-sigular elliptic curve over the finite field F_p, where both p and q are 160 bits. We define some notations as follows:

P: A pairing operation.

M_{G_1}: A scalar multiplication operation in G_1.

M_G: A scalar multiplication operation in G.

E_{G_2}: A exponentiation operation in G_2.

We use a simple method to evaluate the computaion efficiency of the different schemes. For example, the scheme [25] needs 13 pairing operations, 4 scalar multiplication operation in G_1, 7 exponentiation operations in G_2. Therefore, the resulting operation time is $13 \times 32.713 + 4 \times 13.405 + 7 \times 2.249 = 494.632$.

According to the above ways, the resulting operation time of other shemes [13,16,23,25,27] is shown in Table 2.

Table 1: Cryptographic operation time (in milliseconds)

P	M_{G_1}	M_G	E_{G_2}
32.713	13.405	3.335	2.249

6 Conclusion

Although several good results have been achieved in speeding up the computation of bilinear pairing function in recent years. The pairing operation is still relatively expensive and the relative computation cost of the pairing is approximately twenty times higher than that of scalar multiplication over elliptic curve group. So it is still quite significant to design cryptography scheme with less pairing operation. In order to save the running time, in the letter, we constructe an identity based proxy signcryption without bilinear pairings. With the running time being saved greatly, as far as my knowledge is concerned, our scheme is more effective than the previous related schemes in computation.

Table 2: Comparison of several IBPS schemes

Schemes	Delegate	D-Verify	Proxy signcryption	P-unsigncryption	Time
Wu [27]	$2M_{G_1}$	2P+M_{G_1}	$P + 2M_{G_1} + E_{G_2}$	$2P + E_{G_2}$	235.088
Wang [25]	$3M_{G_1}$	$3P + E_{G_2}$	$2P + M_{G_1} + 2E_{G_2}$	$8P + 4E_{G_2}$	494.632
Swapna [23]	$2M_{G_1}$	$2P + M_{G_1}$	$P + 2M_{G_1} + E_{G_2}$	$3P + 2M_{G_1}$	292.362
Lo [16]	M_{G_1}	$2M_{G_1}$	$P + 4M_{G_1}$	$3P + 5M_{G_1}$	291.712
Li [13]	$3M_{G_1}$	$3P + E_{G_2}$	$2P + 2M_{G_1} + 2E_{G_2}$	$8P + 4E_{G_2}$	508.037
Our scheme	M_G	$3M_G$	$4M_G$	$6M_G$	46.69

Acknowledgments

The authors are grateful to the anonymous referees for their helpful comments and suggestions. The research is supported by the National Natural Science Foundation of China under Grants 61562012, the Innovation Group Major Research Projects of Department of Education of Guizhou Province under Grant No. KY[2016]026.

References

[1] M. Bellare and G. Neven, "Multi-signatures in the plain public key model and a general forking lemma," in *Proceedings of 13th ACM Conference on Computer and Communications Security (CCS'06)*, pp. 390–399, 2006.

[2] T. Bhatia and A. K. Verma, "Cryptanalysis and improvement of certificateless proxy signcryption scheme for e-prescription system in mobile cloud computing," *Annals of Telecommunications*, vol. 72, no. 9–10, pp. 563–576, 2017.

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

[4] C. Gamage, J. Leiwo and Y. Zheng, "An efficient scheme for secure message transmission using proxy-signcryption," in *Proceeding of 22nd Australasian Computer Science Conference (ACSC'99)*, pp. 420–431, 1999.

[5] X. Hu, W. Tan, H. Xu and J. Wang, "Short and provably secure designated verifier proxy signature scheme," *IET Information Security* , vol. 10, no. 2, pp. 69–79, 2013.

[6] D. 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, 2017.

[7] D. He, J. Chen and J. Hu, "An ID-based proxy signature scheme without bilinear pairings," *Annual Telecommunications*, vol. 66, no. 11–12, pp. 657–662, 2011.

[8] Y. Huang and J. Yang, "A novel identity-based signcryption scheme in the standard model," *Information*, vol. 8, no. 2, pp. 58, 2017.

[9] 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.

[10] A. Karati and G. P. Biswas, "A practical identity based signcryption scheme from bilinear pairing," *Advances in Computing*, pp. 832–836, 2016.

[11] A. V. N. Krishna, A. H. Nareyana, 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.

[12] 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.

[13] X. Li and K. Chen, "Identity based proxy signcryption scheme from pairings," in *Proceedings of the IEEE International Conference on Services Computing (SCC'04)*, pp. 494–497, 2004.

[14] B. Libert and J. Quisquater, "A new identity based signcryption schemes from pairings," in *Proceedings of the IEEE International Theory Workshop*, pp. 155–158, 2003.

[15] H. Lin, T. Wu, S. Huang and Y. S. Yeh, "Efficient proxy signcryption schemes with provable CCA and CMA security," *Computers and Mathmatics with Applications*, vol. 60, no. 7, pp. 1850–1858, 2010.

[16] N. Lo and J. Tsai, "A provably secure proxy signcryption scheme using bilinear pairings," *Journal of Applied Mathematics* , vol. 2014, pp. 10, 2014. (`http://dx.doi.org/10.1155/2014/454393`)

[17] 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.

[18] M. G. Muñiz and P. Laud, "Strong forward security in identity-based signcryption," *Journal of Discrete Mathematical Sciences and Cryptography*, vol. 16, no. 4–5, pp. 235–258, 2013.

[19] M. Mambo, K. Usuda, and E. Okamoto, "Proxy signature: Delegtion of power to sign messages," *IEICE Trasactions on Fundamentals*, vol. E79-A, no. 9, pp. 1138–1353, 1996.

[20] A. Shamir, "Identity-based cryptosystem and signature scheme," in *Advances in Cryptology*, pp. 47–53, 1985.

[21] S. Shan, "An efficient certificateless signcryption scheme without random oracles," *International Journal of Electronics and Information Engineering*, vol. 11, no. 1, pp. 9–15, 2019.

[22] V. Saraswat, R. A. Sahu and A. k. Awasthi, "A secure anonymous proxy signcryption scheme," *Jouranl of Mathematical Cryptology*, vol. 11, no. 2, pp. 63–84, 2017.

[23] G. Swapna, P. V. S. S. N. Gopal, T. Gowri and P. V. Reddy, "An efficient ID-based proxy signcryption scheme," *International Jouranl of Information and Network Security*, vol. 3, no. 1, pp. 200–206, 2012.

[24] S. Ullah, M. Junaid, F. Habib, Sana, *et al.*," A novel proxy blind signcryption scheme based on hyper elliptic curve," in *Proceedings of the 12th International Conference on Natural Computation Fuzzy Systems and Knowledge Discovery*, pp. 1964–1968, 2016.

[25] M. Wang, H. Li and Z. Liu, "Efficient identity based proxy-signcryption schemes with forward security and public verifiability," *Networking and Mobile Computing*, pp. 982–991, 2005.

[26] F. Wang, C. C. Chang, C. L. Lin, and S. C. Chang, "Secure and efficient identity-based proxy multisignature using cubic residues," *International Journal of Network Security*, vol. 18, no. 1, pp. 90–98, 2016.

[27] J. Wu, "Identity-based proxy signcryption schemes," *Applied Mechanics and Materials*, vol. 380-384, pp. 2605–2608, 2013.

[28] Y. Zheng, "Digitial signcryption or how to achieve cost (Signature and encryption) cost (Signature) + Cost (Encryption)," *Advances in Cryptology*, pp. 165–179, 1997.

[29] C. Zhou, "Identity based generalized proxy signcryption scheme," *Information Technology and Control*, vol. 45, no. 1, pp. 13–26, 2016.

[30] C. Zhou, "A provable secure identity-based generalized proxy signcryption scheme," *International Journal of Network Security*, vol. 20, no. 6, pp. 1183–1193, 2018.

Biography

Hui Guo received her B.S. from Guizhou Normal University, Guiyang, PR China, in 2016; She is now a master student in the School of Mathematical Sciences, Guizhou Normal University, Guiyang, PR China. Her recent interest include cryptography and information safety.

Lunzhi Deng received his B.S. from Guizhou Normal University, Guiyang, PR China, in 2002; M.S. from Guizhou Normal University, Guiyang, PR China, in 2008; and Ph.D. from Xiamen, PR China, in 2012. He is now a professor in the School of Mathematical Sciences, Guizhou Normal University, Guiyang, PR China. His recent interest include algebra and information safety.

Fast Scalar Multiplication Algorithm Based on Co_Z Operations on Elliptic Curves over GF(3^m)

Shuang-Gen Liu, Shi-Mei Lu, and Rui-Wen Gong
(Corresponding author: Shuang-Gen Liu)
School of Communication and Information Engineering, Xi'an University of Posts and Telecommunications
Xi'an 710121, China
(Email: liusgxupt@163.com)
(Received Jan. 11, 2019; Revised and Accepted June 19, 2019; First Online July 30, 2019)

Abstract

In this paper, the Co_Z idea is adopted to propose the Co_Z point addition formulae in Jacobian projective coordinate and Lopez & Dahab projective coordinate on elliptic curve over GF(3^m) in order to improve the efficiency and safety of scalar multiplication algorithm. Considering the optimized symmetric ternary algorithm for scalar multiplication against Simple Power Attack (SPA), the computational efficiency is increased by 16%, 26% and 10% compared with the balanced ternary scalar multiplication algorithm, the symbolic ternary form(STF) algorithm and the optimized symmetric ternary scalar multiplication algorithm. In addition, the new formula for $3P + Q$ operation in Jacobian projective coordinate is also proposed. The efficiency of improved balanced ternary scalar multiplication algorithm is increased by 7% compared with the previous algorithm.

Keywords: Characteristic Three; Co_Z *Operation; Projective Coordinate; Scalar Multiplication*

1 Introduction

Elliptic curve, as an important issue of algebraic geometry, has been studied for more than 100 years. Until in the year of 1985, Kolitz and Miller introduced it into the field of cryptography and constructed ECC (Elliptic Curve Cryptography). Compared with other cryptosystems, ECC has advantages such as short key, less storage space and low bandwidth requirements, which makes it has a superior development prospect. In 2017, Du [5] constructed an ID-based dynamic group communication sign-cryption scheme by using hyperelliptic curve cryptosystem and ID-based sign-cryption model. Elliptic curve cryptosystem mainly includes elliptic curve key generation, key exchange, encryption, decryption, signature and other algorithms. In these elliptic curve cryptography algorithms, scalar multiplication is the most time-consuming operation. How to improve the efficiency of scalar multiplication on the elliptic curve has been a hot topic of public key cryptography. Scalar multiplication, ie, the computation of the point $kP = P + ... + P$, where k is an integer and P is a point on the elliptic curve. There are two main ways to improve the efficiency of scalar multiplication. On the one hand, it is the effective representation of scalar k, such as Binary Scalar Multiplication (BSM), non-adjacent form (NAF) [18] and Balanced Ternary form (BTF) [4]. On the other hand, it is the improvement of point addition and doubling formulas. Point doubling and addition involve operations over prime or ternary fields, In these operations, it is inversion that is the most time-consuming. We can change the coordinates to reduce the number of field inversions, the commonly used coordinate systems are projective coordinate, Jacobian coordinate [3], and Lopez & Dahab coordinate [12].

In recent years, there have been many studies on elliptic curves over GF(2^m) and GF(p). References [1,9,13] provided the fast scalar multiplication algorithms over GF(p), which accelerate the computing speed of elliptic curve cryptosystem. Hisil *et al.* [8] provided a faster mixed addition on modified Jacobiquartic coordinates, and introduced tripling formulae for different forms. In 2007, Meloni [15] proposed the earliest Co_Z point addition operation on Weierstraβ elliptic curves, which used a small amount of computation to calculate the addition operation of two different points with identical Z coordinate. Goundar *et al.* [7] presented the further Co_Z addition formula for various point additions on Weierstraβ elliptic curves. In 2017, Yu *et al.* [19] proposed the Co_Z montgomery algorithm over the finite field of characteristic three by using the same Z-coordinate. In this paper, we develop the fast Co_Z point addition formulas in different projective coordinate systems over GF(3^m), these new operational formulas are used to optimize the existing scalar multiplication algorithms [4,13,14]. As a result, we get efficient scalar multiplication algorithms based on Co_Z point addition using signed ternary representation.

The organizational structure of the paper is as fol-

lows. The next section introduces the basics of elliptic curves over GF(3^m), the point operations under different projective coordinate systems, and the scalar multiplication algorithm based on symmetric ternary representation. The third section describes the specific steps to improve the calculation formula of point addition in different coordinates by using Co_Z idea. We also optimize $3P+Q$ formula for Jacobian projective coordinate, then the improved scalar multiplication algorithm is proposed. The fourth section gives the performance analysis of the scalar multiplication algorithm combined with new formulas. Finally, the fifth section draws a conclusion.

2 Preliminaries

2.1 Basic Knowledge of Elliptic Curve over GF(3^m)

Definition 1. *An elliptic curve over Fp is given by using the generalized Weierstrass equation:*

$$E: y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6,$$

where $a_1, a_2, a_3, a_4, a_6 \in Fp$, and satisfy the discriminant $\triangle \neq 0$ on elliptic curve E. The condition of $a = 0$ ensures the smoothness of the elliptic curve, that is to say, no point on the curve has two or more tangent lines. If the characteristic of Fp is equal to 3 and $a_1^2 \neq -a_2$, elliptic curve E can be given by:

$$y^2 = x^3 + ax^2 + b,$$

where $a, b \in Fp, a, b \neq 0$, and the curve is nonsuper singular. The basic operations of points on the elliptic curve are defined as follows:

Let $P = (x_1, y_1) \in E(GF(3^m))$, $Q = (x_2, y_2) \in E(GF(3^m))$, $P \neq \pm Q$, then $R = P + Q = (x_3, y_3)$

$$\begin{cases} x_3 = (\frac{y_2-y_1}{x_2-x_1})^2 - x_1 - x_2 - a \\ y_3 = \frac{y_2-y_1}{x_2-x_1}(x_1 - x_3) - y_1. \end{cases}$$

Given the point $P = (x_1, y_1) \in E(GF(3^m))$, its double $R = 2P = (x_3, y_3)$ is obtained by:

$$\begin{cases} x_3 = (\frac{ax_1}{y_1})^2 + x_1 - a \\ y_3 = \frac{ax_1}{y_1}(x_1 - x_3) - y_1 \end{cases} \quad (1)$$

It can be seen from Equation (1) that $1I + 2M + 1S$ is required for point addition. If $a = 1$, point doubling needs $1I + 1S + 2M$. We denote field inversion as I, field squaring as S, field multiplication as M, and field cubing as C.

2.2 Projective Coordinate System

In affine coordinate, point doubling and addition on the elliptic curve over GF(3^m) involve the number of field inversions, which is a quite time-consuming operation. In order to avoid inversion operation, projective coordinate systems (X, Y, Z) are introduced. There are mainly two types of projective coordinate systems discussed in this paper, that is Jacobian projective coordinate and Lopez & Dahab projective coordinate.

2.2.1 Jacobian Projective Coordinate

Let $P = (x, y) \in E(GF(3^m))$, the correspondence between the Jacobian projective coordinate and the affine coordinate is: $(x, y) \longmapsto (\frac{X}{Z^2}, \frac{Y}{Z^3})$.

The curve equation in Jacobian projective coordinate is represented by:

$$Y^2 = X^3 + aX^2Z^2 + bZ^6.$$

Given two points $P = (X_1, Y_1, Z_1)$, $Q = (X_2, Y_2, Z_2) \in E(GF(3^m))$, and $P \neq \pm Q$, the point $R = P + Q = (X_3, Y_3, Z_3)$ is obtained by:

$$\begin{cases} X_3 &= U - (U_1 + U_2)H^2 \\ Y_3 &= -RU + (S_1 + S_2)H^3 \\ Z_3 &= HZ_1Z_2 \end{cases} \quad (2)$$

where $U_1 = X_1Z_2^2; U_2 = X_2Z_1^2; S_1 = Y_1Z_2^3; S_2 = Y_2Z_1^3; H = U_2 - U_1; R = S_2 - S_1; U = R^2 - aZ_3^2$.

Let $P = (X_1, Y_1, Z_1) \in E(GF(3^m))$, its double $R = 2P = (X_3, Y_3, Z_3)$ is obtained by

$$\begin{cases} X_3 &= U + X_1Y_1^2 \\ Y_3 &= -(aZ_1^2X_1)U - Y_1^4 \\ Z_3 &= Y_1Z_1 \end{cases}$$

where $U = (aZ_1^2X_1)^2 - aZ_1^2Y_1^2$.

From the above equations, if $a = 1$, the point addition costs $10M + 3C + 3S$, and the point doubling costs $5M + 3S$. The mixed addition $P + Q$ costs $7M + 2C + 2S$ when $Z_2 = 1$.

2.2.2 Lopez &Dahab Projective Coordinate

In the Lopez &Dahab coordinate, the point addition formula is as follows:

Suppose $P = (X_1, Y_1, Z_1)$, $Q = (X_2, Y_2, Z_2) \in E(GF(3^m))$. If $P \neq \pm Q$, the point $R = P + Q = (X_3, Y_3, Z_3)$ is obtained by:

$$\begin{cases} X_3 &= (A - B)^2 - E(C - D)^2(C + D - aE) \\ Y_3 &= E(C - D)(A - B)[DE(C - D)^2 - X_3] \\ & \quad -Z_3B(C - D)^2 \\ Z_3 &= [E(C - D)]^2 \end{cases} \quad (3)$$

where $A = Y_2Z_1^2; B = Y_1Z_2^2; C = X_2Z_1; D = X_1Z_2; E = Z_1Z_2; F = Z_1^2; G = Z_2^2$.

From Equation (3), the point addition costs $13M + 5S$. The mixed addition $P + Q$ costs $10M + 4S$ when $Z_2 = 1$.

2.3 Scalar Multiplication Algorithm Based on Symmetric Ternary

Definition 2. *An arbitrary positive integer* K *is expressed as* $K = a_n a_{n-1}...a_1 a_0$, *where* $a_n a_{n-1}...a_1 a_0$ *are numbers from -1, 0, 1. It is called symmetric ternary representation.*

In 2017, Liu *et al.* [13] optimized the bottom operation using the Jacobian coordinate on elliptic curves over GF(p). And they also optimized the symmetric ternary representation. In Jacobian coordinate system, point doubling is faster than point addition, so the non-zero Hamming weight is reduced by changing operations, and the efficiency of scalar multiplication is improved by replacing point addition with point doubling. Algorithm 1 describes the optimized symmetric ternary scalar multiplication method.

Algorithm 1 Optimized symmetric ternary scalar multiplication algorithm

Input: $K = (k_{n-1}k_{n-2}...k_1k_0)_3, P(X, Y, Z)$
Output: KP
$i = 0, P_1 = \propto, Q = \propto$
while $i \leq n-1$ **do**
  if $k_i = 1$ **then**
    $Q = Q + P, P = 3P$
  else if $k_i = -1$ **then**
    $Q = Q - P, P = 3P$
  else if $k_i = 2$ **then**
    $P_1 = 2P, Q = Q + P_1, P = 3P$
  else if $k_i = -2$ **then**
    $P_1 = 2P, Q = Q - P_1, P = 3P$
  else
    $P = 3P$
  end if
end while
$i = i + 1$
Return $Q = (X_q, Y_q, Z_q)$
End

According to the analysis, the calculation of Algorithm 1 requires $n(9M+7S)+\frac{n}{6}(4M+6S)+\frac{n}{2}(12M+4S)$.

3 The Improved Scalar Multiplication Algorithm

3.1 Scalar Multiplication Algorithm Based on Co_Z Operation

3.1.1 Co_Z Point Addition in Jacobian Projective Coordinate

Let $P = (X, Y, Z), Q = (X, Y, Z)$ be two points with the same Z-coordinate in Jacobian projective coordinate on elliptic curve over GF(3^m). We perform the Co_Z operation on Equation (2) and obtain the Co_Z point addition formula:

$$\begin{cases} X_3 &= Z^6(Y_2 - Y_1)^2 - aZ_3^2 - Z^6(X_1 + X_2) \cdot (X_2 - X_1)^2 \\ Y_3 &= -Z^3(Y_2 - Y_1)[Z^6(Y_2 - Y_1)^2 - aZ_3^2] \\ & \quad + Z^9(Y_1 + Y_2)(X_2 - X_1)^3 \\ Z_3 &= Z^4(X_2 - X_1). \end{cases}$$

After simplification, we get:

$$\begin{cases} X_3 &= (Y_2 - Y_1)^2 - a[Z(X_1 - X_2)]^2 \\ & \quad -(X_1 + X_2) \cdot (X_2 - X_1)^2 \\ Y_3 &= -(Y_2 - Y_1)[(Y_2 - Y_1)^2 - a[Z(X_2 \\ & \quad -X_1)]^2] + (Y_1 + Y_2)(X_2 - X_1)^3 \\ Z_3 &= Z(X_2 - X_1). \end{cases}$$

Algorithm 2 describes the Co_Z point addition operation in Jacobian projective coordinate, where the coordinate representation of output point satisfies $(X_1(X_2 - X_1)^2, Y_1(X_2 - X_1)^3, Z(X_2 - X_1) \sim (X_1, Y_1, Z))$. In Algorithm 2, it requires $4M + 3S + 1C$ for per bit. In this paper, we take $S = 0.8M, C = 1.37M$. Compared with mixed addition, the reduced cost is about $2M + 1C$.

Algorithm 2 Co_Z point addition algorithm in Jacobian projective coordinate(J-ZADD)

Input: $P(X_1, Y_1, Z), Q(X_2, Y_2, Z)$
Output: $P + Q = (X_3, Y_3, Z_3)$
$A \leftarrow (X_2 - X_1), B \leftarrow Z \cdot A$
$C \leftarrow (Y_2 - Y_1)^2, D \leftarrow (X_2 - X_1)^3$
$X_3 = C - aB^2 - (X_1 + X_2) \cdot (X_2 - X_1)^2$
$Y_3 = -(Y_2 - Y_1) \cdot (C - aB^2) + (Y_1 + Y_2) \cdot (X_2 - X_1)^3$
$Z_3 = Z \cdot A$
Return $Q = (X_3, Y_3, Z_3)$
End

3.1.2 Co_Z Point Addition in Lopez &Dahab Projective Coordinate

Let $P = (X_1, Y_1, Z_1)$ and $Q = (X_2, Y_2, Z_2)$ be two points with the same Z coordinate in Lopez & Dahab projective coordinate, then the operation formula of $P + Q = (X_3, Y_3, Z_3)$ can be simplified as:

$$\begin{cases} X_3 &= (Y_2 - Y_1)^2 - Z(X_2 - X_1)^2(X_1 + X_2 + aZ) \\ Y_3 &= Z(X_2 - X_1)(Y_2 - Y_1)[X_1Z(X_2 - X_1)^2 - X_3] \\ & \quad - Y_1[Z(X_2 - X_1)^2]^2 \\ Z_3 &= [Z(X_2 - X_1)]^2. \end{cases}$$

In Algorithm 3, the Co_Z point addition operation in Lopez &Dahab projective coordinate is given, which satisfies $(X_1U, Y_1U^2, ZU) \sim (X_1, Y_1, Z)$ in the process of algorithm execution, so $P + Q = (X_3, Y_3, Z_3)$ has the same Z coordinate as input points P and Q.

The computation of Algorithm 3 requires $8M + 3S$ per bit, and the reduced cost is about $5M + 2S$ compared with previous point addition. Compared with mixed addition, the reduced cost is about $3M$.

Although the Algorithm 1 mentioned in section 2.3 improves the efficiency of scalar multiplication algorithm,

Algorithm 3 Co_Z point addition algorithm in Lopez &Dahab projective coordinate(LD-ZADD)

Input: $P(X_1, Y_1, Z), Q(X_2, Y_2, Z)$
Output: $P + Q = (X_3, Y_3, Z_3)$
$S \leftarrow (X_2 - X_1)^2, U \leftarrow Z \cdot S$
$V \leftarrow Z \cdot (X_2 - X_1), W \leftarrow (Y_2 - Y_1)^2$
$X_3 = W - U \cdot (X_1 + X_2 + aZ)$
$Y_3 = V \cdot (Y_2 - Y_1) \cdot (X_1 \cdot U - X_3) - Y_1 \cdot U^2$
$Z_3 = Z \cdot U$
Return $Q = (X_3, Y_3, Z_3)$
End

it cannot resist SPA attack. Therefore, combining with Co_Z point addition operation in Jacobian projective coordinate, we propose an optimized symmetric ternary scalar multiplication algorithm to resist SPA attacks.

Algorithm 4 Optimized symmetric ternary scalar multiplication algorithm against SPA

Input: $P = (X, Y, Z) \in E(GF(3^m)), and \quad k = \sum_{i=0}^{n-1} k_i 3^i, k_i \in (-2, -1, 0, 1, 2)$
Output: kP
$Q = O, Q_1 = P, Q_{-1} = -P, Q_2 = 2P, Q_{-2} = -2P$
for $i = n - 1, ..., 0$ **do**
  $Q = 3Q$
  $Q = J - ZADD(Q, Q_{ki})$
end for
Return Q
End

In Algorithm 4, each loop performs point addition and tripling(T) operation, and two doublings are required in precomputation. Through the analysis, we can get computation complexity of Algorithm 4 about $n(ZA+T)+2D$, where ZA represents Co_Z addition and D represents doubling operation. The implementation of the process in Algorithm 4 is independent of specific location of scalar k, so attacker can't get the information about scalar k through the side channel information. Algorithm 4 can resist the SPA attack. Compared with Algorithm 1, the proposed Algorithm 4 improves security and efficiency of ECC system.

3.2 Tripling-Add Operation in Jacobian Coordinate

Computing $3P + Q$. The Literature [20] proposed $3P + Q$ formula in affine coordinate, in order to avoid inverse operation, let $3P+Q = (X_3, Y_3, Z_3)$, according to $3P+Q$ formula in affine coordinate, we get:

$$\begin{cases} X_3' &= \frac{C^2 - B(DZ_2)^2 - (Z_1^7 AD)^2 (aZ_2^2 - X_2)}{(Z_1^7 ADZ_2)} \\ Y_3' &= \frac{C}{Z_1^7 ADZ_2} (\frac{X_3}{Z_3^2} - \frac{X_2}{Z_2^2} - \frac{Y_2}{Z_2^3}) \end{cases}$$

where $A = a(X_1 + bZ_1^2); B = (X_1^3 + bZ_1^6)^3 - Z_1^{12} a^3 bX_1^3; C = Z_2^3[Y_1^9 - a^3 Z_1^6 Y_1^3 (X_1^3 + bZ_1^6)^2] - Y_2(Z_1^7 A)^3; D = BZ_2^2 - X_2(Z_1^7 A)^2$, Then we take $Z_3 = Z_1^7 ADZ_2$, The formula of $3P + Q$ is derived as followed:

$$\begin{cases} X_3 &= C^2 - B(DZ_2)^2 - (Z_1^7 AD)(aZ_2^2 - X_2) \\ Y_3 &= C[X_3 - X_2(Z_1^7 AD)^2] + Y_2(Z_1^7 AD)^3 \\ Z_3 &= Z_1^7 ADZ_2. \end{cases}$$

By storing the intermediate results and setting $a = 1$, it can be seen that the computation of $3P + Q$ is $23M + 4C + 8S$ instead of $1I + 13M + 5S + 6C$. We use combined tripling-add operation in Jacobian coordinate to replace point addition and tripling operation, and get an improved balanced ternary scalar multiplication algorithm.

Algorithm 5 Improved balanced ternary scalar multiplication algorithm

Input: $P = (X, Y, Z) \in E(GF(3^m)), and \quad k = \sum_{i=0}^{n-1} k_i 3^i, k_i \in (-1, 0, 1)$
Output: kP
$Q \leftarrow O$
for $i = n - 1, ..., 0$ **do**
  if $k_i = 1$ **then**
    $Q = 3Q + P$
  end if
  if $k_i = -1$ **then**
    $Q = 3Q - P$
  end if
  if $k_i = 0$ **then**
    $Q = 3Q$
  end if
end for
Return Q
End

The computation amount of Algorithm 5 is $\frac{1}{3}nT + \frac{2}{3}n \cdot TA$, where TA represents tripling-add operation.

4 Performance Analysis

In this section, we analyze the computational efficiency of improved scalar multiplication algorithm on the elliptic curve over GF(3^m). First, Table 1 shows the comparisons of calculation costs in different coordinate systems. It can be seen from Table 1, the newly proposed Co_Z point addition formula is more efficient than traditional point addition.

In order to analyze the efficiency better, we choose the ternary length of scalar k to be 101 bits, and consider the typical ratio of inversion and multiplication: $I = 8M$. The calculation comparisons of different scalar multiplication algorithms are given in Table 2, it can be seen from Table 2, that compared with other scalar multiplication algorithms, the computational efficiency of our Algorithm 4 is increased by 16%, 26%, 10%, 3%, respectively.

Second, before we analyze the efficiency of $3P + Q$ in Jacobian coordinate, We define α is the ratio of field in-

Table 1: Comparisons of computation costs in different coordinate systems

Operation	**Jacobian projective coordinate**	**Lopez &Dahab projective coordinate**
Point addition	$10M + 3C + 3S$	$13M + 5S$
Mixed addition($Z_2 = 1$)	$7M + 2C + 2S$	$10M + 4S$
Co_Z point addition	$4M + 1C + 3S$	$8M + 3S$
Tripling	$5M + 2S + 5C$	$7M + 1S + 5C$

Table 2: Computation costs of different scalar multiplication algorithms

Algorithm	**Total cost**	**Anti-SPA**	**n=101bits**
BTF Algorithm [4]	$n(I + 4S + 7M) + \frac{2}{3}n(I + S + 2M)$	no	$2565M$
STF Algorithm [14]	$n(I + 4S + 7M) + n(I + S + 2M)$	yes	$2929M$
Algorithm 3 in Ref. [13]	$n(9M + 7S) + \frac{n}{6}(4M + 6S) + \frac{n}{2}(12M + 4S)$	no	$2390M$
Co_Z Montgomery Algorithm(binary) [19]	$n(10M + C + 3S) + I + 10M + S + C$	yes	$2223M$
Proposed Algorithm 4	$n(5M + 2S + 5C) + n(4M + 1C + 3S) + 2(5M + 3S)$	yes	$2158M$

Table 3: Calculation comparisons of different algorithms

Bit legth	**Algorithms**		
	BTSM Algorithm [20]	Ternary scalar multiplication Algorithm [20]	Proposed Algorithm 5
101	$3005M$	$3407M$	$2801M$
122	$3630M$	$4113M$	$3384M$
142	$4225M$	$4786M$	$3939M$
162	$4820M$	$5458M$	$4493M$

version and multiplication. The efficiency can be derived from the following equation:

$$efficiency = 1 - \frac{34.88}{25.22+\alpha}$$

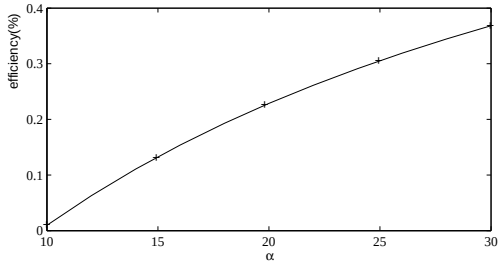


Figure 1: Efficiency of $3P + Q$ formula in Jacobian coordinate

Figure 1 shows the efficiency analysis of $3P + Q$ formula in Jacobian coordinate and affine coordinate. In Table 3, we choose the scalar symmetric ternary length as 101,122,142 and 162. When the scalar length of the ternary expansion is 162, it can be seen from Table 3 that our proposed Algorithm 5 improves the computational efficiency by 7% and 18% compared with the BTSM algorithm and the ternary scalar multiplication algorithm.

5 Conclusions

In this paper, the Co_Z point addition formula in different projective coordinates on elliptic curve over GF(3^m) is proposed. Compared with previously mixed point addition formulas, the computational time is reduced. Combined the new formulaes and optimized symmetric ternary scalar multiplication algorithm against SPA, the computational efficiency is improved. Compared with other algorithms, the efficiency of Algorithm 4 is increased by 16%, 26%, and 10%, respectively. We also proposed $3P+Q$ formula in Jacobian coordinate. The average running time of the improved balanced ternary scalar multiplication algorithm is about 7% faster than that of the BTSM algorithm.

Acknowledgments

The support of NSFC (National Natural Science Foundation of China, No.61872058), Shaanxi Natural Science Foundation (No.2017JQ6010) is gratefully acknowledged.

References

[1] G. Chen, G. Bai, and H. Chen, "A high-performance elliptic curve cryptographic processor for general curves over GF(p) based on a systolic arithmetic

unit," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 54, pp. 412–416, May. 2007.

[2] J. J. Cheng, F. Y. Zheng, and J. Q. Lin, "High-performance implementation of Curve25519 on GPU," *Netinfo Security (in Chinese)*, no. 9, pp. 122–127, 2017.

[3] D. V. Chudnovsky and G. V. Chudnovsky, "Sequences of numbers generated by addition informal groups and new primality and factorization tests," *Advances in Applied Math*, vol. 7, no. 4, pp. 385–434, 1986.

[4] W. Y. Deng and X. H. Miao, "Application of balanced ternary in elliptic curve scalar multiplication," *Computer Engineering (in chinese)*, vol. 38, no. 5, pp. 152–154, 2012.

[5] Q. L. Du, "Identity-based dynamic group communication signcryption scheme," *Netinfo Security(in Chinese)*, no. 9, pp. 42–44, 2017.

[6] Q. F. Duanmu, X. Y. Zhang, Y. B. Wang, and K. Z. Zhang, "Fast arithmetic operations of elliptic curve over GF(3^n)," *Journal of PLA University of Science and Technology (Natural Science Eddition) (in Chinese)*, vol. 12, no. 01, pp. 1–6, 2011.

[7] R. R. Goundar, M. Joye, and A. Miyaji, "Co_Z addition formula and binary ladders on elliptic curves," in *International Workshop on Cryptographic Hardware and Embedded Systems*, pp. 65–79, 2010.

[8] H. Hisil, G. Carter, E. Dawson, "New formulae for efficient elliptic curve arithmetic," in *International Conference on Cryptology in India*, pp. 138–151, 2007.

[9] M. S. Hossain, Y. N. Kong, E. S, and N. V, "High-performance elliptic curve cryptography processor over NIST prime fields," *IET Computers and Digital Techniques*, vol. 11, no. 1, pp. 33–42, 2017.

[10] Z. X. Lai and Z. J. Zhang, "Scalar multiplication on hessian curves based on Co_Z operations," *Bulletin of Science and Technology (in Chinese)*, vol. 32, no. 2, pp. 28–33, 2016.

[11] L. Li and T. Zhan, "Co_Z addition operation of hessian curve," in *Seventh International Conference on Computational Intelligence and Security*, pp. 915–919, Dec. 2011.

[12] Q. W. Li, Z. F. Wang, and X . C. Liu, "Fast point operation architecture for elliptic curve cryptography," in *IEEE Asia-Pacific Conference on Circuits and Systems*, pp. 184–188, Nov. 2008.

[13] H. Liu, Q. Dong, and Y. Li, "Efficient ECC scalar multiplication algorithm based on symmetric ternary in wireless sensor networks," in *Progress in Electromagnetics Research Symposium - Fall (PIERS - FALL)*, pp. 879–885, Nov. 2017.

[14] S. Liu, H. Yao, and X. A. Wang, "SPA resistant scalar multiplication based on addition and tripling indistinguishable on elliptic curve cryptosystem," in *The 10th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC)*, pp. 785 – 790, Nov. 2015.

[15] Meloni and Nicolas, "New point addition formulae for ECC applications," in *International Workshop on the Arithmetic of Finite Fieldss*, pp. 189–201, June. 2007.

[16] B. Y. Peng, Y. C. Hsu, Y. J. Chen, and D. C. Chueh, "Multi-core FPGA implementation of ECC with homogeneous Co-Z coordinate representation," in *The 15th International Conference on Cryptology and Network Security*, pp. 637–647, Nov. 2016.

[17] C. S. Sin, "Regular ternary algorithm for scalar multiplication on elliptic curves over finite fields of characteristic three," in *Technical Report Iacr Cryptology Eprint Archive*, July 2012. `https://eprint.iacr.org/2012/390.pdf`

[18] M. Wang and Z. Wu, "Algorithm of NAF scalar multiplication on ECC against SPA," *Journal on Communications (in Chinese)*, vol. 33, no. S1, pp. 228–232, 2012.

[19] W. Yu, B. Li, K. W. Wang, and W. H. Li, "Co_Z montgomery algorithm on elliptic cureves over finite fields of characteristic three," *Chinese Journal of Computers (in Chinese)*, vol. 40, no. 5, pp. 1121–1131, 2017.

[20] N. Zhang and X. T. Fu, "Ternary method in elliptic curve scalar multiplication," in *The 5th International Conference on Intelligent Networking and Collaborative Systems*, pp. 490–494, Sep. 2013.

[21] Z. B. Zhou, S. B. Zhang, and E. T. Luo, "A group RFID tag ownership transfer protocol without trusted third party," *Netinfo Security (in Chinese)*, no. 6, pp. 18–27, 2018.

Biography

Shuang-Gen Liu, born in 1979, Ph.D, associate professor. A member of the China Computer Federation, and a member of the Chinese Association for Cryptologic Research. His main research interests focus on information security and cryptography.

Shi-Mei Lu, born in 1995. A graduate student of Xi'an University of posts and telecommunications. She is mainly engaged in the research of elliptic curve cryptosystem.

Rui-Wen Gong, born in 1998. An undergraduate student in information security in Xi'an University of posts and telecommunications. Her main research interest is cryptography.

A Reversible Data Hiding Method for SMVQ Indices Based on Improved Locally Adaptive Coding

Chin-Chen Chang[1], Jun-Yong Chen[1], Yan-Hong Chen[2], and Yanjun Liu[1]
(Corresponding author: Yanjun Liu)

Department of Information Engineering and Computer Science, Feng Chia University[1]
Taichung 40724, Taiwan
(Email: yjliu104@gmail.com)
School of Information, Zhejiang University of Finance and Economics[2]
Hangzhou 310018, China
(Received Jan. 31, 2018; Revised and Accepted Aug. 18, 2018; First Online July 16, 2019)

Abstract

In this paper, we propose a reversible data hiding scheme based on improved locally adaptive coding for compressed images by side match vector quantization (SMVQ). In the proposed scheme, an indicator is defined to encode the SMVQ indices and embed the secret data. The smaller the index is, the more bits of secret data could be hidden. In order to improve the capacity, we use an improved adaptive coding method after the assigned indices. Experimental results show that our proposed scheme outperforms state-of-the-art VQ-based data hiding schemes in terms of embedding capacity.

Keywords: Data Hiding; Image Compression; Locally Adaptive Coding; Reversibility; Side Match Vector Quantization (SMVQ)

1 Introduction

Due to the advance of Internet technology and the convenience of transmitting information, data hiding has become a popular research issue in recent years. Generally speaking, data hiding schemes can be classified into three types, *i.e.*, spatial domain, frequency domain and compression domain based schemes. Data hiding schemes in the spatial domain embed secret information by directly modifying pixel values of an image, while those in the frequency domain first transform an image from the spatial domain to the frequency domain and then modify coefficient values. However, the compressed domain based schemes embed data into the compressed codes of digital images [9]. Nowadays, many researchers have proposed various information hiding methods designed for the compression domain, such as block truncation coding (BTC) [12, 21, 22], discrete wavelet transform (DWT) [1, 2, 20], discrete cosine transform (DCT) [6, 14, 15] and vector quantization (VQ) [8, 10].

Wu and Sun [22] presented a data hiding method in which each secret bit is embedded into the bitmap of the BTC compression codes. Lin and Liu [12] proposed a data hiding method in BTC compressed images using the order of each pair of gray levels to embed the secret data. Wang *et al.* [21] presented a reversible data hiding scheme for images compressed by BTC based on the correlation among adjacent blocks and prediction-error expansion. Abdelwahab and Hassaan [1] proposed a data hiding method based on DWT that hides secret images inside the cover image using two secret keys. Vijay and VigneshKumar [20] proposed a method that first employs DWT to transform the cover image from the spatial domain to the frequency domain, and then compresses the secret data by the Huffman code for embedding. Baby *et al.* [2] proposed a DWT-based data hiding scheme using multiple color images to hide the secret data. Chang *et al.* [6] proposed a scheme to hide secret information in the DCT coefficients. In this scheme, the image is divided into 8×8 blocks. If two successive coefficients of the medium-frequency components are zero, the information is hidden in each block. Lin [15] proposed a method that embeds the data using the DCT coefficients and integer mapping. Lin [14] presented a method that uses histogram shifting to embed the secret data based on two-dimensional DCT coefficients.

VQ is another efficient image compression technique that was proposed by Gray [10]. In recent years, many researchers have studied data hiding based on VQ [3–5, 7, 8, 11, 13, 16–19, 23, 24]. In 2009, Chang *et al.* [5, 8] presented two reversible data hiding schemes based on VQ. They used joint neighbor coding to hide data [8] and employed a locally adaptive coding scheme (LAS) to reduce the cost of the compression code [5]. Also in 2009, Yang

and Lin [23] proposed a VQ-based reversible information hiding method in which an indicator for data hiding is defined according to the embedding rule and the corresponding index. To improve the method in Chang and Chou's scheme [5], in 2010, Yang and Lin [24] presented a data hiding scheme that changes the hiding path to further enhance the effect of LAS. In 2011, Chang *et al.* [4] proposed an improved method that modifies the embedding rule to increase the embedding capacity of the method in [5]. In 2013, Pan *et al.* [17] presented another VQ-based method which uses search order coding (SOC) to compress the indices, and then utilizes the remainder of the space to embed the secret data. In 2014, Chang and Nguyen [7] introduced the concept of side match vector quantization (SMVQ) and proposed a novel data hiding scheme based on SMVQ that uses fewer indices than [17] when the same data are embedded. In 2015, Tu and Wang [19] proposed a method that divides all indices into two clusters for embedding. Their scheme can reduce extra bits and increase the hiding capacity. In 2016, Qin and Hu [18] proposed a method that uses an improved searching order coding (ISOC) encoded VQ index table. This method uses a lower bit rate and extends the degree of the index table to achieve higher hiding capacity.

However, the embedding capacities of the aforementioned data hiding schemes based on VQ are limited. Therefore, in this paper, we propose a VQ-based reversible data hiding scheme to enhance the embedding capacity. In the proposed scheme, we combine the SMVQ and LAS techniques to achieve a high embedding capacity.

The remainder of this paper is divided into five sections. Section 2 shows two techniques that our proposed reversible information hiding scheme is based on, *i.e.*, the SMVQ and LAS methods. Section 3 elaborates our proposed scheme. The detailed experimental description and comparative analysis are provided in Section 4, and our conclusions are presented in Section 5.

2 Related Work

In this section, we will introduce VQ, SMVQ and LAS techniques in Sections 2.1, 2.2 and 2.3, respectively.

2.1 VQ

VQ is a simple, lossy compression technology that is commonly used in reversible data hiding schemes [13]. The algorithm of VQ is divided into three phases. In the codebook generation phase, some images are used for training and to generate a codebook using the Linde-Buzo-Gary (LBG) algorithm, which was proposed by Linde *et al.* [16] in 1980. The LBG algorithm is an iterative procedure in which, first, it divides the image into a set of blocks with the size of $r \times r$, and each block can be viewed as an $r \times r-$ dimensional vector. Then, 256 blocks are selected randomly from these blocks as the initial codebook, with these 256 initial vectors as the 256 centroids. The rest of the blocks are grouped into the 256 centroids, *i.e.*, each block is used to find the nearest centroid to form 256 groups. Then, the centroids of these 256 groups are recalculated to get a new codebook until the codebook converges, *i.e.*, when the training of the codebook is complete. In the compression phase, the image is divided into many blocks of the same size containing many pixels. These blocks are used to search for similar codewords in the codebook, and the vector is replaced to form an index table. These procedures are explained by an example in Figure 1. In the decompression phase, the index values are used to extract the corresponding vectors from the codebook, and the vectors are used to form blocks that are then merged into an image.

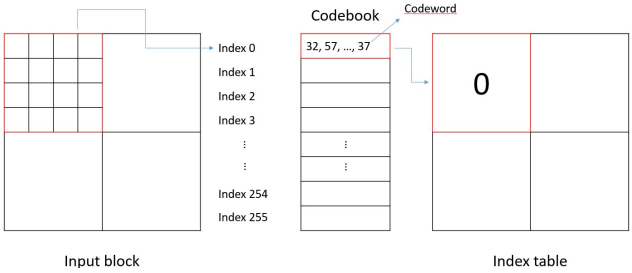


Figure 1: An example of VQ

2.2 SMVQ

As a variant of VQ, SMVQ was proposed by Kim [11] in 1992. In the encoding procedure, the algorithm divides the cover image into a set of non-overlapping blocks with the size of $c \times c$. These blocks are divided into two parts, *i.e.*, seed blocks and residual blocks. The image blocks in the first row and first column are denoted as seed blocks, and the rest of blocks are defined as residual blocks. As shown in Figure 2, the upper block U and the left block L are used to predict the current block X. After the prediction, let $x_1 = \frac{(u_{13}+l_4)}{2}, x_2 = u_{14}, x_3 = u_{15}, x_4 = u_{16}, x_5 = l_8, x_9 = l_{12}$, and $x_{13} = l_{16}$. Then, $x_1, x_2, x_3, x_4, x_5, x_9$ and x_{13} are used to find the most similar codeword from the traditional VQ codebook and a state codebook by using the most similar codewords is constructed. The codeword in the state codebook with the minimum Euclidean distance from X is used to encode X.

2.3 LAS

LAS is a data compression method that was proposed by Bentley *et al.* [3] in 1986. An example of the LAS is shown in Table 1. Suppose we want to compress the message "I HAVE A PEN I HAVE AN APPLE". Then, by using LAS, we can get the compressed message "1 I, 2 HAVE, 3 A, 4 PEN, 4, 4, 5 AN, 6 APPLE." The encoding steps are described as follows. First, list L is empty and is denoted as $L = \{\}$, and the size of list L is $S = 0$.

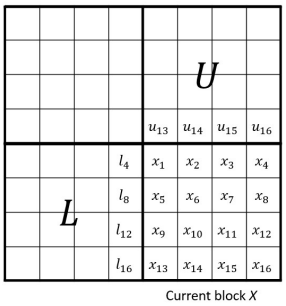


Figure 2: An illustration of SMVQ

We use the list L to encode the input word x. If the input word x doesn't belong to L, then we will use $S+1$ concatenated with x to encode it and x is inserted at the front of list L. Otherwise, the input word x is encoded by S, indicating the position of the input word x in the list L, and x is moved to the front of list L. The encoding process is repeated for the next word until all input words of the message are encoded.

Table 1: An example of LAS

Step	Input	The list $L = \{\}$	Output
1	I	$L = \{$I$\}$	1 I
2	HAVE	$L = \{$HAVE I$\}$	2 HAVE
3	A	$L = \{$A HAVE I$\}$	3 A
4	PEN	$L = \{$PEN A HAVE I$\}$	4 PEN
5	I	$L = \{$I PEN A HAVE$\}$	4
6	HAVE	$L = \{$HAVE I PEN A$\}$	4
7	AN	$L = \{$AN HAVE I PEN A$\}$	5 AN
8	APPLE	$L = \{$APPLE AN HAVE I PEN A$\}$	6 APPLE

3 Proposed Scheme

The proposed scheme is elaborated in this section. First, we propose an improved locally adaptive coding scheme (NILAS) by using the feature of SMVQ. Then, the embedding procedure and the extraction procedure are shown in Sub-sections 3.2 and 3.3, respectively.

3.1 NILAS

In the proposed NILAS, first we set the list L as not empty, and its length as equal to the largest value in the SMVQ indices. (For example, $L = \{0, 1, \ldots, N-1\}$, and N is equal to the largest value in the SMVQ indices. If the state codebook size is equal to 256, then the value of N is 256.) Based on the statistical analysis of different types of images, we found that the referred frequencies of all indices in SMVQ index table were close to the value in the range from zero to nine. Therefore, we divide L into two parts, *i.e.*, $L1$ and $L2$, where $L1 = \{0, 1, \ldots, 9\}$ and $L2 = \{10, 11, \ldots, N-1\}$. The order of $L1$ remains unchanged, and the move-to-front method is only applied in the order of $L2$. An example is provided in Figure 3.

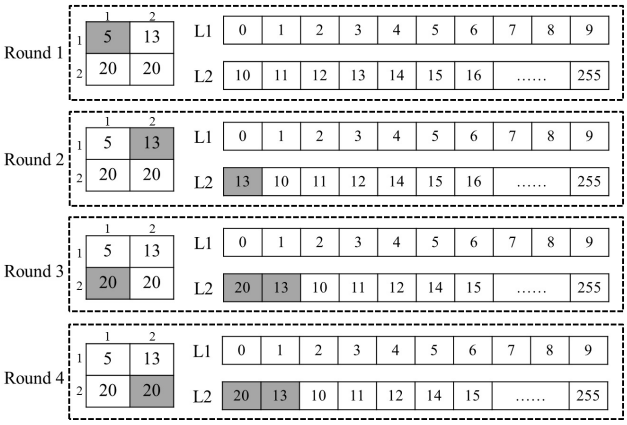


Figure 3: An example of NILAS

In Figure 3, we show a set of 4 blocks, and two lists $L1$ and $L2$. There are four rounds and each round processes a block. In particular, the gray block represents the current processing block. In the first round, the index value of the current block is 5, which belongs to $L1$. Thus, its order is unchanged. In the second round, the index value of the current block is 13. Since it belongs to $L2$, the move-to-front scheme is applied and it moves to the head of $L2$. Then, in the third round, the index value of the current block is 20. Since it belongs to $L2$, it also moves to the head of $L2$. Lastly, the index value of the current block is 20, which belongs to $L2$. Because its order is in the front of $L2$, the sequence will be unchanged.

3.2 Embedding Phase

Figure 4 shows the embedding flowchart of our method. I represents a cover image sized $H \times W$. Secret data S is a bit stream, and s_l represents the bit value in S, where $s_l = 1$ or $0, l = 0, 1, \ldots, M$, and M is the maximum embedding capacity in bits to the cover image I.

Input: A cover image I sized $H \times W$, codebook CB, and secret data S.

Output: The stego code stream CS.

Step 1: Encode I utilizing the SMVQ algorithm to obtain the SMVQ index table, IT. V_i is the value of each index in index table IT, where $i = 0, 1, \ldots, \frac{H}{4} \times \frac{W}{4} - 1$. Set the $L = \{0, 1, \ldots, N-1\}$ and divide L into two parts, $L1 = \{0, 1, \ldots, 9\}$ and $L2 = \{10, 11, \ldots, N-1\}$, where N is the size of codebook CB. Apply the NILAS method to $IT, L1$ and $L2$.

Step 2: Read the current V_i in IT. If the position of the current V_i is in the first row or the first column, i is increased by 1 and read the next V_i.

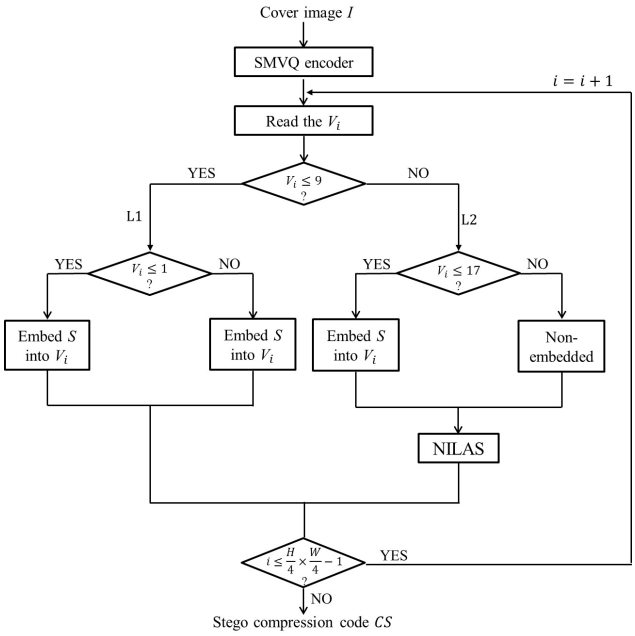


Figure 4: Flowchart of the embedding phase

Step 3: Check the value of V_i. If $V_i \le 1$, go to **Step 4**. If $2 \le V_i \le 9$, go to **Step 5**. Otherwise, go to **Step 6**.

Step 4: Add 00 to the head of the value of V_i as an indicator. That is, the first two bits are indicator bits and the third bit is the value of V_i. The remaining $n-3$ bits are used to embed the next $s_{l,l+1,\ldots,l+(n-3)-1}$ from secret data S, where $n = \log_2\lfloor N \rfloor$. They will be encoded by $00\|V_i\|s_{l,l+1,\cdots,l+(n-3)-1}$, where $\|$ represents the concatenation operation. Go to **Step 8**.

Step 5: Add 01 to the front as an indicator. The next three bits are the value of $V_i - 2$ in binary form. And the remaining $n-5$ bits are used to embed the next $s_{l,l+1,\ldots,l+(n-5)-1}$ from secret data S. They will be encoded by $01\|V_i\|s_{l,l+1,\ldots,l+(n-5)-1}$. Go to **Step 8**.

Step 6: Check the position p_i of current V_i in $L2$, where p_i is the position of V_i in $L2$. If $8 \le p_i \le N-11$, go to **Step 7**. Otherwise, if $0 \le p_i \le 7$, add 10 to the front as an indicator. Then, use three bits to encode the position p_i in binary form. The next $n-5$ bits are used to embed the next $s_{l,l+1,\ldots,l+(n-5)-1}$ from secret data S. They will be encoded by $10\|p_i\|s_{l,l+1,\ldots,l+(n-5)-1}$. Then, move V_i to the front of $L2$. Go to **Step 8**.

Step 7: Add 11 to the front as an indicator.. Then, encode the position p_i in binary form with n bits. They will be encoded by $11\|p_i$. Then, move V_i to the front of $L2$.

Step 8: Repeat **Steps 2** through **7** until all V_i's have been processed.

Step 9: Output the stego compression code CS.

After all of the steps have been completed, we get the stego compression code CS. To further clarify our embedding phase, an example of the embedding process is provided. In Figure 5(a), we assume that our SMVQ index table, IT, has the size of 3×3. Figures 5(b) and (c) show the secret data S that we want to embed in the SMVQ index table, IT, and the result after embedding, respectively.

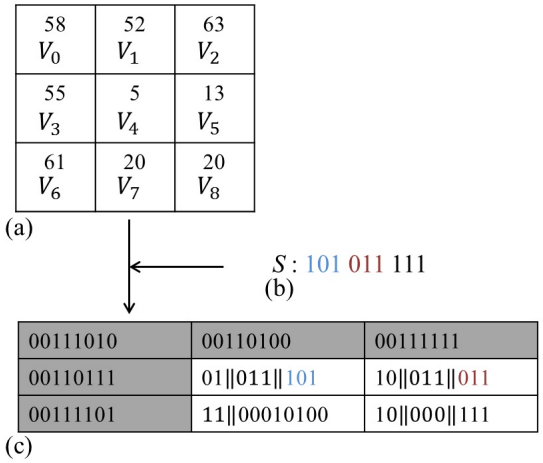


Figure 5: An example of the data embedding phase: (a) SMVQ index table IT; (b) Secret data S; (c) Stego compression code CS

3.3 Extraction Phase

The extraction phase, as shown in Figure 6, is explained in this section. Because the indicator of two bits has been inserted during data embedding phase, a decoder can choose the corresponding extraction operation according to the indicator. If the first bit of the current processing stego index V_i', *i.e.*, the first bit of the indicator, is 0, the current processing stego index V_i' is in $L1$. Otherwise, it is in $L2$. Then, according to the second bit of indicator, it determines the extraction and recovery method. Moreover, if the current processing stego index V_i' is in $L2$, it must be moved to the front of $L2$ after extracting and recovery.

Input: Codebook CB and the stego compression code CS.

Output: The secret data S, reconstructed SMVQ index table IT, and cover image I.

Step 1: Check the position of current processing stego index V_i'. If the position is at the first row or the first column, i is increased by 1 and read the next current processing stego index V_i'.

Step 2: Read the first bit of current processing stego index V_i'. if it is 0, go to **Step 3.** Otherwise, go to **Step 5**.

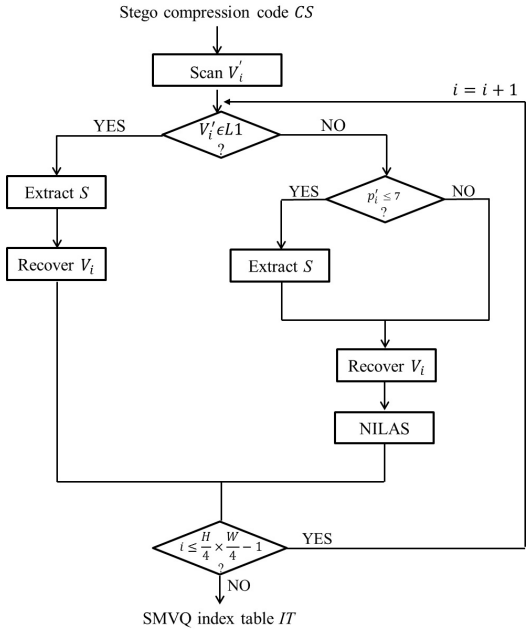


Figure 6: Flowchart of the data extraction phase

Step 3: Check the value of the second bit of current processing stego index V_i'. If it is 0, use the value of the third bit of the current processing stego index V_i' to recover the index V_i. Then, the remainder of $n-3$ bits, *i.e.*, $s_{l,l+1,\ldots,l+(n-3)-1}$, will be extracted and appended to the secret data S. Go to **Step 7**. Otherwise, go to **Step 4**.

Step 4: Check the value of the second bit of current processing stego index V_i'. If it is 1, the next 3 bits are converted into decimal value to recover the index V_i. The remainder, $n-5$ bits, *i.e.*, $s_{l,l+1,\ldots,l+(n-5)-1}$, will be extracted and appended to the secret data S. Go to **Step 7**.

Step 5: Get the second bit of current processing stego index V_i'. If it is equal to 0, the next 3 bits will be transformed into decimal value p_i'. Use the corresponding value of p_i' in $L2$ to recover the index V_i. Extract the remaining $n-5$ bits, $s_{l,l+1,\ldots,l+(n-5)-1}$, and insert them into secret data S. Then, move V_i to the front of $L2$. Go to **Step 7**. Otherwise, go to **Step 6**.

Step 6: Convert the remaining n bits into decimal value p_i'. Utilize p_i'' corresponding value in $L2$ to recover the index V_i and move V_i to the front of $L2$.

Step 7: Repeat **Steps 1** through **6** until all bits in the stego compression code have been read.

Step 8: Obtain secret data S and the SMVQ index table IT.

After all of the steps have been completed, we can obtain the secret data S and the reconstructed SMVQ index table IT. To further clarify our extraction and recovery phase, we present an example of the extracting process. Figures 7(a), (b) and (c) show the stego code stream, the reconstructed SMVQ index table, IT, and secret data S, that we can extract the reconstructed SMVQ index table IT and secret data S from the stego code stream, respectively.

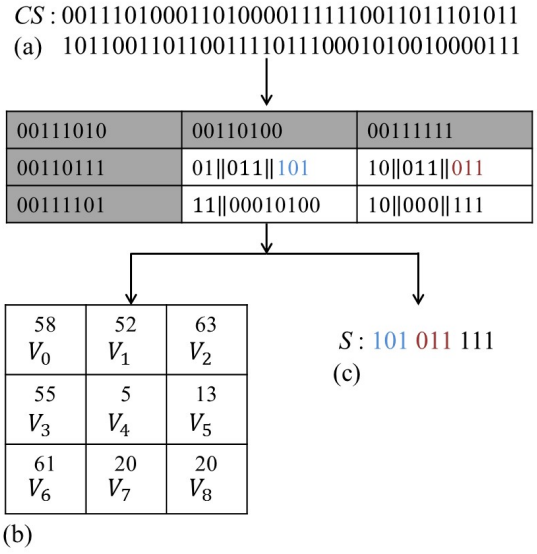


Figure 7: An example of the data extraction phase: (a) Stego code stream; (b) Reconstructed SMVQ index table IT; (c) Secret data S

4 Simulation and Analysis

We chose the software MATLAB R2016 to conduct the experiments, and used a computer with an Intel-Core i7-6700 3.40GHz and 32GB RAM. The operating system was the 64-bit Windows 10 Pro. To verify the efficiency of the proposed scheme, several experiments are performed. Six 512×512-sized, 8-bit grayscale images were used as the test images. These images were "Lena", "Peppers", "F-16", "Toys", "Girl" and "Sailboat", as shown in Figure 8. These images were smooth images, which have been used extensively in image processing. Rough images would generate a negative value of payload, so we used the smooth images to keep the value positive, the bit rate lower, and the embedding rate higher. The images were partitioned into non-overlapping 4×4 blocks, and the codebook with a size of 256 was trained using the LBG algorithm. In our experiments, the secret data S were in binary form and the data were generated randomly by using a pseudo-random number generator and encrypted by conventional methods.

4.1 Experimental Results

In these experiments, the embedding efficiency was evaluated by the payload and the embedding rate, and the

Figure 8: Test images: (a) Lena; (b) Peppers; (c) F-16; (d) Toys; (e) Girl; (f) Sailboat

compression efficiency was evaluated by the bit rate. In order to show that our proposed scheme can hide more secret data than other methods, payload was defined as the number of secret bits that can be hidden when the bit rate is fixed at 0.5, *i.e.*, the larger the payload is, the better the result is, as shown in Equation (1). The bit rate indicates the compression rate of the cover image by using bits per pixel (bpp). So, the smaller the bit rate is, the greater the compression efficiency is, as defined in Equation (2). The embedding rate is a ratio of the size of embedded secret data to the size of code stream, and it denotes whether the size of the code stream is fixed. The higher the embedding rate is, the more secret bits can be embedded, as defined in Equation (3).

$$Payload = Capacity - (0.5 - bit\,rate) \times Size\,of\,original\,image. \quad (1)$$

$$Bit\,rate = \frac{Size\,of\,code\,stream}{Size\,of\,original\,image}. \quad (2)$$

$$Embedding\,rate = \frac{Size\,of\,embedded\,secret\,data}{Size\,of\,code\,stream} \times 100\%. \quad (3)$$

To investigate the superiority of our proposed scheme, we compared it with some VQ-compressed-based data hiding methods, and the compressed code was used to achieve compression and transformation [4, 5, 7, 18, 19, 23, 24]. Because the capacities of other methods are not high, we proposed a high capacity method combined with SMVQ and NILAS, and, even though the bit rates were identical, our method still provided the highest capacity. Figure 9, Table 2, Table 3 and Figure 10 compare our proposed scheme's hiding capacity, payload, embedding rate and bit rate with those of Chang *et al.*'s schemes [4, 5, 7], Yang and Lin's schemes [23,24], Qin and Hu's scheme [18], and Tu and Wang's scheme [19].

As shown in Table 2 and Figure 9, our experimental results indicated that the capacity of our method was higher than that of all of the other methods. We achieved this

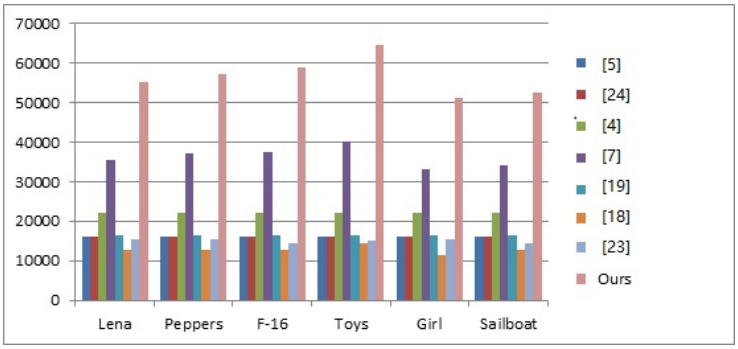


Figure 9: Comparisons of capacity (bits) between our method and other methods

result because we used the frequency of the occurrences of the SMVQ indices and clever application of the indices to hide more secret bits. The capacity of our method was still greater than that of other methods even under the same bit rate. The payload is defined that, when the value of bit rate is 0.5bpp, the number of secret data that can be hidden. When the sizes of code stream of all schemes are identical, we can get the value of payload according to the total number of the embedded secret bits. The experimental results showed that the capacity and payload of our scheme were much better than the other VQ-based methods.

Table 3 compares the embedding rate of our method and other methods. The higher the embedding rate was, the more secret bits could be hidden. It means that the embedding rate depends on the code stream. It can be seen that our proposed method is better than the other methods irrespective of how many code streams are embedded.

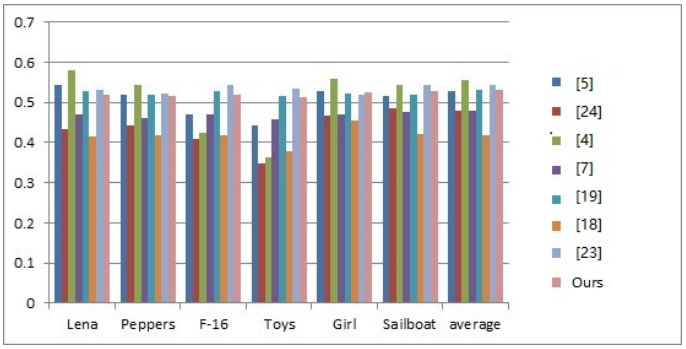


Figure 10: Comparisons of the bit rate (bpp) of our method with other methods

Figure 10 compares the bit rate of our method with other methods. According to Equation (2), we can know that the bit rate represents the compression rate. If the value of bit rate is low, the compression rate is better. Although compression rate of some methods are better than our method, the capacity of our method is better

Table 2: Comparison of the payload (bits) of our method and other methods

Image	[5]	[24]	[4]	[7]	[19]	[18]	[23]	Ours
Lena	4360	33982	850	43442	8738	34692	7521	49794
Peppers	11166	31089	10277	47700	11403	33940	9362	52614
F-16	23763	39754	42274	45421	9426	34555	3546	53786
Toys	31086	55728	58244	51191	11852	46770	6280	61025
Girl	9077	25068	6617	41147	10878	23449	10264	44906
Sailboat	11951	20339	10536	40632	11141	33684	2627	45215
average	8604	21740	7655	37404	8752	34558	2758	42244

Table 3: Comparison of the embedding rate (bpp) of our method and other methods

Image	[5]	[24]	[4]	[7]	[19]	[18]	[23]	Ours
Lena	0.113	0.143	0.145	0.289	0.118	0.116	0.111	0.404
Peppers	0.119	0.139	0.154	0.309	0.120	0.116	0.112	0.421
F-16	0.131	0.150	0.199	0.305	0.119	0.117	0.102	0.433
Toys	0.139	0.176	0.233	0.335	0.121	0.147	0.107	0.480
Girl	0.117	0.132	0.151	0.270	0.120	0.096	0.114	0.372
Sailboat	0.119	0.127	0.155	0.274	0.120	0.115	0.101	0.379
average	0.118	0.133	0.159	0.259	0.118	0.118	0.102	0.363

under the same bit rate. That is the reason why the payload is more important to represent the embedding effect of the data hiding scheme.

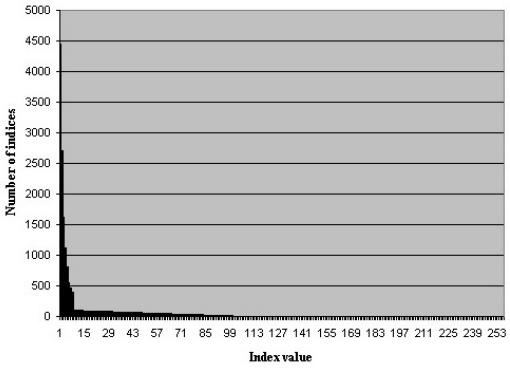


Figure 11: Histogram of the frequency of the occurrences of SMVQ indices

Figure 11 shows that we obtained the highest frequency of SMVQ index occurrences, *i.e.*, from 0 to 17. Therefore, we embedded most of the secret data in indices 0 through 9, and, after index 10, we used NILAS to improve the embedding capacity. Then, an indicator was defined and used to encode the indices and embed the secret data as much as possible in indices 0 or 1. That is to say, for indices 0 and 1, we can hide 5 bits; for indices 2 to 9 and 10 to 17, we can hide 3 bits. Then, in order to improve the embedding capacity after index 10, we used improved adaptive coding to increase the embedding capacity. This technology can use the feature of recurring occurrences of SMVQ indices to increase the embedding capacity. If the test image is smooth, we can hide more bits. Therefore, the experimental results showed that more secret data were embedded while the bit rate was the same, providing a higher capacity than schemes in the previous literature.

4.2 Discussions

In Chang and Wu's method [8], each index only can be embedded with one secret bit, but their method requires more bits to present the index in some cases. So, the payload of their method is lower than that of our method. In Chang and Chou's method [5], the largest capacity of each index is 1, and some indices cannot be embedded with secret bits if the index is not in the list L. Although their method can reduce the bit rate, the payload of their method is lower. Yang and Lin's method [24] improved the method in [5] by changing the run path of the normal VQ encoder. Although their bit rate is lower than Chang and Chou's method [5], their payload is still lower than that of our proposed method. The improved run path and the normal path are shown in Figure 12.

In Chang *et al.*'s method [4], although two secret bits can be embedded in some cases, it still requires more bits to represent the index. The largest capacity of their method is two, which is lower than that of our method. In Chang and Nguyen's method [7], the largest capacity of the indices is 3, while it is 5 for our proposed method. In their method, although the indices belong to the cases that can embed secret bits, the indices only need to use 7 bits to represent the index. But the payload of their method is lower than that of our proposed method. In Tu and Wang's method [19], although they used trained images to obtain a better distribution of codewords, clus-

ter 2 and cluster 3 had to embed an extra 2 bits; so the payload of our proposed method was better than their method. In Qin and Hu's method [18], an improved search order coding-encoded VQ index table was used that could effectively embed secret data. However, their bit rate and payload were lower than those of our proposed method. In Yang and Lin's method [23], they resorted the VQ codebook according to the referred frequency of each index. , Their method can be mapped directly to the other clusters though the index is in the first cluster, thereby reducing the presented bits of the index. If the index is not in the first cluster, their method must use more bits to represent the index. The largest capacity of the indices in their method is 2, which is lower than that of our method. So our method is better than Yang and Lin's method [23]. However, our proposed method utilized only one bit to represent the indices and hided five secret bits most frequently, and three secret bits can be hidden combined with NILAS and SMVQ. Therefore, our proposed method achieved a higher capacity than any of the other methods while simultaneously providing a greater embedding rate and payload than the other methods.

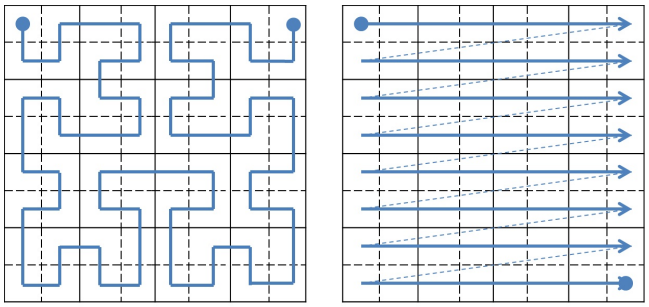


Figure 12: Run paths: (a) Improved path in Yang and Lin's method; (b) Normal VQ encoder path

5 Conclusions

In this paper, we proposed a reversible data hiding method for SMVQ indices. In the proposed method, using SMVQ indices can embed five bits or three bits at a time. In order to increase the embedding capacity, we combined SMVQ with improved locally adaptive coding. The experimental results showed that our proposed method effectively improved the capacity and outperformed state-of-the-art VQ-based data hiding schemes. In the future, we will focus on studying a simpler technique than locally adaptive coding to realize data hiding for compressed images while keeping high embedding capacity.

References

[1] A. A. Abdelwahab and L. A. Hassaan, "A discrete wavelet transform based technique for image data hiding," in *The 25th National Ratio Science Conference*, pp. 1–9, Jun. 2008.

[2] D. Baby, J. Thomas, G. Augustine, E. George, and N. R. Michael, "A novel DWT based image securing method using steganography," *Procedia Computer Science*, vol. 46, pp. 612–618, 2015.

[3] J. L. Bentley, D. D. Sleator, R. E. Tarjan, and V. K. Wei, "A locally adaptive data compression scheme," *Communications of the ACM*, vol. 29, no. 4, pp. 320–330, 1986.

[4] T. S. Nguyen C. C. Chang and C. C. Lin, "A reversible data hiding scheme for vq indices using locally adaptive coding," *Journal of Visual Communication and Image Representation*, vol. 22, no. 7, pp. 664–672, 2011.

[5] C. C. Chang and Y. C. Chou, "Reversible information hiding for vq indices based on locally adaptive coding," *Journal of Visual Communication and Image Representation*, vol. 20, no. 1, pp. 57–64, 2009.

[6] C. C. Chang, C. C. Lin, C. S. Tseng, and W. L. Tai, "Reversible hiding in DCT-based compressed images," *Information Sciences*, vol. 177, no. 13, pp. 2768–2786, 2007.

[7] C. C. Chang and T. S. Nguyen, "A reversible data hiding scheme for smvq indices," *Informatica*, vol. 25, no. 4, pp. 523–540, 2014.

[8] C. C. Chang and W. C. Wu, "A lossless data embedding technique by joint neighboring coding," *Pattern Recognition*, vol. 42, no. 7, pp. 1597–1603, 2009.

[9] I. C. Chang, Y. C. Hu, W. L. Chen, and C. C. Lo, "High capacity reversible data hiding scheme based on residual histogram shifting for block truncation coding," *Signal Processing*, vol. 108, pp. 376–388, 2015.

[10] R. M. Gray, *Vector Quantization*, Readings in Speech Recognition, 1990. (`https://en.wikipedia.org/wiki/Vector_quantization`)

[11] T. Kim, "Side match and overlap match vector quantizers for images," *IEEE Transactions on Image Processing*, vol. 1, no. 2, pp. 170–185, 1992.

[12] C. C. Lin and X. L. Liu, "A reversible data hiding scheme for block truncation compressions based on histogram modification," in *The 6th International Conference on Genetic and Evolutionary Computing*, pp. 157–160, Aug. 2012.

[13] C. C. Lin, X. L. Liu, and S. M. Yuan, "Reversible data hiding for vq-compressed images based on search-order coding and state-codebook mapping," *Information Sciences*, vol. 293, no. 1, pp. 314–326, 2015.

[14] Y. K. Lin, "High capacity reversible data hiding scheme based upon discrete cosine transformation," *Journal of Systems and Software*, vol. 85, no. 10, pp. 2395–2404, 2012.

[15] Y. K. Lin, "A data hiding scheme based upon dct coefficient modification," *Computer Standards & Interfaces*, vol. 36, no. 5, pp. 855–862, 2014.

[16] Y. Linde, A. Buzo, and R. Gray, "An algorithm for vector quantizer design," *IEEE Transactions on Communications*, vol. 28, no. 1, pp. 84–95, 1980.
[17] Z. Pan, X. Ma, X. Deng, and S. Hu, "Low bit-rate information hiding method based on search-order-coding technique," *The Journal of Systems and Software*, vol. 86, no. 11, pp. 2863–2869, 2013.
[18] C. Qin and Y. C. Hu, "Reversible data hiding in vq index table with lossless coding and adaptive switching mechanism," *Signal Processing*, vol. 129, pp. 48–55, 2016.
[19] T. Y. Tu and C. H. Wang, "Reversible data hiding with high payload based on referred frequency for VQ compressed codes index," *Signal Processing*, vol. 108, pp. 278–287, 2015.
[20] M. Vijay and V. VigneshKumar, "Image steganography algorithm based on huffman encoding and transform domain method," in *The 15th International Conference on Advanced Computing*, pp. 517–522, Oct. 2013.
[21] K. Wang, Y. Hu, and Z.M. Lu, "Reversible data hiding for block truncation coding compressed images based on prediction-error expansion," in *The 8th International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pp. 317–320, Aug. 2012.
[22] X. Wu and W. Sun, "Data hiding in block truncation coding," in *International Conference on Computational Intelligence and Security*, pp. 406–410, Dec. 2010.
[23] C. H. Yang and Y. C. Lin, "Reversible data hiding of a vq index table based on referred counts," *Journal of Visual Communication and Image Representation*, vol. 20, no. 6, pp. 399–407, 2009.
[24] C. H. Yang and Y. C. Lin, "Fractal curves to improve the reversible data embedding for vq-indexes based on locally adaptive coding," *Journal of Visual Communication and Image Representation*, vol. 21, no. 4, pp. 334–342, 2010.

Biography

Chin-Chen Chang received his Ph.D. degree in computer engineering from National Chiao Tung University. His current title is Chair Professor in Department of Information Engineering and Computer Science, Feng Chia University, from February 2005. He is currently a Fellow of IEEE and a Fellow of IEE, UK. And, since his early years of career development, he consecutively won Outstanding Talent in Information Sciences of the R. O. C., AceR Dragon Award of the Ten Most Outstanding Talents, Outstanding Scholar Award of the R. O. C., Outstanding Engineering Professor Award of the R. O. C., Distinguished Research Awards of National Science Council of the R. O. C., Top Fifteen Scholars in Systems and Software Engineering of the Journal of Systems and Software, and so on. On numerous occasions, he was invited to serve as Visiting Professor, Chair Professor, Honorary Professor, Honorary Director, Honorary Chairman, Distinguished Alumnus, Distinguished Researcher, Research Fellow by universities and research institutes. His current research interests include database design, computer cryptography, image compression, and data structures.

Jun-Yong Chen received his B.S. degree in 2015, in Department of Computer Science and Information Engineering from Chung Hua University, Hsinchu, Taiwan. He is currently studying in Feng Chia University. His current research interests include image processing and data hiding.

Yan-Hong Chen was born in September 1980 and received his master degree in computer applications from Northeast Dianli University, China in March 2005. He is currently a lecturer at School of Information, Zhejiang University of Finance and Economics. His current research interests include data hiding, image retrieval, evolutionary algorithms etc.

Yanjun Liu received her Ph.D. degree in 2010, in School of Computer Science and Technology from University of Science and Technology of China (USTC), Hefei, China. She has been an assistant professor serving in Anhui University in China since 2010. She currently serves as a senior research fellow in Feng Chia University in Taiwan. Her specialties include E-Business security and electronic imaging techniques.

Security Analysis of a Certificateless Public Provable Data Possession Scheme with Privacy Preserving for Cloud-Based Smart Grid Data Management System

Caixue Zhou

(Corresponding author: Caixue Zhou)

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

(Received Jan. 18, 2019; Revised and Accepted Aug. 8, 2019; First Online Sept. 21, 2019)

Abstract

The certificateless public key cryptosystem not only reduces the high cost of public key management, but also eliminates the private key escrow problem. The cloud-based smart grid data management system can release the burden of big data storage in power enterprises. Provable data possession (PDP) can ensure the integrity of data stored in the cloud with a high probability. Recently, a certificateless public PDP scheme with privacy preserving for cloud-based smart grid data management system was proposed. However, we find the scheme insecure. We give two concrete attacks to the scheme - the first attack shows that a malicious cloud storage provider (CSP) can forge a valid tag of any file block modified at his will, and the second one shows that CSP can produce a valid proof without storing any file blocks. Then, we point out the flaws in their proof and the key reason why their scheme is insecure.

Keywords: Certificateless Cryptosystem; Provable Data Possession; Smart Grid

1 Introduction

With the rapid development of computer network, communication technology and sensor technology, the smart grid [1, 6, 11] is gradually entering people's life as the next-generation power system. Build on the integrated and high-speed two-way communication work, it is aimed to achieve reliable, safe, economic, efficient and environmental-friendly operations. So far, many countries have launched smart grid projects [13, 16].

However, with the application of smart meters and other smart devices, the volume of electric power data is increasing exponentially. As a result, the traditional electric power information management system can no longer be able to process them in real time, prompting the birth of the cloud-based smart grid data management system [3]. It is flexible, scalable and reliable with a high equipment utilization rate and can help the smart grid achieve the storage of massive data. However, data stored in the cloud may be lost or damaged due to soft/hardware failures, human errors or hacker attacks. Thus, it has become an essential step to verify the integrity of data stored in the cloud.

Provable data possession (PDP) [5] can help check the integrity of cloud data without downloading it. It is a lightweight cloud data integrity probabilistic checking model. There are two kinds of auditing methods in PDP, i.e., public auditing [9] and private auditing [15]. In the former, Anyone with public information can audit the data, and therefore the cloud user can delegate the verification process to a third-party verifier (TPV) to ensure that his data is intact in the cloud; while in the latter, the auditor must use some private information to audit the data. At present, public auditing is becoming a popular trend. But in this method, the TPV should not deduce the cloud user's data when they check the integrity of it, and in this case, a public verifiable scheme with privacy preserving can be used [14]. However, all the above schemes are based on the public key infrastructure (PKI), which has the complex public key management problem. In order to reduce the high cost of public key management, the identity-based PDP [12] was proposed. However, the identity-based public key cryptosystem brings a new problem about key escrow - the trusted third party knows all users' private keys.

Regarding this problem, the certificateless public key cryptosystem [10] has great superiority. It not only reduces the high cost of public key management, but also solves the private key escrow problem. Many certifictateless PDP schemes have been proposed in the literature.

For example, in 2017, Kang *et al.* [7] proposed a certificateless public PDP scheme with privacy preserving, and applied it to the cloud-assisted wireless body area network. In the same year, He *et al.* [2] proposed another certificateless public PDP scheme with privacy preserving. Kim *et al.* [8] also proposed a certificateless public PDP scheme. In 2018, He *et al.* [4] proposed a certificateless public PDP scheme, and applied it to the cloud-assisted wireless body area network. However, their scheme does not support privacy preserving. In the same year, He *et al.* [3] proposed another certificateless public PDP scheme with privacy preserving, and applied it to the cloud-based smart grid data management system.

In this paper, we point out that scheme [3] is insecure and give two concrete attacks against it. The first attack shows that a malicious CSP can modify a file block and produce the corresponding tag. The second shows that a malicious CSP can produce a proof to pass the integrity verification without having to hold any data blocks.

The rest of this paper is organized as follows. Section 2 provides the formal definition and security model of certificateless provable data possession. Section 3 describes He *et al.*'s scheme. Section 4 gives two concrete attacks against their scheme, and then it points out the flaws in their proof and the key reason why their scheme is insecure. At last, the conclusion is given in Section 5.

2 Preliminaries

2.1 Formal Definition of Certificateless Provable Data Possession

There are four entities in the system: cloud users, who have huge data to be stored in the cloud; a cloud storage provider (CSP), which provides data storage service; a third party verifier (TPV), which is delegated by the cloud users to verify the cloud data integrity; and a key generation center (KGC), which produces system parameters and cloud users' partial private keys.

A certificateless provable data possession scheme consists of the following five algorithms:

1) Setup: Given a security parameter 1^k, KGC generates a master private key s and a common public parameter $Params$. For cloud user ID, KGC uses s and $Params$ to generate a partial private key PSK_{ID} and sends it to him secretly. Then, the cloud user ID randomly selects a secret value x_{ID}, and computes his public key UPK_{ID}. The cloud user ID's full private key consists of two parts: the partial private key PSK_{ID} and the secret value x_{ID}.

2) Store: Given a file $m_F = \{m_1, m_2, ..., m_n\}$, the cloud user uses his full private key to generate $\{m_i\}(i = 1, 2, ..., n)$'s tag $\{\sigma_i\}(i = 1, 2, ..., n)$. Then, he sends $\{m_i, \sigma_i\}(i = 1, 2, ..., n)$ to CSP, which checks whether $\{\sigma_i\}(i = 1, 2, ..., n)$ are valid. If they are invalid, CSP will ask the cloud user to re-produce them.

3) ChalGen: TPV randomly chooses a subset $I \in \{1, 2, ..., n\}$ and generates a challenge message to CSP.

4) ProGen: CSP produces a proof according to the challenge message, file $\{m_i\}(i = 1, 2, ..., n)$, tags $\{\sigma_i\}(i = 1, 2, ..., n)$ and sends it to TPV.

5) ProVer: TPV checks whether the proof is valid. If it is invalid, TPV will inform the cloud user that his file is corrupted.

Note 1. *Cloud users use the Store algorithm to produce file blocks' tags. With these tags and file blocks, CSP can produce a proof of data possession and cloud users can delete file blocks from their local copies. TPV uses the ChalGen algorithm to produce random file blocks to be audited. CSP uses the ProGen algorithm to produce a proof of data possession, which demonstrates that CSP stores users' files intactly. TPV uses the ProVer algorithm to check whether the proof is valid. If the proof is valid, it demonstrates that the data is intact in the cloud server.*

2.2 Security Model of Certificateless Provable Data Possession

Our security model is exactly the same as He *et al.*'s. There are two types of attackers in the certificateless public key cryptosystem [10]. The type-I attacker A_I can replace anyone's public key, but does not know the master private key. The type-II attacker A_{II} knows the master private key, but cannot replace anyone's public key. A certificateless PDP scheme must be unforgeable under both type-I and type-II adversaries.

Definition 1. *A CL-PDP scheme is unforgeable if no probabilistic polynomial time (PPT) adversary A (A_I or A_{II}) has a non-negligible advantage in the following game:*

Setup: Given a security parameter 1^k, challenger C produces the system's parameters $Params$ and a master private key s. If A is a type-I adversary, C gives the parameters $Params$ to A_I. If A is a type-II adversary, C gives the parameters $Params$ and master private key s to A_{II}.

Queries: A can adaptively make a polynomially bounded number of queries as follows:

1) Create-User Query: A supplies an identity ID. If ID's key pair has not been created, C produces ID's partial private key PSK_{ID} and secret value x_{ID}, and computes ID's public key UPK_{ID}. Then, C returns the public key UPK_{ID} to A.

2) Replace-Public-Key Query: A supplies an already created identity ID and a new public key UPK'_{ID}. C replaces the current public key UPK_{ID} with the new key UPK'_{ID}. If A is a type-II adversary, he cannot make such query.

3) Extract-Partial-Private-Key Query: A supplies an already created identity ID. C returns ID's partial private key PSK_{ID} to A. If A is a type-II adversary, he does not need to make such a query.
4) Extract-Secret-Value Query: A supplies an already created identity ID. C returns ID's secret value x_{ID} to A.
5) Tag-Gen Query: A supplies an already created identity ID and a file block m_i, and C computes the corresponding tag σ_i and returns it to A.

Forgery: At last, A outputs a forged tag σ^* corresponding the cloud user's identity ID^*. A wins the game if σ^* is valid and the following conditions hold.

1) If A is a type-I adversary, A_I cannot extract the partial private key of ID^*. If A is a type-II adversary, A_{II} cannot extract the secret value of ID^*.
2) σ^* is not the output of the Tag-Gen query.

3 He *et al.*'s Scheme

He *et al.*'s scheme consists of the following five algorithms.

Setup Algorithm

Step 1:

1) Given a security parameter k, KGC chooses two cyclic groups G_1 and G_2 of prime order q, a random generator P of G_1, a bilinear map $e : G_1 \times G_1 \to G_2$.
2) KGC randomly chooses $s \in Z_q^*$ as the system private key and computes the system public key $P_{pub} = sP$.
3) KGC chooses five secure hash functions $h_i : \{0,1\}^* \to Z_q^* (i = 1, 2, 3, 4)$ and $H : \{0,1\}^* \to G_1$.
4) KGC publishes the system parameters $\{G_1, G_2, e, P, q, P_{pub}, h_1, h_2, h_3, h_4, H\}$ and saves s secretly.

Step 2:

1) KGC randomly chooses $\overline{y}_{DO} \in Z_q^*$ and computes $\overline{Y}_{DO} = \overline{y}_{DO} \cdot P$.
2) KGC computes $\alpha_{DO} = h_1(ID_{DO}, \overline{Y}_{DO})$, and $y_{DO} = \alpha_{DO} \cdot \overline{y}_{DO} + s \mod q$.
3) KGC sends the partial private key y_{DO} to data owner (DO) secretly.

Step 3:

1) DO randomly chooses $x_{DO} \in Z_q^*$ as his secret value.
2) DO computes his public key $X_{DO} = x_{DO} \cdot P$.

Store Algorithm

Step 1:

1) DO randomly chooses $x_F \in Z_q^*$ and computes $X_F = x_F \cdot P$.
2) DO computes $\beta_{DO} = h_2(ID_{DO}, X_{DO}, \overline{Y}_{DO})$, $\alpha_F = h_3(ID_{DO}, X_{DO}, \overline{Y}_{DO}, X_F)$, and $s_F = \alpha_F \cdot x_F + \beta_{DO} \cdot x_{DO} + y_{DO} \mod q$.
3) DO saves x_F and X_F as a one-time signing key and verification key, respectively.

Step 2:

1) DO computes $V_{pub} = H(P_{pub})$ and $V_i = H(name_F, i)$ for $i = 1, ..., n$.
2) DO computes $\Phi_i = x_F \cdot (m_i \cdot V_{pub} + V_i)$ for $i = 1, ..., n$ and $\Phi_F = x_F \cdot V_{pub}$.
3) DO outputs Φ_i as m_i's tag for $i = 1, ..., n$.

Step 3: DO sends $\overline{F} = \{\{m_i\}_{i=1}^n, \{\Phi_i\}_{i=1}^n, s_F, X_F, \Phi_F\}$ to CSP.

Step 4: CSP checks if the equations $s_F \cdot P = \alpha_F \cdot X_F + \beta_{DO} \cdot X_{DO} + \alpha_{DO} \cdot \overline{Y}_{DO} + P_{pub}$ and $e(\sum_{i=1}^n \Phi_i, P) = e((\sum_{i=1}^n m_i) \cdot V_{pub} + \sum_{i=1}^n V_i, X_F)$ hold.

ChalGen Algorithm

Step 1: TPV randomly chooses a subset $I \in \{1, 2, ..., n\}$.

Step 2: TPV randomly chooses $w_i \in Z_q^*$ for each $i \in I$.

Step 3: TPV outputs $(\{i, w_i\}_{i \in I})$ as the challenge message and sends it to CSP.

ProGen Algorithm

Step 1: CSP randomly chooses $r_{CS} \in Z_q^*$ and computes $R_{CS} = r_{CS} \cdot \Phi_F$, $\Phi_{CS} = \sum_{i \in I} w_i \cdot \Phi_i$, $\alpha_{CS} = h_4(ID_{DO}, X_{DO}, \overline{Y}_{DO}, X_F, R_{CS}, \Phi_{CS})$ and $s_{CS} = \alpha_{CS} \cdot r_{CS} + \sum_{i \in I} w_i \cdot m_i \mod q$.

Step 2: CSP outputs the proof $\{X_F, \Phi_F, R_{CS}, \Phi_{CS}, s_{CS}\}$ and sends it to TPV.

ProVer Algorithm

TPV checks if the equations $s_F \cdot P = \alpha_F \cdot X_F + \beta_{DO} \cdot X_{DO} + \alpha_{DO} \cdot \overline{Y}_{DO} + P_{pub}$ and $e(\alpha_{CS} \cdot R_{CS} + \Phi_{CS}, P) = e(s_{CS} \cdot V_{pub} + \sum_{i \in I} w_i \cdot V_i, X_F)$ hold.

4 Security Analysis of He *et al.*'s Scheme

4.1 Two Concrete Attacks

Attack 1: Tag forging attack. According to Definition 1, in the Queries stage, a malicious CSP makes a Create-User query to ensure that the ID is created. Then he chooses a file block m_i and makes a Tag-Gen

query for (ID, m_i). After that, challenge C computes the corresponding tag $\overline{F} = \{m_i, \Phi_i, s_F, X_F, \Phi_F\}$ and returns it to him. Now, the malicious CSP can modify m_i to m_i^* and forge its corresponding tag Φ_i^* as follows. $\Phi_i^* = \Phi_i - m_i \cdot \Phi_F + m_i^* \cdot \Phi_F = x_F \cdot (m_i^* \cdot V_{pub} + V_i)$. The malicious CSP forges a valid tag Φ_i^* of m_i^* with a probability of 1.

Attack 2: Data loss hiding attack. Setup and Store algorithms are run as normal. After the Store algorithm, CSP gets file blocks and tags $\overline{F} = \{\{m_i\}_{i=1}^n, \{\Phi_i\}_{i=1}^n, s_F, X_F, \Phi_F\}$. Then, TPV runs the ChalGen algorithm to produce a challenge message $(\{i, w_i\}_{i \in I})$. After that, the malicious CSP computes $t_i = \Phi_i - m_i \cdot \Phi_F = x_F \cdot V_i$ for $i = 1, ..., n$. Then, he deletes all the file blocks $\{m_i\}_{i=1}^n$ and runs the ProGen algorithm as follows. He randomly chooses $r_{CS} \in Z_q^*$ and computes $R_{CS} = r_{CS} \cdot \Phi_F$, $\Phi_{CS} = \sum_{i \in I} w_i \cdot t_i$, $\alpha_{CS} = h_4(ID_{DO}, X_{DO}, \overline{Y}_{DO}, X_F, R_{CS}, \Phi_{CS})$, and $s_{CS} = \alpha_{CS} \cdot r_{CS} \mod q$. At last, the malicious CSP outputs the proof $\{X_F, \Phi_F, R_{CS}, \Phi_{CS}, s_{CS}\}$ and sends it to TPV. Obviously, the equation $e(\alpha_{CS} \cdot R_{CS} + \Phi_{CS}, P) = e(s_{CS} \cdot V_{pub} + \sum_{i \in I} w_i \cdot V_i, X_F)$ holds, meaning that the proof can pass the validation of the ProVer algorithm.

Note 2. *In the above attack 2, the malicious CSP can compute a valid proof without the cloud user's file, that is, the malicious CSP can delete all the cloud user's data file blocks.*

4.2 Flaws in the Proof of Lemma 1

In He *et al.*'s scheme, the proof of Lemma 1 is based on the security model which is defined in Subsection 2.2. In the Create-Data-Owner query, they divided it into two cases:

1) $ID_i = ID^*$. C stores $(ID^*, x_i, y_i, \perp, X_i, \overline{Y}_i)$ into L_K;

2) $ID_i \neq ID^*$. C stores $(ID_i, x_i, \perp, \overline{y}_i, X_i, \overline{Y}_i)$ into L_K.

Therefore, C knows the partial private key y_i in (1) and does not know the partial private key y_i in (2). Then, in the Extract-Partial-Private-Key query, C cannot give an answer when $ID_i \neq ID^*$. In other words, in most cases, C cannot answer this query. The same situation happens to Generate-Tag query - in most cases, C cannot answer this query, either.

In addition, in the proof part of C solving the CDH problem, the authors require $ID_i = ID^*$. In fact, our attack 1 shows that a malicious CSP can forge a valid tag when $ID_i \neq ID^*$. In other words, C can never solve the CDH problem.

Therefore, the simulation made by C is distinguishable from a true challenger, indicating that the proof is questionable.

4.3 Key Reason for the Insecurity

The key reason why He *et al.*'s scheme is insecure is that they compute a Φ_F in the Store algorithm. Then, the cloud user sends Φ_F to CSP along with the tags $\{\Phi_i\}_{i=1}^n$ and file blocks $\{m_i\}_{i=1}^n$. Because $\Phi_i = x_F \cdot (m_i \cdot V_{pub} + V_i) = m_i \cdot \Phi_F + x_F \cdot V_i$, CSP can modify file block m_i to m_i^* and compute $\Phi_i^* = \Phi_i - m_i \cdot \Phi_F + m_i^* \cdot \Phi_F = x_F \cdot V_i + m_i^* \cdot \Phi_F = x_F \cdot (m_i^* \cdot V_{pub} + V_i)$, that is, he can forge a valid tag in the above attack 1. At the same time, CSP can compute $t_i = \Phi_i - m_i \cdot \Phi_F = x_F \cdot V_i$ and produce a valid proof in the above attack 2. Obviously, if Φ_F is unknown to CSP, he cannot do the above attack computing.

Meanwhile, the ProGen algorithm must be run by CSP, but if Φ_F is unknown to CSP, it will not be able to run the algorithm. Therefore, He *et al.*'s scheme has a logic error.

5 Conclusions

In this paper, we give two attacks to a recently proposed certificateless public PDP scheme with privacy preserving for cloud-based smart grid data management system. In the first attack, a malicious CSP can forge a valid tag for any modified file block; and in the second one, a malicious CSP can produce a valid proof without storing any file blocks. We also point out the flaws in their proof and the key reason why their scheme is insecure.

Acknowledgments

This study was supported by the National Natural S cience Foundation of China [Grant no. 61462048] and the Natural Science Foundation of Jiangxi Province, China (No. 20181BAB202011). We thank to Ms. Yan Di, who checked our manuscript.

References

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

[2] D. B. He, N. Kumar, H. Q. Wang, L. N. Wang, and K. K. R. Choo, "Privacy-preserving certificateless provable data possession scheme for big data storage on cloud," *Applied Mathematic and Computation*, vol. 314, pp. 31–43, 2017.

[3] D. B. He, N. Kumar, S. Zeadally, and H. Q. Wang, "Certificateless provable data possession scheme for cloud-based smart grid data management systems," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 3, pp. 1232–1241, 2018.

[4] D. B. He, S. Zeadally, and L. B. Wu, "Certificateless public auditing scheme for cloud-assisted wireless

body area networks," *IEEE Systems Journal*, vol. 12, no. 1, pp. 64–73, 2018.

[5] W. F. Hsien, C. C. Yang, and M. S. Hwang, "A survey of public auditing for secure data storage in cloud computing," *International Journal of Network Security*, vol. 18, no. 1, pp. 133–142, 2016.

[6] M. Inam, Z. Li, A. Ali, and A. Zahoor, "A novel protocol for vehicle cluster formation and vehicle head selection in vehicular ad-hoc networks," *International Journal of Electronics and Information Engineering*, vol. 10, no. 2, pp. 103–119, 2019.

[7] B. Y. Kang, J. Q. Wang, and D. Y. Shao, "Certificateless public auditing with privacy preserving for cloud-assisted wireless body area networks," *Mobile Information Systems*, 2017. DOI: 10.1155/2017/2925465.

[8] D. M. Kim and I. R. Jeong, "Certificateless public auditing protocol with constant verification time," *Security and Communication Networks*, vol. 2017, no. 5, pp. 1-14, 2017.

[9] 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.

[10] 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.

[11] N. S. Nafi, K. Ahmed, M. A. Gregory, and M. Datta, "A survey of smart grid architectures, applications, benefits and standardization," *Journal of Network and Computer Applications*, vol. 76, pp. 23–36, 2016.

[12] S. Peng, F. C. Zhou, J. Li, Q. Wang, and Z. F. Xu, "Efficient, dynamic and identity-based remote data integrity checking for multiple replicas," *Journal of Network and Computer Applications*, vol. 134, pp. 72–88, 2019.

[13] 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.

[14] S. Thokchom and D. K. Saikia, "Privacy preserving and public auditable integrity checking on dynamic cloud data," *International Journal of Network Security*, vol. 21, no. 2, pp. 221–229, 2019.

[15] T. Y. Wu, Y. M. Tseng, S. S. Huang, and Y. C. Lai, "Non-repudiable provable data possession scheme with designated verifier in cloud storage systems," *IEEE Access*, vol. 5, pp. 19333–19341, 2017.

[16] Y. Zhang, W. Chen, and W. J. Gao, "A survey on the development status and challenges of smart grids in main driver countries," *Renewable and Sustainable Energy Reviews*, vol. 79, pp. 137–147, 2017.

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.

Multi-Parameter and Time Series Based Trust for IoT Smart Sensors

Zhi-Ge He

(Corresponding author: Zhi-ge He)

School of Computer Science & Engineering, University of Electronic Science & Technology of China
No. 2006, Xiyuan Ave, West Hi-Tech Zone, 611731, Chengdu, Sichuan, P. R. China
(Email: 578301541@qq.com)

(Received Jan. 18, 2019; Revised and Accepted June 19, 2019; First Online July 16, 2019)

Abstract

The Internet of Things, or IoT has achieved much attention in the past few years with many concrete applications. Among various IoT components, smart sensors play a vital role for things' tracking and monitoring, but due to the absence of centralized administration, those sensors may encounter various security issues which hinder IoT further development. Trust computing provides dynamic behavior perceiving capability and can take precautionary measures against malicious actions. In this study, unlike traditional binary parameter trust, we first propose a multi-parameter trust computing method so that trust states can be more accurately and practically described, then according to the theory of time series, a favorable trust data sequence and an unknown trust data sequence are generated so that nodes' malicious actions can be observed and detected from the context of a time period. Simulation results show that the proposed method can generate a fast detection of malicious nodes, a higher data packet delivery ratio, and a more trusted network environment ideal for transactions among sensor nodes.

Keywords: IoT; Multi-Parameter Trust; Smart Sensors; Time Series

1 Introduction

As one of the most emerging technologies in computer science, the Internet of Things, or IoT has achieved much popularity in the past few years and many IoT applications are being implemented in areas like logistics, traffic surveillance, and smart families. IoT can incorporate seamlessly and transparently a large number of heterogeneous smart devices or end systems, while providing open access to selected subsets of data for the development of a great many of digital services [21,24]. The term IoT is initially used to refer to the interoperability of uniquely identifiable objects with radio frequency identification (RFID) technology [16]. Later, the definition of IoT has been expanded to refer to a network of interconnected objects or devices such as RFID tags, sensors, actuators, and smart phones with the object to collect data and interact with the physical world [3, 22].

Among various IoT components, smart sensors play a vital role in the current IoT applications. Programmable smart sensors equipped with processing unit, storage memory, and wireless communication module are able to autonomously join in or construct a certain IoT network. Those sensors usually work in a completely distributed manner so as to collaboratively collect ambient data and monitor certain events. But due to the lack of fixed infrastructure, the absence of centralized administration, and the inherent characteristics of these sensors such as limited computing resources, short radio range, and dynamic topology, IoT composed by those sensors may encounter various security issues, *e.g.*, an entity may become malicious and launch packet dropping or select forwarding attack to gain its own benefits, which poses new security challenges for IoT applications [10, 15, 20].

As a complementary solution to the traditional network security, trust mechanism provides access control by judging the quality of the service and makes traditional security services more reliable by ensuring that all communicating devices are trustworthy during service cooperation [13, 14]. In this study, unlike traditional binary parameter trust methods [8], we first propose a multi-parameter trust computing method so that trust states can be more accurately and practically described, then according to the theory of time series, a favorable trust data sequence and an unknown trust data sequence are generated so that nodes' malicious actions can be observed and detected from the context of a time period.

The rest of this study is organized as follows. Section 2 explores recent representative trust-based models implemented for the IoT. Section 3 describes the preliminaries about trust computing and theory of time series. Sections 4 and 5 present our proposed trust method and related simulation tests. Section 6 concludes this article and suggests directions for future research.

2 Related Work

In this section, some latest and representative literatures about trust schemes in IoT are discussed, ranging from data aggregation/fusion, edge computing, malicious infiltration, information sharing, data routing, node classification, to trust estimation and assessment.

Data aggregating techniques using external IoT mobile elements (MEs) have been recently proposed in some studies where MEs collect data from stationary sensors and relay the collected data to the base station. These MEs could be regular mobile sensors or any mobile devices with sensing capability. Ali *et al.* [3] proposed a scheme on selecting trusted MEs for data aggregation in IoT enabled wireless sensor networks. When passing through the network, only trusted MEs were recruited, then they acted as anonymous agents and served as the cluster heads in order to increase the life span of the network. The trust vales placed on MEs were completely based on the direct interactions between the MEs and the base station at the end of each aggregation round. Regarding the trust calculation, [3] also uses the classic Beta trust model [8] which is of two trust parameter based and has been utilized by many reputation systems for its simplicity and flexibility. After that, all the trust values and management are handled by the base station and each sensor node maintains a local copy of the trust vales for other nodes in the network.

The integration of IoT and edge computing is currently a hot research direction, but the lack of trust among IoT edge devices has somewhat hindered the acceptance of IoT edge computing [19]. To facilitate the IoT edge computing applications, Yuan *et al.* [23] proposed a reliable and lightweight trust mechanism for IoT edge devices based on multi-source feedback information fusion so that efficient trust calculation mechanism can be established in the IoT edge computing architecture. The proposed scheme uses a feedback information fusion algorithm based on objective information entropy theory to overcome the limitations of traditional trust schemes, and the trust factors can be weighted manually or subjectively. In [23], the trust calculation falls into direct trust calculation and feedback trust calculation. The former uses the similar Beta trust model and the latter maintains the trust vales in a matrix.

Infiltration from malicious devices that can temporarily stop the provided services is one of the main issues faced by the current IoT networks and these malicious devices may also launch coordinated attacks. To find out the malicious behavior of IoT nodes, Khan [12] proposed an intrusion detection system based on the trust management where a node monitored the receivers of its messages checking if they had forwarded them correctly. Behavior following the scheme can improve the trust of a node in another one, but trust deteriorates if the observer detects that its peer behaves maliciously. [12] built the trust relation by using the opinion triangles in Jøsang's subjective logic [11] which allows to aggregate the trust values of various other IoT devices.

For information sharing in a health IoT system comprising IoT devices carried by members of an environmental health community, Al-Hamadi *et al.* [2] proposed a trust management system that could guide IoT devices to use the most trustworthy environmental health information for decision making. In [2], a collective knowledge base can be built to rate the environment at a particular location and time, and this knowledge could enable an IoT device to act on behalf of its user to decide whether or not the user should visit this place for health reasons. The proposed system considers the risk classification, reliability trust, and loss of health probability for decision making in the health IoT system.

With large amount of IoT devices likely to be interconnected globally, an important issue is how to secure the routing of data in the underlying networks from various attacks. Airehrour *et al.* [1] proposed a lightweight secure trust-based routing framework for IoT sensor nodes to identify and isolate common routing attacks in IoTs. The proposed framework incorporates the concept of trust among different IoT sensor nodes and utilizes the successful and unsuccessful node interactions among IoT nodes to evaluate a neighbor's trustworthiness. Further, the framework also considers a recovery period for nodes that are classified as untrusted ones owing to lossy network links or low battery power which could result in the decrease of their trust values.

Fragkiadakis *et al.* [7] proposed a centralized trust-based scheme employing evidence reasoning for IoT architecture where all nodes monitor their one-hop neighbors and report their findings to a single fusion center. The proposed scheme considers nodes' behavior with regards to their forwarding capability, thus each node observes its neighbors and estimates their packet drop ratio, then all nodes create direct trust reports for specific criteria regarding their neighbors, and the fusion is performed by employing a belief distribution using an evidence reasoning algorithm.

Asiri *et al.* [4] proposed an IoT trust and reputation model which used distributed probabilistic neural networks to classify trustworthy nodes from malicious ones. The proposed model is based on a recommender system which helps an IoT node decide to connect to another one based on previous observed behaviors. The proposed model also tackles the cold start problem in IoT environment by predicting ratings for newly joined nodes based on their characteristics over time, and the processing is completely distributed and is handled by the nodes themselves.

Gwak *et al.* [9] proposed an IoT trust estimation scheme making a user evaluate the trust value of an IoT device in an unknown place. The proposed scheme is on the basis that a user's subjective experience can be substituted by those social friends sharing identical subjective experiences with the user. It first finds a collection of past subjective experiences of a user relevant to the target device, then it discovers the friends of a target user who

have past subjective experiences closely matching with the collection. Based on the subjective trust value and the level of the subjective experience identity of the sharing friends, a user's trust value of a target device with the objective opinion of all the users who have interacted with the device is estimated.

To establish the initial trust level that a device places on another at their first encounter in IoTs, Nguyen *et al.* [17] proposed a challenge-response-based initial trust assessment scheme. The proposed scheme creates the knowledge about the device by learning the uncertainty level in its behaviors, then it relies on the results of the challenge-response process to assess if a device can be trusted to a level for its admission to the network. The proposed mechanism allows a device to generate the evidence for trust computation instead of waiting for the recommendations or actual interactions for long period.

3 Preliminaries

In this section, as the basis of our proposed trust method, we first discuss the traditional trust computing method and then shortly introduce the theory of time series.

3.1 Trust Computing

In trust computing, Bayesian analysis [5] has been widely used, and as a representative of such a approach, Ganeriwal *et al.* [8] proposed a classical reputation based framework for high integrity sensor networks (RFSN) where sensor nodes use Beta reputation to evaluate other's trust values.

Suppose that in a packet relay cooperation, a sensor node i has the probability φ to pass a packet to another node in the following jth round, let α and β denote the historical number of successful and unsuccessful cooperations respectively, φ is an unknown parameter and is equally to take all the values between 0 and 1 inclusive, then according to Bayesian analysis, $P(\varphi)$ is defined by

$$P(\varphi) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\varphi^{\alpha-1}(1-\varphi)^{\beta-1} \tag{1}$$

Let $K(=0/1)$ denote the outcome of the jth round, then $P(K|\varphi)$ is defined by

$$P(K|\varphi) = \varphi^K(1-\varphi)^{1-K} \tag{2}$$

Once the jth round is completed, according to Bayesian theorem, the posterior distribution of φ is defined by

$$P(\varphi|K) = \frac{P(K|\varphi)P(\varphi)}{\int P(K|\varphi)P(\varphi)d\varphi} \tag{3}$$

Put Equation 1 and Equation 2 into Equation 3, then we get

$$\begin{aligned} P(\varphi|K) &= \frac{\varphi^K(1-\varphi)^K\frac{\Gamma(\alpha+\beta)\cdot\varphi^{\alpha-1}(1-\varphi)^{\beta-1}}{\Gamma(\alpha)\Gamma(\beta)}}{\int \varphi^K(1-\varphi)^{1-K}\cdot\frac{\Gamma(\alpha+\beta)\cdot\varphi^{\alpha-1}(1-\varphi)^{\beta-1}}{\Gamma(\alpha)\Gamma(\beta)}d\varphi} \\ &= \frac{\Gamma(\alpha+\beta+1)\varphi^{\alpha+K-1}(1-\varphi)^{\beta+1-K-1}}{\Gamma(\alpha+K)\Gamma(\beta+1-K)} \end{aligned} \tag{4}$$

Equation 4 is the update of φ after the jth round. It can be noticed that in Equation 4, the posterior probability of φ still has a Beta distribution, i.e. before the jth round, $P(\varphi) \backsim Beta(\alpha,\beta)$ (Equation 1) ; after the jth round, $P(\varphi) \backsim Beta(\alpha+K, \beta+1-K)$. Therefore, before the jth round, $E(\varphi)$ is defined by

$$E(\varphi) = \frac{\alpha}{\alpha+\beta} \tag{5}$$

and after the jth round, $E(\varphi)$ is redefined by

$$E(\varphi) = \frac{\alpha+K}{\alpha+\beta+1} \tag{6}$$

In practice, $E(\varphi)$ is the trust value of node i and (α,β) are the only two trust parameters characteristic of Beta reputation that are computed and maintained by the neighboring nodes. This kind of trust computing is also called direct observation computing, and many literatures like [3,17] either directly or indirectly extend and modify such a method. In [18], an indirect method based on the belief discounting is used in the trust system, which is mapped into Dempster-Shafer belief theory [6] where the two trust parameters are defined as follows.

$$\alpha_i+ = \frac{2\alpha_h\alpha_i^h}{(\beta_h+2)+(\alpha_i^h+\beta_i^h+2)+2\alpha_h} \tag{7}$$

$$\beta_i+ = \frac{2\alpha_h\beta_i^h}{(\beta_h+2)+(\alpha_i^h+\beta_i^h+2)+2\alpha_h} \tag{8}$$

In Equation 7 and Equation 8, j receives the trust about i from h, let (α_i^h,β_i^h) denote the indirect trust and j has the past trust values about i and h denoted by (α_i,β_i) and (α_h,β_h) respectively. One of the advantages of the indirect method is that malicious nodes are prevented from colluding with each other to feed false trust information, but it can also result in the energy exhaustion of the network system.

To sum up, Beta reputation first computes the prior probability of an event, then updates the probability by using a posterior inference according to the relevant evidences, and (α,β) are trust parameters used to represent the positive and negative outcome in a transaction. Although Beta reputation model is widely used, it only considers two parameters to describe an event, which limits its applications to a large extent.

3.2 Time Series

Time series is a statistics tool for processing dynamic data sequence by which meaningful or abnormal facts can be analyzed and discovered. The data sequence is usually measured at successive time instants and spaced at predefined time intervals. For example, let Y denote a random variable and its time series is defined by $Y=\{y_1,...,y_n\}$ where y_n is the value of Y at time instant n.

In an IoT network, a node's trust is directly related to its *attitude* towards certain task, *e.g.*, faithfully relay data

packets as requested, or maliciously drop some or all the data packets. To some extent, malicious actions always end up with lower trust, but smart sensors can switch between *good* and *bad* so as to keep their trust and cover their malicious actions. Such a switch is easy to result in trust fluctuation over time. Thus, the fluctuated trust values can be regarded as a data sequence, and time series can be used to find out whether a node is of malicious actions or not.

In a trust data time series, there are three components: a trust data sequence to be checked, a standard sequence to be compared with, and a sequence checking mechanism. The trust data sequence is the outcomes of a certain node actions over time, *e.g.*, node i's trust data sequence is defined by

$$\mathcal{T}_i = \{t_i(t_1), t_i(t_2), ..., t_i(t_n)\}. \tag{9}$$

The standard sequence consists of a series of comparing data, each of which will be compared with its counterpart of the same time instant in the trust sequence. The standard sequence is denoted by

$$\mathcal{S} = \{s(t_1), s(t_2), ..., s(t_n)\}. \tag{10}$$

Both the trust data sequence and the standard sequence should have the same length, and the sequence checking mechanism used in this article will be introduced in the following section.

4 The Proposed Method

In this section, our proposed multi-parameter and time series base trust method is presented. The proposed method consists of two components: a trust computing module and a time series checking module.

4.1 Trust Computing Module

Assume there are k outcomes in one transaction denoted by $\{o_1, ..., o_k\}$ with the probability $\Theta = \{\theta_1, ..., \theta_k\}$ where $P(o_i) = \theta_i$, and n_i is the number of occurrence of o_i where $n_1 + n_2 + ... + n_k = N$, then according to the multinomial distribution, $P(Y = N|\Theta)$ is defined by

$$P(Y = N|\Theta) = \Sigma_{i=1}^{k-1} n_i \cdot (N-1)! \cdot \frac{\Pi_{i=1}^{k} \theta_i^{n_i}}{\Pi_{i=1}^{k} n_i!} \tag{11}$$

Based on the Dirichlet distribution, the conjugate prior probability of Θ is defined by

$$P(\Theta) = \frac{\Gamma(\Sigma_{i=1}^{k} \alpha_i)}{\Pi_{i=1}^{k} \Gamma(\alpha_i)} \cdot \Pi_{i=1}^{k} \theta_i^{\alpha_i - 1} \tag{12}$$

In Equation 12 as in Equation 1, α_i is the prior or historical counts of o_i, and the posterior of Θ is defined by

$$P(\Theta|Y = N) = \frac{P(Y = N|\Theta)P(\Theta)}{\int P(Y = N|\Theta) d\Theta} \tag{13}$$

Put Equation 11 and Equation 12 into Equation 13, we get

$$P(\Theta|Y = N) = \frac{\Sigma_{i=1}^{k-1} n_i \cdot (N-1)! \cdot \frac{\Pi_{i=1}^{k} \theta_i^{n_i}}{\Pi_{i=1}^{k} n_i!} P(\Theta)}{\int \Sigma_{i=1}^{k-1} n_i \cdot (N-1)! \cdot \frac{\Pi_{i=1}^{k} \theta_i^{n_i}}{\Pi_{i=1}^{k} n_i!} d\Theta} \tag{14}$$

$$= \frac{\Sigma_{i=1}^{k-1} n_i \cdot (N-1)! \cdot \frac{\Pi_{i=1}^{k} \theta_i^{n_i}}{\Pi_{i=1}^{k} n_i!} \cdot \frac{\Gamma(\Sigma_{i=1}^{k} \alpha_i)}{\Pi_{i=1}^{k} \Gamma(\alpha_i)} \cdot \Pi_{i=1}^{k} \theta_i^{\alpha_i - 1}}{\int \Sigma_{i=1}^{k-1} n_i \cdot (N-1)! \cdot \frac{\Pi_{i=1}^{k} \theta_i^{n_i}}{\Pi_{i=1}^{k} n_i!} d\Theta}$$

$$= \frac{\Gamma(\Sigma_{i=1}^{k} (\alpha_i + n_i))}{\Pi_{i=1}^{k} \Gamma(\alpha_i + n_i)}$$

Then, $E(\Theta)$ is defined by

$$E(\Theta) = (\frac{\alpha_1 + n_1}{\sum_{i=1}^{k} (\alpha_i + n_i)}, ..., \frac{\alpha_k + n_k}{\sum_{i=1}^{k} (\alpha_i + n_i)}). \tag{15}$$

In Equation 15 as in Equation 6, $E(\Theta)$ is the trust value set of node i and $(\alpha_1, ..., \alpha_n)$ are the multi trust parameters characteristic of the trust computing model in the proposed method where in $\alpha_i + n_i$, $n_i = 0, 1, ...N$.

Consider an example of three kinds of outcomes, assume that they are {*excellent, good, average*} denoted respectively by $\{o_1, o_2, o_3\}$ with the occurrence number $\{n_1, n_2, n_3\}$ after certain transactions, and the historical occurrence numbers are $\{\alpha_1, \alpha_2, \alpha_3\}$, based on Equation 15, the trust value set of {*excellent, good, average*} are computed as follows.

$$E(\theta_1) = \frac{\alpha_1 + n_1}{\alpha_1 + \alpha_2 + \alpha_3 + n_1 + n_2 + n_3} \tag{16}$$

$$E(\theta_2) = \frac{\alpha_2 + n_2}{\alpha_1 + \alpha_2 + \alpha_3 + n_1 + n_2 + n_3} \tag{17}$$

$$E(\theta_3) = \frac{\alpha_3 + n_3}{\alpha_1 + \alpha_2 + \alpha_3 + n_1 + n_2 + n_3} \tag{18}$$

Compared with the Beta reputation based trust, the multi parameter based trust has more trust parameters to present more outcome states in the actual applications.

4.2 Sequence Checking Module

In the checking module, the trust data sequence and the standard sequence are regarded as two vectors, the cosine angle $\lambda(v_1, v_2)$ of the two vectors are computed to measure their similarity, regarding the trust data sequence and the standard sequence, their cosine angle is defined by

$$\lambda(\mathcal{T}_i, \mathcal{S}) = \frac{\mathcal{T}_i \cdot \mathcal{S}}{\| \mathcal{T}_i \| \| \mathcal{S} \|} = \frac{\sum_{j=1}^{n} t_i(t_j) \times s(t_j)}{\sqrt{\sum_{j=1}^{n} (t_i(t_j))^2} \times \sqrt{\sum_{j=1}^{n} (s(t_j))^2}} \tag{19}$$

Further, λ is formalized as follows so that its value is mapped into $[0, 1]$.

$$\dot{\lambda}(\mathcal{T}_i, \mathcal{S}) = 1 - \frac{\cos^{-1}(\lambda(\mathcal{T}_i, \mathcal{S}))}{\pi} \tag{20}$$

In Equation 20, the closer $\dot{\lambda}$ gets to 0, the less similar the trust data sequence and the standard sequence become, which means that the trust data sequence deviates from the standard sequence to a large extent, and the sensor node is highly likely of malicious actions within the time period.

In practice, considering the unknown events such as packet loss during the transmission and to fully take the advantage of the multi trust parameters, we use four data sequences: favorable trust data sequence fT_i (like successful data relay) and its counterpart comparing standard sequence S_{fT}, unknown trust data sequence uT_i and its comparing standard sequence S_{uT}.

The working algorithm of the proposed trust method is shown in Algorithm 1.

Algorithm 1 Working of the proposed method

α_1: historical favorable outcome number
α_2: historical unfavorable outcome number
α_3: historical unknown outcome number
n_1: current favorable outcome number
n_2: current unfavorable outcome number
n_3: current unknown outcome number
Len: segment length of time series measured by the number of transactions ($Len \geq n_1 + n_2 + n_3$)
$\varphi_1 \in [0, 1]$: threshold of $\dot{\lambda}(fT_i, S_{fT})$
$\varphi_2 \in [0, 1]$: threshold of $\dot{\lambda}(uT_i, S_{uT})$
Begin
Node j initiates a certain transaction such as packets relay within its one hop neighbors, assume Node i responds, j first checks i's (favorable) trust value, if i is qualified then j starts the transaction with i and observes i's transaction outcomes
for(count=0, count <= *Len*, count++)
  {j observes the transaction outcomes and records them in (n_1,n_2,n_3)}
if ($\dot{\lambda}(fT_i, S_{fT}) > \varphi_1$)
  {$\alpha_1 += n_1, \alpha_2 += n_2, \alpha_3 += n_3$, compute $E(\theta_1)$}
else
{
  if ($\dot{\lambda}(uT_i, S_{uT}) > \varphi_2$)// i is of malicious actions.
    {delete i from j's transaction partners' list, or reset i's trust value for its redemption}
  else// too many unknown outcomes exist.
    {$\alpha_1 = \alpha_1, \alpha_2 = \alpha_2, \alpha_3 += n_3$, compute $E(\theta_1)$}
}
End

5 Simulations

Suppose that in an IoT packet relay task, there exist three kinds of smart sensor nodes, i.e. legitimate nodes (65%), malicious nodes (25%), and selfish nodes (10%). A transaction is defined as a data packet relay. Legitimate nodes are of good actions and they faithfully relay all the received packets to the others as requested; to attack the integrity of network is the first priority of the malicious nodes, they intelligently and selectively drop some or all the received packets and try to keep their trust values to an acceptable level so as to cover their malicious actions; selfish nodes sometimes drop packets or deny request not out of malicious actions but to gain its own benefit such as saving their energy. Each node generates 1 data packet containing its ID on every 10 seconds, and a base station locates on the border of the test area to collect all the packets from the network. It is also assumed that sensor nodes are capable of bidirectional communication and their NICs work in a promiscuous mode. NS–2 is used for simulation and the classical binary reputation based trust used in [3] is selected for comparing. Simulation parameters are presented in Table 1.

Table 1: Simulation parameters

Parameters	Values
Simulation time	500s
Number of nodes	100
Test area	$200\times200m^2$
Transmission range	50m
Node placement	random
MAC protocol	IEEE 802.11
Packet size	100 bytes
Communication error	5%
S_{fT}	randomly $\subset [0.75, 0.95]$
uT_i	randomly $\subset [0, 0.15]$
Initial trust value	0.5

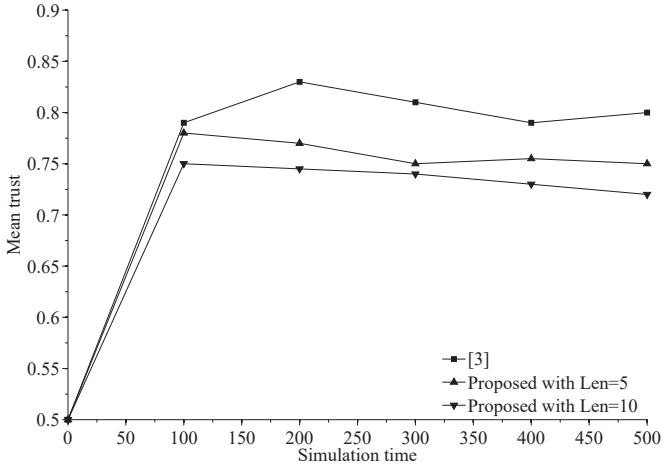


Figure 1: Mean trust with $\varphi_1 = 0.8, \varphi_2 = 0.8$

5.1 Test 1

In this section, the mean trust value is tested between the compared method and the proposed method, results are shown in Figure 1 and Figure 2.

In Figure 1, as the simulation time goes by, the mean

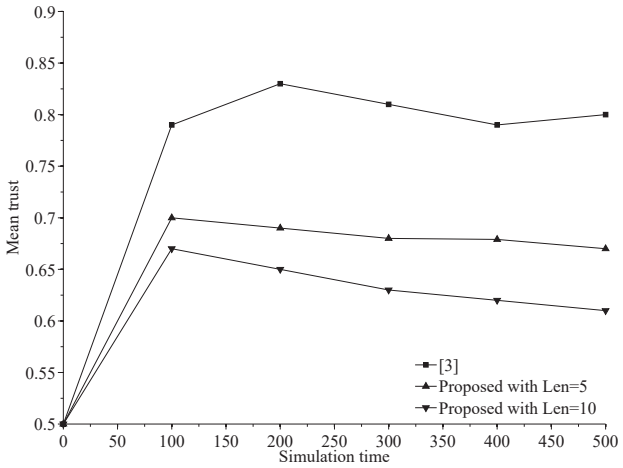


Figure 2: Mean trust with $\varphi_1 = 0.9, \varphi_2 = 0.9$

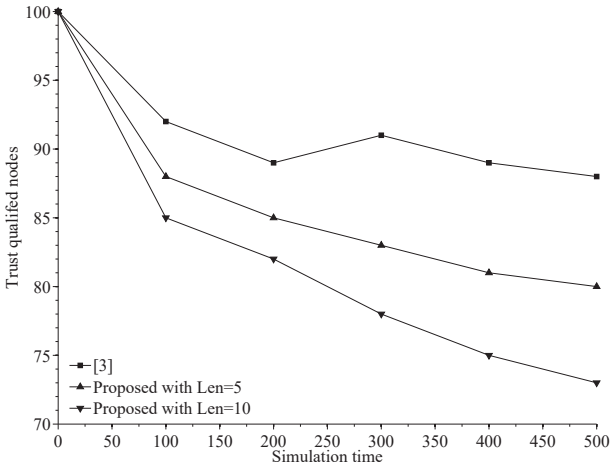


Figure 3: Trust qualified nodes with $\varphi_1 = 0.8, \varphi_2 = 0.8$

trust value in [3] begins to go upward and reaches about 0.83 on the 200th second, then drops and fluctuates around 0.8 till to the 500th second. Although there exist 25% malicious nodes and 10% selfish nodes, the mean trust value still keeps high as 0.8. This is because in [3], malicious nodes can intelligently switch between passing and dropping the packets, which helps them maintain an acceptable trust value that can be considered as trusted relaying nodes.

While in the proposed method, trust is computed based on *Len*–segment length of time series which is measured by the number of transactions. It means that trust in the proposed method is computed according to the segment length instead of upon the completion of a transaction. This helps to keep broader perspective on the target nodes. When malicious nodes intelligently switch between *good* and *bad*, its trust fluctuate accordingly, when its trust data sequence and standard data sequence are applied into the checking module, if the result is less than the threshold φ_1 and meanwhile there is no enough unknown outcomes, then according to the algorithm presented above, such nodes are treated as malicious ones. In Figure 1, it can be noticed that due to the successfully spotting the malicious nodes, the mean trust value in the proposed method goes downward gradually, *e.g.*, around 0.67 ($Len = 5$) and 0.63($Len = 10$) on the 500th second. In Figure 1, the mean trust is the lowest in the propose method when the $Len = 10$, this is because when *Len* becomes larger, more malicious actions can be observed, if any, and malicious nodes are more difficult to cover their actions.

Similar results can be found in Figure 2. The difference is that in Figure 2, the thresholds φ_1 and φ_2 are set as 0.9 instead of 0.8 in Figure 1. Figure 2 shows that when these two thresholds are set larger, the checking module is becoming stricter, meaning that more malicious nodes can be detected resulting in much lower mean trust of the the network, *e.g.*, around 0.6 when ($Len = 5$) on the 500th second in Figure 2.

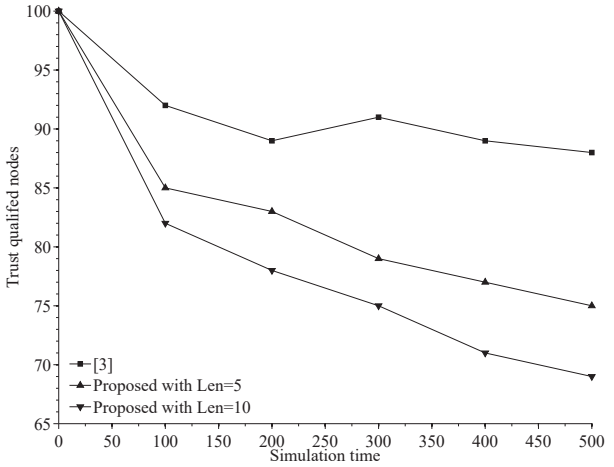


Figure 4: Trust qualified nodes with $\varphi_1 = 0.9, \varphi_2 = 0.9$

5.2 Test 2

In this section, the number of trust qualified nodes is tested between the compared method and the proposed method, results are shown in Figure 3 and Figure 4.

As is shown in Figure 3 and Figure 4, as the simulation continues, the number of trust qualified nodes drops in both the compared methods, *e.g.*, in Figure 3, the number in [3] is around 90 on the 200th second and around 88 on the 500th second. Such a number varies slightly from the 200th second to the 500th second, and the reason is that the switching actions of malicious nodes make them difficult to be spotted by the method in [3]. In addition, the number of trust qualified nodes of [3] in Figure 3 or Figure 4 is not the actual number which consists of many malicious nodes.

On contrast, the number of trust qualified nodes drops

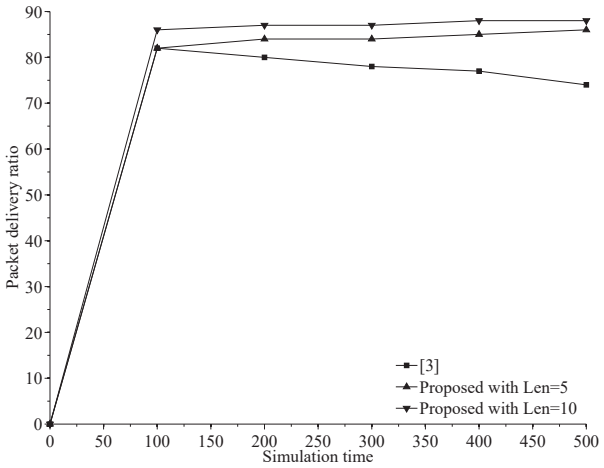


Figure 5: Packet delivery ratio with $\varphi_1 = 0.8, \varphi_2 = 0.8$

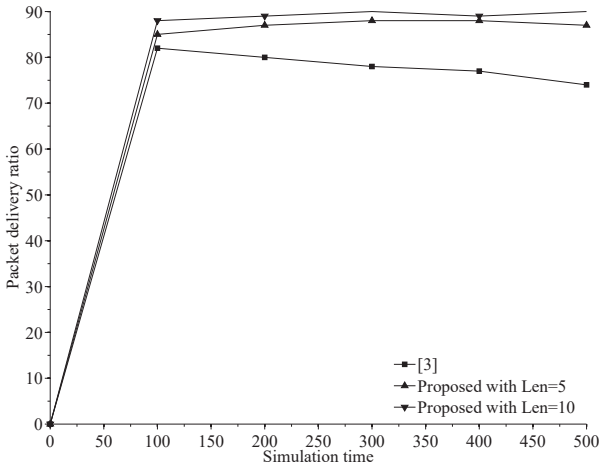


Figure 6: Packet delivery ratio with $\varphi_1 = 0.9, \varphi_2 = 0.9$

faster in the proposed method, *e.g.*, in Figure 3, around 86 on the 200th second and around 81 on the 500th second when $Len = 5$. In Figure 3 and Figure 4, it can be found that both the segment length Len and the two thresholds φ_1, φ_2 influence the number of trust qualified nodes. Under the same conditions, the larger the segment length and the two thresholds get, the less number of the trust qualified nodes becomes. For example, in Figure 4, when $Len = 10$, $\varphi_1 = 0.9$, and $\varphi_2 = 0.9$, the number in the proposed method is around 69. However, such a number in the proposed method approximates the actual number of legitimate nodes (65%), meaning that more and more malicious nodes including some selfish nodes are detected in the proposed method and most of the remaining nodes are legitimate.

5.3 Test 3

In this section, the packet delivery ratio is tested between the compared methods and results are shown in Figure 5 and Figure 6.

Due to the existence of malicious nodes and selfish nodes, not all the data packets generated by the legitimate nodes can be received by the base station. Take Figure 6 as an example, on the 500th seconds, only about 75% data packets reach the base station, and most of the rest 25% are dropped by the malicious nodes; while in the proposed method, because of the timely detection of malicious nodes, the packet delivery ration can reach as high as 85% ($Len = 5$) or 90% ($Len = 10$). Figure 5 and Figure 6 further indicate that with the increase of segment length and the two thresholds φ_1, φ_2, so does the packet delivery ratio.

These three tests also indicate that compared with the method in [3], although the proposed method generates a lower mean trust value and less trust qualified nodes in the network, it does result in a fast detection of malicious nodes, a higher data packet delivery ratio, and a more trusted network environment ideal for transactions among sensor nodes.

6 Conclusions

Due to the lightweight but powerful mechanism, trust scheme is a promising technology to establish security for the resource-constrained devices that are characteristic of the IoT smart sensors. In this study, we propose a multi-parameter trust computing method combined with the theory of time series. Through simulation tests, the feasibility and effectiveness of the proposed method have been confirmed. But in the proposed method, the segment length, the two thresholds cannot be selected freely, a longer segment length would exhaust the buffer of a sensor node; a larger threshold would not tolerate any mistakes such as a single packet dropping in the test case; a smaller threshold would be availed by malicious nodes to switch their actions, all of which would be our future research.

Acknowledgments

This study was supported by the NSF of China under grant No. 61772115 and Sichuan Miaozi Project under grant No.2019009. The author gratefully acknowledges the anonymous reviewers for their valuable comments.

References

[1] D. Airehrour and *et al.*, "A lightweight trust design for IoT routing," in *The 14th IEEE International Conference on Pervasive Intelligent and Computing*, pp. 552-557, 2016.

[2] H. Al-Hamadi and I. R. Chen, "Trust-based decision making for health IoT systems," *IEEE Internet of Things Journal*, vol. 4, no. 5, pp. 1408-1419, 2017.

[3] B. A. Ali, H. M. Abdulsalam and A. Alghemlas, "Trust based scheme for IoT enabled wireless sensor networks," *Wireless Personal Communications*, vol. 99, no. 4, pp. 1061-1080, 2018.

[4] S. Asiri and A. Miri, "An IoT trust and reputation model based on recommender systems," in *Privacy, Security & Trust*, pp. 561-568, 2017.

[5] S. Che and *et al.*, "A lightweight trust management based on Bayesian and entropy for wireless sensor networks," *Security and Communication Networks*, vol. 8, no. 2, pp. 168-175, 2015.

[6] A. Dempster, "Upper and lower probabilities induced by multivalued mapping," *The Annals of Mathematical Statistics*, vol. 38, no. 2, pp. 325-339, 1967.

[7] A. Fragkiadakis and E. Tragos, "A trust-based scheme employing evidence reasoning for IoT architectures," in *IEEE World Forum on Internet of Things*, pp. 559-564, 2017.

[8] S. Ganeriwal and *et al.*, "Reputation based framework for high integrity sensor networks," *ACM Transactions on Sensor Networks*, vol. 4, no. 3, pp. 15-37, 2008.

[9] B. Gwak and *et al.*, "IoT trust estimation in an unknown place using the opinions of i-sharing friends," in *IEEE Trustcom*, pp. 602-609, 2017.

[10] R. Jhaveri and *et al.*, "Sensitivity analysis of an attack-pattern discovery based trusted routing scheme for mobile Ad-Hoc networks in industrial IoT," *IEEE Access*, vol. 6, pp. 20085-20103, 2018.

[11] A. Jøsang, "A logic for uncertain probabilities," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 9, pp. 279-311, 2001.

[12] Z. A. Khan, "Using energy-efficient trust management to protect IoT networks for smart cities," *Sustainable Cities and Society*, vol. 40, pp. 1-15, 2018.

[13] C. T. Li, M. S. Hwang and Y. P. Chu, "An efficient sensor-to-sensor authenticated path-key establishment scheme for secure communications in wireless sensor networks", *International Journal of Innovative Computing, Information and Control*, vol. 5, no. 8, pp. 2107-2124, Aug. 2009.

[14] X. Li and *et al.*, "T-broker: A trust-aware service brokering scheme for multiple cloud collaborative services," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 7, pp. 1402-1415, 2015.

[15] L. Liu, Z. Cao, O. Markowitch, "A note on design flaws in one aggregated-proof based hierarchical authentication scheme for the internet of things," *International Journal of Electronics and Information Engineering*, vol. 5, no. 2, pp. 88–92, 2016.

[16] A. Mayzaud, R. Badonnel and I. Chrisment, "A taxonomy of attacks in rpl-based internet of things," *International Journal of Network Security*, vol. 18, no. 3, pp. 459-473, 2016.

[17] T. Nguyen and *et al.*, "Initial trust establishment for personal space IoT systems," in *IEEE Conference on Computer Communications Workshops*, pp. 784-789, 2017.

[18] S. Ozdemir, "Functional reputation based reliable data aggregation and transmission for wireless sensor networks," *Computer Communications*, vol. 31, no. 17, pp. 3941-3953, 2008.

[19] S. Pinto and *et al.*, "IIoTEED: An enhanced, trusted execution environment for industrial IoT edge devices," *IEEE Internet Computing*, vol. 21, no. 1, pp. 40-47, 2017.

[20] J. Shen and *et al.*, "Cloud aided lightweight certificateless authentication protocol with anonymity for wireless body area networks," *Journal of Network and Computer Applications*, vol. 105, pp. 117-123, 2018.

[21] 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.

[22] C. H. Wei, M. S. Hwang, A. Y. H. Chin, "A mutual authentication protocol for RFID", *IEEE IT Professional*, vol. 13, no. 2, pp. 20–24, Mar. 2011.

[23] J. Yuan and X. Li, "A reliable and lightweight trust computing mechanism for IoT edge devices based on multi-source feedback information fusion," *IEEE Access*, vol. 6, pp. 23626-23638, 2018.

[24] A. Zanella and *et al.*, "Internet of things for smart cities," *Internet of things for smart cities*, vol. 1, no. 1, pp. 22-32, 2014.

Biography

Zhi-ge He. He received his bachelor degree in computer science and application from University of Electronic Science and Technology of China in 2017 and now he is a master candidate in computer science and application. His main research interest includes network security and big data.

Security Analysis of Two Unbalancing Pairing-free Identity-based Authenticated Key Exchange Protocols

Qingfeng Cheng[1,2], Yuting Li[1,2], Qi Jiang[3], and Xiong Li[4]
(Corresponding author: Qingfeng Cheng)

Strategic Support Force Information Engineering University[1]
Zhengzhou 450001, P. R. China
State Key Laboratory of Mathematical Engineering and Advanced Computing[2]
Zhengzhou 450001, P. R. China
School of Cyber Engineering, Xidian University[3]
Xi'an, 710071, P. R. China
School of Computer Science and Engineering, Hunan University of Science and Technology[4]
Xiangtan, 411201, P.R. China
(Email: qingfengc2008@sina.com)
(Received Feb. 13, 2019; Revised and Accepted Sept. 16, 2019; First Online Feb. 9, 2020)

Abstract

The Internet of Things plays an increasingly important role in various fields. However, there are many devices in the Internet of Things that are unbalanced in terms of computing and storage capacity, which should be given full consideration. Recently, Zhang *et al.* proposed two unbalancing pairing-free identity-based authenticated key exchange (AKE) protocols for disaster scenarios, which was claimed to achieve forward security and impersonation attack resilience. In this paper, we show that two proposed AKE protocols are lack of forward security and also cannot resist key compromise impersonation attack.

Keywords: Authenticated Key Exchange; Forward Security; Key Compromise Impersonation Attack; Pair-Free

1 Introduction

The Internet of Things has been developing rapidly in recent years, bringing a lot of convenience services to people in various fields of the society. In the environment of Internet of Things, there are a substantial number of sensors, radio frequency cards and other devices with different computing and storage capabilities. In order to ensure secure communications among these devices, we usually use authenticated key exchange (AKE) protocols [2,6–13] to generate the session keys for encrypting messages over public network.

Although there are many AKE protocols for Internet of Things, they are seldom designed for disaster scenarios. In such scenarios, secure data transmissions among unbalanced devices are very important. Recently, Zhang *et al.* [14] proposed two pairing-free identity-based AKE protocols, called UPIAP1 protocol and UPIAP2 protocol. Both of them were designed for the limited devices with unbalanced computational ability. Zhang *et al.* proved their two UPIAP protocols' security in the mBR model [2] and compared the performance with pairing-free AKE protocols in [3, 5, 12]. However, in this paper, we will analyze the security of the UPIAP1 protocol and UPIAP2 protocol, and show that both of them still exist some security flaws. In details, if the adversary can learn two parties' long-term private keys, he can recover the previous session keys. In addition, if the adversary can learn a party's secret key or partial secret key, he can impersonate the other party to cheat the party, who divulges his own long-term private key.

The remainder of this paper will firstly introduce some notations and desirable security attributes in Section 2. Then we briefly review UPIAP1 protocol and UPIAP2 protocol in Section 3. Further, Section 4 points out the weaknesses of UPIAP1 protocol and UPIAP2 protocol. Conclusion will be given in Section 5.

2 Preliminaries

This section briefly introduces some notations and security attributes in Table 1, which are used in the UPIAP1 protocol and UPIAP2 protocol. More details can refer to [14].

In general, the basic desirable attributes of secure AKE protocols include key compromise impersonation (KCI)

Table 1: Notations

Notations	Description
τ	security parameter
Z_p^*	$\{1, 2, \cdots, p-1\}$
$\mathcal{G}$	a cyclic additive group of order p, P is a generator of this group
$\mathcal{M}$	the adversary
s	Key Generation Center (KGC)'s master private key
P_{pub}	KGC's master public key, where $P_{pub} = sP$
$\hat{X}$	party who involves in the AKE protocol
$(s_{\hat{X}}, v_{\hat{X}})$	party $\hat{X}$'s long-term private key, where $s_{\hat{X}}P = R_{\hat{X}} + H_1(\hat{X} \parallel R_{\hat{X}}) \cdot P_{pub}$ and $v_{\hat{X}} \in Z_p^*$
$(R_{\hat{X}}, V_{\hat{X}})$	party $\hat{X}$'s long-term public key, where $R_{\hat{X}} = r_{\hat{X}} \cdot P$, $r_{\hat{X}} \in Z_p^*$ and $V_{\hat{X}} = v_{\hat{X}} \cdot P$
H_1	a hash function from $\{0,1\}^*$ to Z_p^*
H_2	a hash function from $\{0,1\}^*$ to $\{0,1\}^\tau$
$HMAC$	a verification hash function from $\{0,1\}^*$ to $\{0,1\}^\tau$

security, key control security and forward security [4], *etc.* In this section, we only describe some security attributes involved in the analysis of two UPIAP protocols.

- **Forward security**. If two parties' long-term private keys are compromised simultaneously, the adversary cannot recover previous session keys.
- **KCI security**. Suppose $\hat{A}$'s private key $(s_{\hat{A}}, v_{\hat{A}})$ is compromised. The adversary cannot impersonate $\hat{B}$ to cheat $\hat{A}$.
- **Partial KCI security**. Suppose $\hat{A}$'s partial key $v_{\hat{A}}$ is compromised. The adversary cannot impersonate $\hat{B}$ to cheat $\hat{A}$.

3 Review of Two UPIAP Protocols

This section describes Zhang *et al.*'s two UPIAP protocols for disaster scenarios, which are claimed to achieve forward security and impersonation attack resilience. Two UPIAP protocols include a KGC and two parties respectively, where the KGC initializes the key exchange system parameters. For the sake of brevity, we omit some unnecessary descriptions.

3.1 UPIAP1 Protocol

In this subsection, we briefly review Zhang *et al.*'s UPIAP1 protocol.

Step 1. The initiator $\hat{A}$ randomly generates a value $\hat{a} \in Z_p^*$. Then $\hat{A}$ sends the message $\hat{I}_1$ to the responder $\hat{B}$ as follows:

$$\hat{A} \rightarrow \hat{B} : \hat{I}_1 = \{R_{\hat{A}}, V_{\hat{A}}, Eph_{\hat{A}}\},$$

where $Eph_{\hat{A}} = \hat{a} + v_{\hat{A}}$.

Step 2. After receiving the message $\hat{I}_1$, $\hat{B}$ randomly generates a value $\hat{b} \in Z_p^*$ and computes $Eph_{\hat{B}} = \hat{b} + v_{\hat{B}}$. Then $\hat{B}$ computes the session key components as follows:

$$K_{\hat{B}1} = s_{\hat{B}} \cdot (T_{\hat{A}} - V_{\hat{A}}) + \hat{b}(R_{\hat{A}} + H_1(\hat{A} \parallel R_{\hat{A}}) \cdot P_{pub}),$$

$$K_{\hat{B}2} = \hat{b} \cdot (T_{\hat{A}} - V_{\hat{A}}),$$

where $T_{\hat{A}} = Eph_{\hat{A}} \cdot P$ and $T_{\hat{B}} = Eph_{\hat{B}} \cdot P$.

Finally, $\hat{B}$ sends the message $\hat{R}$ to $\hat{A}$ as follows:

$$\hat{B} \rightarrow \hat{A} : \hat{R} = \{R_{\hat{B}}, V_{\hat{B}}, T_{\hat{B}}, T_{\hat{A}}, MAC_{\hat{B}}\},$$

where $MAC_{\hat{B}} = HMAC(K_{\hat{B}1} \parallel K_{\hat{B}2}, R_{\hat{B}} \parallel V_{\hat{B}} \parallel T_{\hat{B}} \parallel T_{\hat{A}})$.

Step 3. After receiving the message $\hat{R}$, $\hat{A}$ generates the session key components as follows:

$$K_{\hat{A}1} = s_{\hat{A}} \cdot (T_{\hat{B}} - V_{\hat{B}}) + \hat{a}(R_{\hat{B}} + H_1(\hat{B} \parallel R_{\hat{B}}) \cdot P_{pub}),$$

$$K_{\hat{A}2} = \hat{a} \cdot (T_{\hat{B}} - V_{\hat{B}}).$$

Then $\hat{A}$ checks $MAC_{\hat{B}}$. If $VER(K_{\hat{A}1} \parallel K_{\hat{A}2}, R_{\hat{B}} \parallel V_{\hat{B}} \parallel Eph_{\hat{B}}, MAC_{\hat{B}})$ equals to 1, it is valid. $\hat{A}$ generates the session key $SK_{\hat{A}\hat{B}} = H_2(\hat{A} \parallel \hat{B} \parallel T_{\hat{A}} \parallel T_{\hat{B}} \parallel K_{\hat{A}1} \parallel K_{\hat{A}2})$ and sends the message $\hat{I}_2$ to $\hat{B}$:

$$\hat{A} \rightarrow \hat{B} : \hat{I}_2 = \{MAC_{\hat{A}}\},$$

where $MAC_{\hat{A}} = HMAC(K_{\hat{A}1} \parallel K_{\hat{A}2}, R_{\hat{A}} \parallel V_{\hat{A}} \parallel Eph_{\hat{A}})$.

Step 4. After receiving the message $\hat{I}_2$, $\hat{B}$ checks $MAC_{\hat{A}}$. If $VER(K_{\hat{B}1} \parallel K_{\hat{B}2}, R_{\hat{A}} \parallel V_{\hat{A}} \parallel Eph_{\hat{A}}, MAC_{\hat{A}})$ equals to 1, it is valid. $\hat{B}$ generates the session key $SK_{\hat{B}\hat{A}} = H_2(\hat{A} \parallel \hat{B} \parallel T_{\hat{A}} \parallel T_{\hat{B}} \parallel K_{\hat{B}1} \parallel K_{\hat{B}2})$.

If $VER(K_{\hat{B}1} \parallel K_{\hat{B}2}, R_{\hat{A}} \parallel V_{\hat{A}} \parallel Eph_{\hat{A}}, MAC_{\hat{A}})$ equals to 0, it is invalid. $\hat{B}$ aborts the session.

3.2 UPIAP2 Protocol

In this subsection, we briefly review Zhang *et al.*'s UPIAP2 protocol.

Step 1. The initiator $\hat{A}$ randomly generates a value $\hat{a} \in Z_p^*$. Then $\hat{A}$ sends the message $\hat{I}_1$ to the responder $\hat{B}$ as follows:

$$\hat{A} \rightarrow \hat{B} : \hat{I}_1 = \{R_{\hat{A}}, V_{\hat{A}}, T_{\hat{A}}\},$$

where $Eph_{\hat{A}} = \hat{a} + v_{\hat{A}}$, $T_{\hat{A}} = Eph_{\hat{A}} \cdot P$.

Step 2. After receiving the message $\hat{I}_1$, $\hat{B}$ randomly generates a value $\hat{b} \in Z_p^*$ and computes $Eph_{\hat{B}} = \hat{b} + v_{\hat{B}}$. Then $\hat{B}$ computes the session key components as follows:

$$K_{\hat{B}1} = s_{\hat{B}} \cdot (T_{\hat{A}} - V_{\hat{A}}) + \hat{b}(R_{\hat{A}} + H_1(\hat{A} \parallel R_{\hat{A}}) \cdot P_{pub}),$$

$$K_{\hat{B}2} = \hat{b} \cdot (T_{\hat{A}} - V_{\hat{A}}).$$

Finally, $\hat{B}$ sends the message $\hat{R}$ to $\hat{A}$ as follows:

$$\hat{B} \rightarrow \hat{A} : \hat{R} = \{R_{\hat{B}}, V_{\hat{B}}, Eph_{\hat{B}}, MAC_{\hat{B}}\},$$

where $MAC_{\hat{B}} = HMAC(K_{\hat{B}1} \parallel K_{\hat{B}2}, R_{\hat{B}} \parallel V_{\hat{B}} \parallel Eph_{\hat{B}})$.

Step 3. After receiving the message $\hat{R}$, $\hat{A}$ generates the session key components as follows:

$$K_{\hat{A}1} = s_{\hat{A}} \cdot (T_{\hat{B}} - V_{\hat{B}}) + \hat{a}(R_{\hat{B}} + H_1(\hat{B} \parallel R_{\hat{B}}) \cdot P_{pub}),$$

$$K_{\hat{A}2} = \hat{a} \cdot (T_{\hat{B}} - V_{\hat{B}}).$$

Then $\hat{A}$ checks $MAC_{\hat{B}}$. If $VER(K_{\hat{A}1} \parallel K_{\hat{A}2}, R_{\hat{B}} \parallel V_{\hat{B}} \parallel Eph_{\hat{B}}, MAC_{\hat{B}})$ equals to 1, it is valid. $\hat{A}$ generates the session key $SK_{\hat{A}\hat{B}} = H_2(\hat{A} \parallel \hat{B} \parallel T_{\hat{A}} \parallel T_{\hat{B}} \parallel K_{\hat{A}1} \parallel K_{\hat{A}2})$ and sends the message $\hat{I}_2$ to $\hat{B}$:

$$\hat{A} \rightarrow \hat{B} : \hat{I}_2 = \{T_{\hat{B}}, MAC_{\hat{A}}\},$$

where $MAC_{\hat{A}} = HMAC(K_{\hat{A}1} \parallel K_{\hat{A}2}, R_{\hat{A}} \parallel V_{\hat{A}} \parallel Eph_{\hat{A}})$.

If $VER(K_{\hat{A}1} \parallel K_{\hat{A}2}, R_{\hat{B}} \parallel V_{\hat{B}} \parallel Eph_{\hat{B}}, MAC_{\hat{B}})$ equals to 0, it is invalid. $\hat{A}$ aborts the session.

Step 4. After receiving the message $\hat{I}_2$, $\hat{B}$ checks $MAC_{\hat{A}}$. If $VER(K_{\hat{B}1} \parallel K_{\hat{B}2}, R_{\hat{A}} \parallel V_{\hat{A}} \parallel Eph_{\hat{A}}, MAC_{\hat{A}})$ equals to 1, it is valid. $\hat{B}$ generates the session key $SK_{\hat{B}\hat{A}} = H_2(\hat{A} \parallel \hat{B} \parallel T_{\hat{A}} \parallel T_{\hat{B}} \parallel K_{\hat{B}1} \parallel K_{\hat{B}2})$.

If $VER(K_{\hat{B}1} \parallel K_{\hat{B}2}, R_{\hat{A}} \parallel V_{\hat{A}} \parallel Eph_{\hat{A}}, MAC_{\hat{A}})$ equals to 0, it is invalid. $\hat{B}$ aborts the session.

4 Analysis of Two UPIAP Protocols

This section will analyze Zhang *et al.*'s two UPIAP protocols and point out security flaws of two UPIAP protocols. Since UPIAP1 protocol and UPIAP2 protocol are similar in structure, we only describe the analysis of UPIAP1 protocol.

4.1 Analysis of Forward Security

In the UPIAP1 protocol, Zhang *et al.* claimed that the adversary could not obtain previous session keys, even if the adversary could get $\hat{A}$ and $\hat{B}$'s long-term private keys by stolen device attack. However, we carefully analyze the UPIAP1 protocol, and prove this protocol without forward security.

The adversary $\mathcal{M}$ can obtain $\hat{A}$'s secret key $(s_{\hat{A}}, v_{\hat{A}})$. Since the public ephemeral message $Eph_{\hat{A}} = \hat{a} + v_{\hat{A}}$, $\mathcal{M}$ can compute the value of $\hat{a}$ through the public ephemeral message $Eph_{\hat{A}}$. Further, $\mathcal{M}$ can use $s_{\hat{A}}$ and $v_{\hat{A}}$ to compute the session key components $K_{\hat{A}1}$ and $K_{\hat{A}2}$. Finally, $\mathcal{M}$ can use $K_{\hat{A}1}$ and $K_{\hat{A}2}$ to recover the previous session key $SK_{\hat{A}\hat{B}} = H_2(\hat{A} \parallel \hat{B} \parallel T_{\hat{A}} \parallel T_{\hat{B}} \parallel K_{\hat{A}1} \parallel K_{\hat{A}2})$, because $\hat{A}, \hat{B}, T_{\hat{A}}$ and $T_{\hat{B}}$ are also public messages.

Similarly, the adversary $\mathcal{M}$ can mount the attack to the UPIAP2 protocol successfully. So two proposed protocols are lack of forward security.

4.2 KCI Attack

The key compromise impersonation (KCI) attack resilience is a basic attribute for AKE protocols. In this subsection, we will prove that the UPIAP1 protocol cannot resist KCI attack. We assume the adversary $\mathcal{M}$ has obtained party $\hat{A}$'s secret key $(s_{\hat{A}}, v_{\hat{A}})$. Then the adversary $\mathcal{M}$ impersonates party $\hat{B}$ to cheat $\hat{A}$. The KCI attack's details are as follows.

Step 1. The initiator $\hat{A}$ randomly generates a value $\hat{a} \in Z_p^*$. Then $\hat{A}$ sends the message $\hat{I}_1$ to the responder $\hat{B}$ as follows:

$$\hat{A} \rightarrow \hat{B} : \hat{I}_1 = \{R_{\hat{A}}, V_{\hat{A}}, Eph_{\hat{A}}\},$$

where $Eph_{\hat{A}} = \hat{a} + v_{\hat{A}}$.

Step 2. The adversary $\mathcal{M}$ intercepts the message $\hat{I}_1$, $\mathcal{M}$ randomly generates a value $\hat{m} \in Z_p^*$ and computes $T_{\mathcal{M}} = \hat{m} \cdot P + V_{\hat{B}}$. Since $\mathcal{M}$ has $s_{\hat{A}}$ and $v_{\hat{A}}$, $\mathcal{M}$ can compute $\hat{a}$ from $Eph_{\hat{A}}$ and $v_{\hat{A}}$. Then $\mathcal{M}$ computes the session key components as follows:

$$K_{\mathcal{M}1} = s_{\hat{A}} \cdot (T_{\mathcal{M}} - V_{\hat{B}}) + \hat{a}(R_{\hat{B}} + H_1(\hat{B} \parallel R_{\hat{B}}) \cdot P_{pub}),$$

$$K_{\mathcal{M}2} = \hat{a} \cdot (T_{\mathcal{M}} - V_{\hat{B}}),$$

Finally, the adversary $\mathcal{M}$ impersonates $\hat{B}$ to send $\hat{R}_{\mathcal{M}}$ to $\hat{A}$ as follows:

$$\hat{B}(\mathcal{M}) \rightarrow \hat{A} : \hat{R}_{\mathcal{M}} = \{R_{\hat{B}}, V_{\hat{B}}, T_{\mathcal{M}}, T_{\hat{A}}, MAC_{\mathcal{M}}\},$$

where $MAC_{\mathcal{M}} = HMAC(K_{\mathcal{M}1} \parallel K_{\mathcal{M}2}, R_{\hat{B}} \parallel V_{\hat{B}} \parallel T_{\mathcal{M}} \parallel T_{\hat{A}})$.

Step 3. After receiving the message $\hat{R}_{\mathcal{M}}$, $\hat{A}$ computes the session key components as follows:

$$K_{\hat{A}1} = s_{\hat{A}} \cdot (T_{\mathcal{M}} - V_{\hat{B}}) + \hat{a}(R_{\hat{B}} + H_1(\hat{B} \parallel R_{\hat{B}}) \cdot P_{pub}),$$

$$K_{\hat{A}2} = \hat{a} \cdot (T_{\mathcal{M}} - V_{\hat{B}}).$$

Then $\hat{A}$ checks $MAC_{\mathcal{M}}$. If $VER(K_{\hat{A}1} \parallel K_{\hat{A}2}, R_{\hat{B}} \parallel V_{\hat{B}} \parallel Eph_{\hat{B}}, MAC_{\mathcal{M}})$ equals to 1, it is valid. $\hat{A}$ generates the session key $SK_{\hat{A}\hat{B}} = H_2(\hat{A} \parallel \hat{B} \parallel T_{\hat{A}} \parallel T_{\mathcal{M}} \parallel K_{\hat{A}1} \parallel K_{\hat{A}2})$ and sends the message $\hat{I}_2$ to $\hat{B}$:

$$\hat{A} \rightarrow \hat{B} : \hat{I}_2 = \{MAC_{\hat{A}}\},$$

where $MAC_{\hat{A}} = HMAC(K_{\hat{A}1} \parallel K_{\hat{A}2}, R_{\hat{A}} \parallel V_{\hat{A}} \parallel Eph_{\hat{A}})$.

Step 4. After intercepting the message $\hat{I}_2$, the adversary $\mathcal{M}$ also computes the session key $SK_{\mathcal{M}\hat{A}} = H_2(\hat{A} \parallel \hat{B} \parallel T_{\hat{A}} \parallel T_{\mathcal{M}} \parallel K_{\mathcal{M}1} \parallel K_{\mathcal{M}2})$.

Since we have $K_{\hat{A}1} = K_{\mathcal{M}1}$ and $K_{\hat{A}2} = K_{\mathcal{M}2}$, it means that the adversary $\mathcal{M}$ can pass $\hat{A}$'s verification successfully and generate the same session key as $\hat{A}$.

Similarly, the adversary $\mathcal{M}$ can mount KCI attack to the UPIAP2 protocol successfully.

4.3 Partial KCI Attack

In this subsection, we will prove that the UPIAP1 protocol cannot resist partial KCI attack either. We assume the adversary $\mathcal{M}$ has only obtained party $\hat{A}$'s partial secret key $v_{\hat{A}}$. Then the adversary $\mathcal{M}$ impersonates party $\hat{B}$ to cheat $\hat{A}$. The partial KCI attack's details are as follows.

Step 1. The initiator $\hat{A}$ randomly generates a value $\hat{a} \in Z_p^*$. Then $\hat{A}$ sends the message $\hat{I}_1$ to the responder $\hat{B}$ as follows:

$$\hat{A} \rightarrow \hat{B} : \hat{I}_1 = \{R_{\hat{A}}, V_{\hat{A}}, Eph_{\hat{A}}\},$$

where $Eph_{\hat{A}} = \hat{a} + v_{\hat{A}}$.

Step 2. The adversary $\mathcal{M}$ intercepts the message $\hat{I}_1$, $\mathcal{M}$ randomly generates a value $\hat{m} \in Z_p^*$ and computes $T_{\mathcal{M}} = \hat{m} \cdot P + V_{\hat{B}}$. Since $\mathcal{M}$ has obtained $v_{\hat{A}}$, $\mathcal{M}$ can compute $\hat{a}$ from $Eph_{\hat{A}}$ and $v_{\hat{A}}$. Then $\mathcal{M}$ computes the session key components as follows:

$$\begin{aligned} K_{\mathcal{M}1} &= \hat{m} \cdot (R_{\hat{A}} + H_1(\hat{A} \parallel R_{\hat{A}}) \cdot P_{pub}) + \\ &\quad \hat{a}(R_{\hat{B}} + H_1(\hat{B} \parallel R_{\hat{B}}) \cdot P_{pub}), \\ K_{\mathcal{M}2} &= \hat{a} \cdot (T_{\mathcal{M}} - V_{\hat{B}}). \end{aligned}$$

Finally, the adversary $\mathcal{M}$ impersonates $\hat{B}$ and sends $\hat{R}_{\mathcal{M}}$ to $\hat{A}$ as follows:

$$\hat{B}(\mathcal{M}) \rightarrow \hat{A} : \hat{R}_{\mathcal{M}} = \{R_{\hat{B}}, V_{\hat{B}}, T_{\mathcal{M}}, T_{\hat{A}}, MAC_{\mathcal{M}}\},$$

where $MAC_{\mathcal{M}} = HMAC(K_{\mathcal{M}1} \parallel K_{\mathcal{M}2}, R_{\hat{B}} \parallel V_{\hat{B}} \parallel T_{\mathcal{M}} \parallel T_{\hat{A}})$.

Step 3. After receiving the message $\hat{R}_{\mathcal{M}}$, $\hat{A}$ computes the session key components as follows:

$$K_{\hat{A}1} = s_{\hat{A}} \cdot (T_{\mathcal{M}} - V_{\hat{B}}) + \hat{a}(R_{\hat{B}} + H_1(\hat{B} \parallel R_{\hat{B}}) \cdot P_{pub}),$$

$$K_{\hat{A}2} = \hat{a} \cdot (T_{\mathcal{M}} - V_{\hat{B}}).$$

Then $\hat{A}$ checks $MAC_{\mathcal{M}}$. We have

$$\begin{aligned} K_{\mathcal{M}1} &= \hat{m} \cdot (R_{\hat{A}} + H_1(\hat{A} \parallel R_{\hat{A}}) \cdot P_{pub}) \\ &\quad + \hat{a}(R_{\hat{B}} + H_1(\hat{B} \parallel R_{\hat{B}}) \cdot P_{pub}) \\ &= \hat{m}s_{\hat{A}}P + \hat{a}s_{\hat{B}}P \\ &= s_{\hat{A}} \cdot (T_{\mathcal{M}} - V_{\hat{B}}) \\ &\quad + \hat{a}(R_{\hat{B}} + H_1(\hat{B} \parallel R_{\hat{B}}) \cdot P_{pub}) \\ &= K_{\hat{A}1}, \\ K_{\mathcal{M}2} &= \hat{a} \cdot (T_{\mathcal{M}} - V_{\hat{B}}) = \hat{a}\hat{m}P \\ &= \hat{m}\hat{a}P \\ &= \hat{a} \cdot (T_{\mathcal{M}} - V_{\hat{B}}) \\ &= K_{\hat{A}2}. \end{aligned}$$

So $VER(K_{\hat{A}1} \parallel K_{\hat{A}2}, R_{\hat{B}} \parallel V_{\hat{B}} \parallel Eph_{\hat{B}}, MAC_{\mathcal{M}})$ equals to 1, it is valid. $\hat{A}$ generates the session key $SK_{\hat{A}\hat{B}} = H_2(\hat{A} \parallel \hat{B} \parallel T_{\hat{A}} \parallel T_{\mathcal{M}} \parallel K_{\hat{A}1} \parallel K_{\hat{A}2})$ and sends the message $\hat{I}_2$ to $\hat{B}$:

$$\hat{A} \rightarrow \hat{B} : \hat{I}_2 = \{MAC_{\hat{A}}\},$$

where $MAC_{\hat{A}} = HMAC(K_{\hat{A}1} \parallel K_{\hat{A}2}, R_{\hat{A}} \parallel V_{\hat{A}} \parallel Eph_{\hat{A}})$.

Step 4. After intercepting the message $\hat{I}_2$, the adversary $\mathcal{M}$ also computes the session key $SK_{\mathcal{M}\hat{A}} = H_2(\hat{A} \parallel \hat{B} \parallel T_{\hat{A}} \parallel T_{\mathcal{M}} \parallel K_{\mathcal{M}1} \parallel K_{\mathcal{M}2})$.

Since we have $K_{\hat{A}1} = K_{\mathcal{M}1}$ and $K_{\hat{A}2} = K_{\mathcal{M}2}$, it means that the adversary $\mathcal{M}$ can generate the same session key as $\hat{A}$.

Similarly, the adversary $\mathcal{M}$ can mount partial KCI attack to the UPIAP2 protocol successfully.

5 Conclusion

Secure communication is a vital point in disaster environment, and encryption is the basic guarantees for communication messages. There have existed many AKE protocols to generate session keys for encryption. Especially, pairing-free identity-based AKE protocols are more adapt for Internet of Things to generate these session keys. In this paper, we analyzes the UPIAP1 protocol and UPIAP2 protocol, which are two pairing-free identity-based AKE protocols proposed by Zhang *et al.* in 2019. The analysis results show that two UPIAP protocols cannot obtain the attribute of forward security, or resist KCI attack as well as partial KCI attack. The main reason for this situation is that there are some security flaws in the misusage of ephemeral key and long-term private key. For designing better protocols to remedy these flaws, we recommend to use the method in [1, 12].

Acknowledgments

The authors would like to thank Prof. Min-Shiang Hwang and the anonymous referees for their helpful comments.

This work was supported by the National Natural Science Foundation of China (No. 61872449).

References

[1] S. Bala, G. Sharma, A. Verma, "PF-ID-2PAKA: pairing free identity-based two-party authenticated key agreement protocol for wireless sensor networks," *Wireless Personal Communications*, vol. 87, no. 3, pp. 995-1012, 2016.

[2] M. Bellare, D. Pointcheval, P. Rogaway, "Authenticated key exchange secure against dictionary attacks," in *Proceedings of Advances in Cryptology*, pp. 139-155, 2000.

[3] X. Cao, W. Kou, X. Du, "A pairing-free identity-based authenticated key agreement protocol with minimal message exchanges," *Information Sciences*, vol. 180, no. 15, pp. 2895-2903, 2010.

[4] Q. Cheng, X. Zhang, "Comments on privacy-preserving Yoking proof with key exchange in the three-party setting," *International Journal of Network Security*, vol. 21, no. 2, pp. 355-358, 2019.

[5] L. Dang, J. Xu, X. Cao, H. Li, J. Chen, Y. Zhang, X. Fu, "Efficient identity-based authenticated key agreement protocol with provable security for vehicular ad hoc networks," *International Journal of Distributed Sensor Networks*, vol. 14, no. 4, pp. 1-16, 2018.

[6] C. Kudla, K. G. Paterson, "Modular security proofs for key agreement protocols," in *Proceedings of Advances in Cryptology*, pp. 549-565, 2005.

[7] B. LaMacchia, K. Lauter, A. Mityagin, "Stronger security of authenticated key exchange," in *Proceedings of ProvSec, First International Cconference on Provable Security*, pp. 1-16, 2007.

[8] C. C. Lee, M. S. Hwang, L. H. Li, "A new key authentication scheme based on discrete logarithms", *Applied Mathematics and Computation*, vol. 139, no. 2, pp. 343-349, July 2003.

[9] I. C. Lin, C. C. Chang, M. S. Hwang, "Security enhancement for the simple authentication key agreement algorithm", in *Proceedings 24th Annual International Computer Software and Applications Conference (COMPSAC'00)*, 2000.

[10] C. Ling, S. Chen, M. Hwang, "Cryptanalysis of Tseng-Wu group key exchange protocol," *International Journal of Network Security*, vol. 18, no. 3, pp. 590-593, 2016.

[11] S. Nathani, B. Tripathi, S. Khatoon, "A dynamic ID based authenticated group key agreement protocol from pairing," *International Journal of Network Security*, vol. 21, no. 4, pp. 582-591, 2019.

[12] L. Ni, G. Chen, J. Li, Y. Hao, "Strongly secure identity-based authenticated key agreement protocols without bilinear pairings," *Information Sciences*, vol. 367, no. 11, pp. 176-193, 2016.

[13] G. R. Thomas, P. Armstrong, A. Boulgakov, A. Roscoe, "FDR3: A modern refinement checker for CSP," in *Proceedings of Tools and Algorithms for the Construction and Analysis of Systems*, pp. 187-201, 2014.

[14] J. Zhang, X. Huang, W. Wang, Y. Yue, "Unbalancing pairing-free identity-based authenticated key exchange protocols for disaster scenarios," *IEEE Internet of Things Journal*, vol. 6, no. 1, pp. 878-890, 2019.

Biography

Qingfeng Cheng received his B. A. degree in 2000 and M. S. degree in 2004 from National University of Defense Technology, and Ph.D. degree in 2011 from Zhengzhou Information Science and Technology Institute. He is now an associate professor in the State Key Laboratory of Mathematical Engineering and Advanced Computing. His research interests include cryptography and information security.

Yuting Li is a graduate student in the State Key Laboratory of Mathematical Engineering and Advanced Computing. Her main research interests include cryptography, edge computing and information security.

Qi Jiang received the B. S. degree in computer science from Shaanxi Normal University in 2005 and Ph.D. degree in computer science from Xidian University in 2011. He is now a professor at School of Cyber Engineering, Xidian University. His research interests include security protocols, wireless network security, cloud security, *etc.*

Xiong Li received the Ph.D. degree in computer science and technology from the Beijing University of Posts and Telecommunications, Beijing, China, in 2012. He is currently an associate professor with the School of Computer Science and Engineering, Hunan University of Science and Technology, Xiangtan, China. He has authored over 100 referred papers. His current research interests include cryptography and information security. He was a recipient of the 2015 Journal of Network and Computer Applications Best Research Paper Award.

An Energy-Efficient Protocol Based on Semi-Random Deployment Algorithm in Wireless Sensors Networks

Alain Bertrand Bomgni and Garrik Brel Jagho Mdemaya
(Corresponding author: Alain Bertrand Bomgni)
Department of Mathematics and Computer Science, University of Dschang
Dschang, Cameroon
(Email: alain.bomgni@gmail.com)
(Received Sept. 12, 2019; Revised and Accepted Jan. 3, 2020; First Online Apr. 19, 2020)

Abstract

The merger between embedded systems and wireless communications has given birth to a new technology called wireless sensor networks. The main purpose of these types of networks is to be able to monitor the area in which the sensors are deployed, in order to collect information and make decisions. The decisions made by the end user thus depend on the quantity and quality of information received at the base station. Therefore, the sensors must be able to collect as much information as possible in the area of interest (AoI), resulting in maximum coverage of this area. Due to the low capacities of sensors, coverage and data collection algorithms need to be energy efficient to ensure a fairly long network lifetime. In this paper we focus on maximizing network lifetime while collecting and sending a big quantity of data to the base station. Our solution is executed in two stages; the first of which is to cover the network as much as possible using static nodes and mobile nodes, and the second presents the process of collecting and processing data to the base station. Compared to many other algorithms in the literature, our solution is better in terms of coverage percentage of the AoI, data received by the base station and in terms of energy consumption.

Keywords: Area Coverage; Energy Consumption; Sensors Deployment Problem; Wireless Sensor Network

1 Introduction

In recent years, the need to observe and control physical phenomena such as temperature, pressure or brightness is essential for many industrial and scientific applications. As a result, many technical and technological advancements in the fields of microelectronics, Micro-mechanical and wireless communication technologies have made it possible to create small communicating objects equipped with a measurement unit, a computing unit, a memory unit and a radio unit for communicating [16]. The massive deployment of these devices in a given area, allows to establish a network whose nodes are sensors: it is a wireless sensors network. With their various advantages, this technology has established itself as a key player in today's communication network architectures [17]. A wireless Sensor Network (WSN), which is a targeted wireless network, consists of a significant number of miniaturized electronic devices, called sensors, distributed over a specified area in order to sense the environment and communicate the accumulated information from the monitored field to other networks (*e.g.*, the internet) [8]. These networks have been extensively used for monitoring of various physical or environmental conditions. These networks are typically deployed in hard-to-reach areas for humans, and once deployed, sensors must work unattended. A WSN has several application perspectives and each application has its own constraints. However, in all areas, the role of a sensor network is almost always the same: the sensors must monitor certain phenomena and send information to a base station, which in turn relays them to an end user via internet [15].

A network of sensors suffers from several technical weak points such as communication range, monitoring range, low battery, and network deployment circumstance problems such as the difficulty of building a sensor network in volcanoes, mountains, or in the oceans [4]. Sensor deployment can either be deterministic or random. In deterministic deployment, coverage can be maximized as a result of optimal placement of sensor nodes. Random deployments are preferred when the region information is not known apriori [1].

In such systems, maximum coverage of the AoI and full connectivity between the deployed nodes are two important factors in sending as much good quality information as possible to the end user through the base station for better decision-making. This can be illustrated when resolving problems like detecting and tracking of intruders

in restricted areas or monitoring volcanic zones. Such applications require full area coverage. Furthermore, the most critical zones should be covered by more than one sensor node.

Various works have been done in the literature in order to solve this type of problems after deployment of sensor nodes. Some of them consider random deployment, others consider a deterministic deployment while others consider both. In this paper, we propose a method to cover as much as possible the AoI after semi-random deployment, using both static and mobile sensors while ensuring full connectivity between the deployed nodes. This method is followed by an algorithm of scheduling node activity that minimizes the energy consumption of the nodes while collecting and sending data to the base station [21]. Despite the encouraging results presented in this paper, what is left is to include a security mechanism to secure the information exchanged by the nodes of the network [14].

The remainder of this paper is organized as follow: In Section 2 we present the various works dealing with the deployment, coverage and connectivity problems; in Section 3, we describe our contribution, then in Section 4 we present differences between our protocol and some other protocols existing in literature. Section 5 deals with some experimental results. A conclusion with open problems ends the paper.

2 Related Works

In the literature, the coverage problem is separated in three types of coverage: Area coverage, barrier coverage and point coverage. Works presented in [10] have been done in order to introduce basic concepts related to coverage and connectivity.

2.1 Area Coverage and Connectivity

The goal in the area coverage problem is to cover the whole area. Therefore, in some cases, the number of sensors is not sufficient; the goal of area coverage becomes maximizing the coverage rate. Works intended to resolve area coverage and connectivity problem are massively done. Recently, [9] proposed a solution which guarantees maximum coverage of the AoI and connectivity between sensors. A schedule algorithm is also proposed in that paper in order to minimize energy consumption of both static and mobile nodes, and both normal nodes and CH. The clustering protocol used and the strategy of feeding empty clusters are not optimal and therefore, sensors exchange too much messages during the first stage of the algorithm. In [19], authors propose a Distributed Scheduling Medium Access Control (DSMAC) algorithm for optimizing the network lifetime of sensor nodes. The geographic distribution of sensor nodes takes into account coverage and network connectivity constraints. Furthermore, DSMAC algorithm allows a full coverage of the monitoring area; but the process of sending data to the base station is not scheduled. In [4] authors achieve both random and deterministic deployment in order to cover as much as possible the area of interest. After deployment, they propose a random node activity scheduling which relies on a random number P_i that helps to determine the next node to be activated to monitor information in a cluster. Thus a node whose residual energy is finished can be chosen to be activated and, since this node is the one that has to select the next node to be activated in a cluster using P_i, this cluster can be paralyzed and sensors in this cluster won't be able to collect information anymore. Connectivity between sensors of this cluster and sensors of the other clusters is therefore impossible. [2] proceeds to a random deployment of static nodes and thereafter, proceeds to deployment of some mobile nodes that are used to repair the coverage holes after initial deployment of the static nodes. This solution ensures a good coverage ratio but not connectivity between sensors. [11] proposed an algorithm that guarantees full coverage and multiple connectivity [10] after regular sensors deployment. But this solution assumes that the AoI is regular. In [8], a deployment approach based on flower pollination algorithm (FPCOA) was proposed. This approach can find the optimal placement topology in terms one QoS metric and ensures simple connectivity between sensors but it did not incorporate other QoS metrics like energy consumption

2.2 Barrier Coverage and Connectivity

Wireless sensors networks are not only designed to sense events occurring in the deployment area; they can also be used to detect intruders that attempt to penetrate in this area. So, the goal of barrier coverage is to guarantee that every intruder crossing the barrier of sensor will be detected. Few works are present in the literature for barrier coverage and connectivity. Nevertheless, we can cite the solutions of [20] and [13]. [20] provides partial coverage after a centralized and probabilistic deployment. The connectivity in this case is intermittent; Meaning that, some of the deployed sensors can not communicate with the base station. [13] made The assumption that $R \geq r$ where R is the communication radius and r is the sensing radius in order to ensure full coverage and permanent connectivity after a distributed and deterministic deployment.

2.3 Point Coverage and Connectivity

It is often unnecessary to monitor the whole area in many applications; thus monitoring some specific points is sufficient. Each of these points (called point of interest (PoI)) should therefore be covered by at least one sensor node. [6] assume that PoI are static and guarantee temporary coverage and intermittent connectivity with random deployment and distributed algorithm. [5] resolved a similar problem differently, and consider that PoI are not static. [7] considers the problem of full coverage with permanent connectivity. In fact, the authors ensure full

1	2	3	4	5
6	7	8	9	10
11	12	13	14	15
16	17	18	19	20
21	22	23	24	25

Figure 1: Subdivision of AoI in sub-areas

coverage of the PoI using forced based deployment algorithm. These solutions do not propose a node scheduling activity to minimize the energy consumption of the sensors.

Each of the works presented in this section addresses the problem of coverage and connectivity in different ways. However, depending on the constraints related to the deployment environment or the types of sensors, the proposed protocols rarely take into account the energy consumption of the sensors and their activity during the lifetime of the network. The solution proposed in this paper ensures full maximum coverage of the AoI, connectivity between sensors and minimizes energy consumption of the sensors. It also guarantees that the roles of the different CHs can be exchanged because of the use of ICP [12]. Finally the scheduling algorithm allows to collect any event occurring in the AoI and send it to the base station.

3 Our Contribution

3.1 Assumptions and Notations

3.1.1 Assumptions

In this work, it was assumed that:

- The area of interest is a square of side C;
- The area of interest must be divided geographically by N_z sub-areas of dimension L×L and diagonal D as shown in Figure 1;
- Using [12], the number of CHs can be estimated;
- All the sensor have the same sensing range and the same communication radius and are able to know in which sub-area it has been deployed.

3.1.2 Notations

In the rest of this work, we will use some notations that we define in this section:

- T_i: Awakening time of a normal node;
- T_s: Time used by a normal node to send data to its CH and receive an acknowledgment;
- T_c: Time used by a normal node to collect data in its cluster;
- T_{ne}: Next awakening time of a normal node;
- N_c: Number of sensors of a cluster;
- T_{sSB}: Time after which a CH should send data to the base station;
- R_c: Communication radius of a sensor;
- R_s: Sensing radius of a sensor;
- n_{zc}: Number of sub-areas covered by a CH;
- n_{znc}: Number of non covered sub-areas;
- N_z: Total number of sub-areas;
- S: The set of sub-areas;
- S_{CH}: The set of sub-areas covered by a CH.

3.2 Mathematics Models

We represent the WSN by a graph G = (V;E), where V represents all nodes of the network and E represents the set of edges giving all possible communications.

3.2.1 Coverage Model

Let A represent the AoI and q a point located in A. The area covered by a sensor S_i∈V is defined as the total area located within R_s [17]. Analytically, the area covered by a sensor S_i∈V is given by Equation (1):

$$C(S_i) = \{q \in A / d(S_i; q) \leq R_s\} \quad (1)$$

So, the area covered by a set of sensors S = {S_1, S_2, ..., S_k} is analytically defined by Equation (2):

$$C(S) = \bigcup C(S_k), k = \{1, ..., |S|\} \quad (2)$$

3.2.2 Connectivity Model

Let us consider S_i and S_j two sensors nodes deployed in the AoI. S_i and S_j are directly connected (one-hop connectivity) if and only if d(S_i;S_j)≤R_c. According to [18], a WSN is considered to be connected if there is at least one path between the sink and each node in the considered area.

3.2.3 Lifetime Model

Let M = {S_1, S_2, ..., S_n} be the set of nodes of a wireless sensor network; S_i ∈ M a given node with lifetime T_i. In [17], network lifetime is defined by the duration within which the network is deployed and the first node loses all its residual energy. So, if T_n is the network lifetime, it is computed as follows (equation (3)):

$$T_n = min T_i \quad (3)$$

3.3 First Stage: Area Coverage Procedure

The first stage of our solution consists in covering as much as possible the AoI in order to collect a maximum number of information. To achieve this, we proceed as follows:

- First deploy deterministically the different CHs in the AoI such as d(CH_i,CH_j)≤2R_c and such as each CH is placed at the center of its sub-area;
- Static nodes are then randomly deployed in the AoI. The idea here is to allow each sensor to belong to the cluster of a CH;
- Application of ICP [12] to initiate clustering. In ICP, acknowledgments are deleted in order to reduce energy consumption during the clustering process. But in our case, acknowledgments will be allowed in order to permit to each sensor to send an acknowledgment to its CH. So, the clustering process becomes:
  - Each CH broadcasts its id to neighbors sensors;
  - If a sensor receives one message from one CH, it becomes a cluster member (CM) of this CH; but if it receives many messages from many CH, it becomes a gateway (GW) node for all the clusters of these CH;
  - Sensors can then send an acknowledgment to the CH containing its id, its role (CM or GW) and the identifier of the sub-area in which it is located;
  - When a CH receives an acknowledgment, it increments the variable n_{zc};
  - After reception of all acknowledgments, each CH then broadcasts the ordered list of its cluster's members to all its cluster's members with parameters T_i, T_s, T_c and N_c. Thus, each sensor will be able to know its CH and all its neighbors in a cluster.

Theorem 1. *Let S_{CH} be the set of sub-areas covered by a CH. The number of non covered sub-areas n_{znc} can be computed by $n_{znc} = N_z$ - $\boldsymbol{C}(\bigcup(S_{CH}))$ where $\boldsymbol{C}$ is the function to determine the cardinal (number of elements) of a set.*

Proof. Since N_z is the total number of sub-areas, n_{znc} can be obtained by a substraction between N_z and the total number of sub-areas covered by the different CHs. Since two CHs can cover the same sub-areas, the function **C** such a way that it removes duplicates entries. Finally, the application of the function C($\bigcup(S_{CH})$) makes it possible to obtain the exact number **n** of sub-areas covered by the different CHs; and therefore, doing N_z-n yields the number of uncovered sub-areas. □

At the end of the previous steps, each static sensor knows in which cluster it belongs. if n_{znc} is equal to zero, the AoI is fully covered. Else, the challenge is to find a way to cover non covered sub-areas using mobile nodes. To do this, we proceed as follows:

- First we determine the identifiers of all non covered sub-areas using formula: S \ $\bigcup(S_{CH})$);
- Secondly, we deploy mobile nodes in these sub-areas.

The algorithm of this stage is given by algorithm 1.

Algorithm 1 AoI coverage

```
Begin
Deterministic deployment of CHs.
Random deployment of static nodes.
Each CH broadcasts its id.
if sensor receives only one message then
    Become a cluster member of the CH.
end if
if sensor receives many messages then
    Become a gateway node.
end if
Sensors can then send an acknowledgment to the CH containing its id, its role (CM or GW) and the identifier of its sub-area.
Each CH broadcasts a list of its cluster's members with the parameters T_i, T_s and T_c to its cluster's members.
Computation of n_znc.
if n_znc == 0 then
    End of the coverage process.
end if
if n_znc ≠ 0 then
    determine the identifiers of all non covered sub-areas.
    deploy mobile nodes in non covered sub-areas.
end if
End
```

3.4 Second Stage: Node Scheduling Algorithm and Sending Data to the Base Station

In this section, we describe how the nodes will be scheduled in order to collect and send data to the base station. Since normal nodes and CH are scheduled differently, we thus propose two algorithms that will permit us to manage both CH and normal nodes simultaneously.

3.4.1 Normal Nodes Scheduling Algorithm and Sending Data to the CH

Each node has in its memory the ordered list of its neighbors; so it knows when it should wake up and begin collecting or sending data. According to our notations, a normal node remains awake during T_i. We therefore pose $T_i = T_s + T_c$. So, a normal node executes these instructions when it is awakened:

1) It starts by computing the next time after which it should be awaken with the formula: T_{ne}=(N_c-1)T_i;

2) If this node has data collected previously in its memory, it sends it to its CH and waits for an acknowledgment during the time T_s;

3) It remains awake during the time T_c waiting for an event to occur in the AoI;

4) The sensor falls asleep after T_i.

The pseudo-code of our description above is given by the algorithm 2.

Algorithm 2 Normal nodes scheduling algorithm and sending data to the CH

```
Begin
Computation of T_ne.
if node has data in its memory then
    Sending data to the CH during T_s.
    Stay awake during T_c.
    Fall asleep after T_i.
end if
if node has no data in its memory then
    Sending data to the base CH during T_i.
    Fall asleep after T_i.
end if
End
```

3.4.2 CH Scheduling Algorithm and Sending Data to the Base Station

Our solution recommends that every T_{sSB}, a CH must send data to the base station. T_{sSB} is computed with the formula: $T_{sSB} = N_c{*}T_s$; which means that, after one round of diffusion of its cluster members, it starts sending data to the base station. Before sending these data, the CH starts by executing the second part of the DSMAC algorithm [17] which will permit them to synchronize sensors belonging to the path relying the CH and the base station by sending beacon frames. This will permit us to know all the nodes that will remain awake during the transmission of data to the base station. The CH can then initiate the transmission. Algorithm 3 describes the pseudo-code of this solution.

4 Comparative Study Of Our Protocol With Some Others Existing Protocols

In Table 1, we make a comparative study between our protocol and some others.

Algorithm 3 CH scheduling algorithm and sending data to the base station

```
Begin
i=1.
while i ≤ N_c do
    Waking up every T_i and stay awake during T_s.
    Receive data from a normal node and send an acknowledgment to this node.
    T = T_s*i.
    if T == T_sSB then
        Determining the nodes in charge of forwarding data to the BS.
        Sending data to the BS.
        i=1.
    end if
    if T ≠ T_sSB then
        i=i+1.
    end if
end while
End
```

5 Performance Evaluation

In this section we evaluate the performance of our approach and compare it to other approaches. The simulation conditions are shown in the Table 2.

The following curves are the result of at least 100 experiments. In our implementation, the MAC layer is managed in such a way that a node can only receive one message at a time.

5.1 Coverage Ratio

In Figure 2, we make a comparison between our protocol and several others in terms of coverage ratio. In fact, the comparison is made between our protocol and FPCOA [8], SRDP [4] and A2CDC [9].

Because of the semi-random and semi-deterministic deployment, our protocol has the best coverage ratio compared to FPCOA and A2CDC. Since SRDP uses a deployment strategy similar to ours, our algorithm used to feed empty clusters allows us to obtain a better coverage ratio.

5.2 Number of Transmissions During Clustering Stage

The major improvement highlighted in this paper concerns the partitioning protocol and therefore the coverage algorithm. Indeed, it was a question of increasing the coverage ratio while minimizing the energy consumption spent by the sensors during the clustering and coverage phase. This was done by reducing the number of transmissions and messages exchanged during the clustering phase. Figure 3 illustrates graphically what we are explaining.

Table 1: Comparison between our protocol and some others

Protocols	Deployment strategy	Node scheduling algorithm	Clustering algorithm
FCOA [8]	Random	No	No
SRDP [4]	Semi random and semi deterministic (square based)	Random selection of the next activated node	Yes
DSMAC [19]	Deterministic (square based)	Deterministic selection of the next activated node	No
A2CDC [9]	Random	Deterministic selection of the next activated node	Wadaa et al. [22] and Bomgni et al. [3]
Our protocol	Semi random and semi deterministic (square based)	Deterministic selection of the next activated node	ICP [12]

Table 2: Conditions of the simulations

Configurations	Value
Communication and sensing radius	8 m
Area of interest (AoI)	100m×100m
Initial sensor's energy	1000 J
Deploy sensor nodes number	Up to 200

5.3 Network's Lifetime

We compared the efficiency of our protocol with three other protocols named Flower Pollination Coverage Optimization approach (FPCOA) [8], Semi-Random Deployment Protocol (SRDP) [4], DSMAC [19] and A2CDC [9] in terms of energy consumption. The results are shown in Figure 4.

Our protocol is clearly better than the one of SRDP, FPCOA, DSMAC and A2CDC in terms of energy consumption. Since FPCOA doesn't use a clustering scheme to maintain connectivity and reduce energy's consumption of the sensors while exchanging messages, it consumes more energy. The SRDP protocol certainly uses a clustering algorithm, but the latter is not really efficient. In fact, clusters are formed by exchanging *hello* messages between CH and its members. Furthermore, this protocol guarantees connectivity and data harvest by randomly activating a sensor which will collect data in the cluster each time. The fact that the active sensor is determined randomly after a computation of a random parameter P consumes more energy at each time that a sensor has to be activated. Finally, the clustering protocol used in A2CDC is more expensive in terms of energy consumption than the one used in our protocol; Which results in very low power consumption from the beginning of our protocol, due to the very small number of messages exchanged during the clustering phase.

5.4 Average Packets Received By The Sink

Figure 5 illustrates that our protocol outperforms DSMAC, SRDP, FPCOA and A2CDC according to the num-

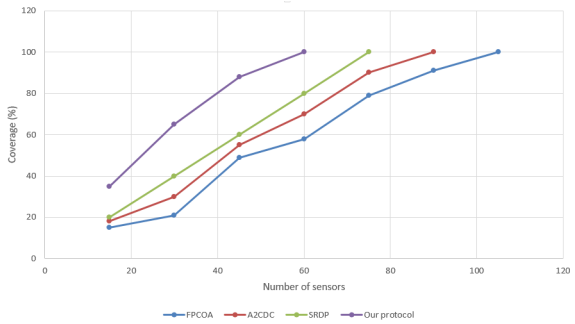


Figure 2: Coverage ratio

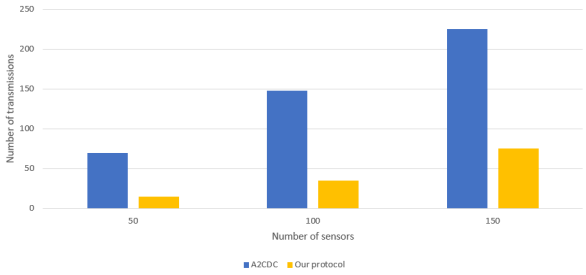


Figure 3: Transmission amount

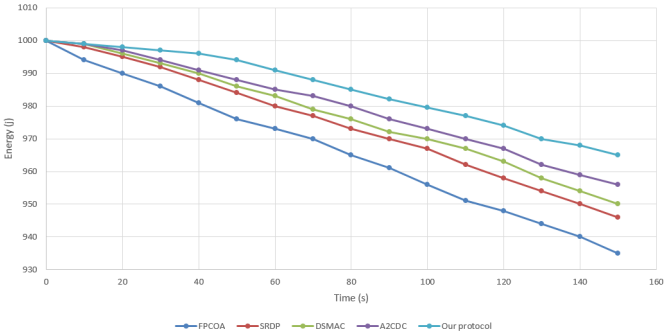


Figure 4: Network's lifetime

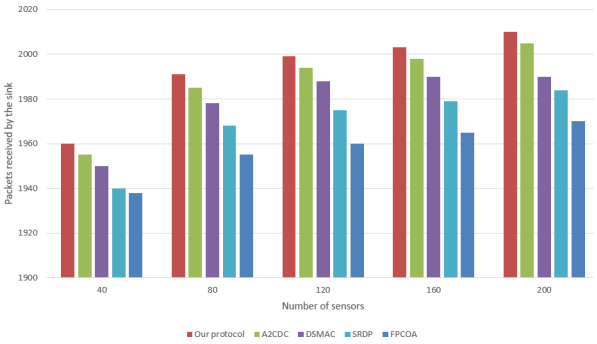


Figure 5: Average packets received by the sink

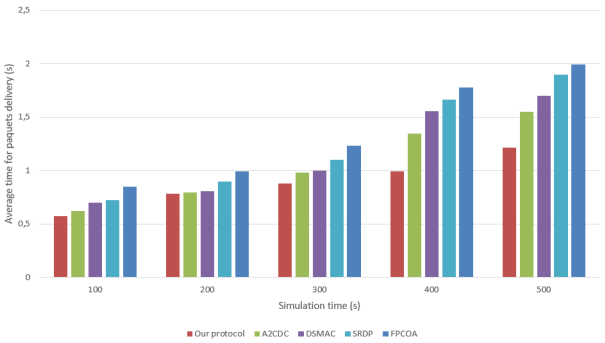


Figure 6: End to end delay

ber of packets received by the Sink. The main reason is due to the fact that our protocol avoids collisions because of the implementation of the CSMA/CA protocol and is based on DSMAC algorithm which mitigates the number of collisions. Our protocol is better than A2CDC only because of the type of deployment. In fact, deterministic deployment ensures more connectivity than random deployment.

5.5 End-to-End Delay

End-to-End delay refers to the time taken for a packet to be transmitted across a network from source to destination. Figure 6 shows that our protocol outperforms DSMAC, SRDP, FPCOA and A2CDC, due to better connectivity between the sensors and better sensors positions within the network.

6 Conclusion

In this paper, we propose an **energy-efficient protocol based on semi-random deployment algorithm ensuring better quality of service and connectivity in wireless sensors networks**, a protocol that aims to optimize coverage and network connectivity while minimizing the energy consumption of sensors during information exchange. To solve the problem, our protocol takes place in two phases: we firstly present our approach to guarantee full coverage of the AoI based on both deterministic and random deployment of sensors, and secondly, we use an algorithm similar to the one presented in [9] to schedule normal nodes and CHs during the phase of collecting data in the monitored area and the phase of sending data to the base station. The proposed approach has been compared with several other approaches in the literature in term of energy consumption, total number of transmissions and average number of packets received by the BS. Experiments show that our solution is better than the other approaches, guarantees connectivity, reduces the number of transmissions and messages and avoids collision of messages.

The results presented in this paper are really encouraging, but several open problems remain. In future work, we plan to introduce a security protocol to ensure the integrity of the data circulating in the network.

Acknowledgments

This study was supported by URIFIA (Unite de Recherche en Informatique Fondamentale Ingenierie et Application). The authors gratefully acknowledge the anonymous reviewers for their valuable comments.

References

[1] J. N. AL-Karaki and A. Gawanmeh, "The optimal deployment, coverage, and connectivity problems in wireless sensor networks: Revisited," *IEEE*, vol. 5, pp. 18051,18065, 2017.

[2] O. Banimelhem, M. Mowafi, and W. Aljoby, "Genetic algorithm based node deployment in hybrid wireless sensor networks," *Communications and Network*, vol. 5, no. 4, 2013.

[3] A. B. Bomgni, G. B. Jagho, E. T. Fute, and C. D. Tayou, "Secure and energy-efficient permutation routing protocol for wireless sensors network deployed in space (3d)," *International Journal of Computer Science and Information Security*, vol. 16, Aug. 2018.

[4] A. Boualem, Y. Dahmani, A. Maatoug, and C. Derunz, "Area coverage optimization in wireless sensor network by semi-random deployment," *Computer Science*, pp. 85–90, 2018.

[5] W. Cheng, M. Li, K. Liu, Y. Liu, X. Li, and X. Liao, "Sweep coverage with mobile sensors," *Mobile Computing IEEE Transactions*, vol. 10, no. 11, 2011.

[6] M. Erdelj, E. Natalizio, and T. Razafindralambo, "Multiple point of interest discovery and coverage with mobile wireless sensors," in *International Conference on Computing, Networking and Communications (ICNC'12)*, 2012. DOI: 10.1109/ICCNC.2012.6167394.

[7] M. Erdelj, T. Razafindralambo, and D. S. Ryl, "Covering points of interest with mobile sensors," *IEEE Transaction on Parallel and Distributed Systems*, vol. 24, no. 1, pp. 32–43, 2012.

[8] F. Hajjej, R. Ejbali, and M. Zaied, "An efficient deployment approach for improved coverage in wireless sensor networks based on flower pollination algorithm," *Computer Science & Information Technology (CS & IT)*, pp. 117,129, 2016. (https://pdfs.semanticscholar.org/842f/0beb97ba1282dac2aae78484454a43dc21eb.pdf)

[9] G. B. Jagho and A. B. Bomgni, "A2cdc: Area coverage, connectivity and data collection in wireless sensor networks," *Network Protocols and Algorithms*, vol. 10, no. 4, pp. 20–34, 2018.

[10] I. Khoufi, P. Minet, A. Laouiti, and S. Mahfoudh, "Survey of deployment algorithms in wireless sensor networks: Coverage and connectivity issues and challenges," *International Journal of Autonomous and Adaptive Communications Systems (IJAACS'17)*, pp. 341,390, 2017.

[11] Y. H. Kim, C. M. Kim, D. S. Yang, Y. J. Oh, and Y. H. Han, "Regular sensor deployment patterns for p-coverage and q-connectivity in wireless sensor networks," *The International Conference on Information Network*, 2012. DOI: 10.1109/ICOIN.2012.6164394.

[12] L. Kong, Q. Xiang, X. Liu, X-Y. Liu, X. Gao, G. Chen, and M-Y. Wu, "ICP: Instantaneous clustering protocol for wireless sensor networks," *Computer Networks*, vol. 101, pp. 144-157, 2016.

[13] L. Kong, Y. Zhu, M. Y. Wu, and W. Shu, "Mobile barrier coverage for dynamic objects in wireless sensor networks," *IEEE 9th International Conference on Mobile Ad-Hoc and Sensor Systems*, 2012. DOI: 10.1109/MASS.2012.6502499

[14] C. T. Li, M. S. Hwang and Y. P. Chu, "An efficient sensor-to-sensor authenticated path-key establishment scheme for secure communications in wireless sensor networks", *International Journal of Innovative Computing, Information and Control*, vol. 5, no. 8, pp. 2107-2124, Aug. 2009.

[15] A. Mansouri and M. S. Bouhlel, "Trust in ad hoc networks: A new model based on clustering algorithm," *International Journal of Network Security*, vol. 21, no. 3, pp. 483–493, May 2019.

[16] L. T. Ngoc and V. T. Tu, "Aodvdc: An improved protocol prevents whirlwind attacks in mobile ad hoc network," *International Journal of Network Security*, vol. 21, no. 2, pp. 333–341, Mar. 2019.

[17] D. Ngom, *Optimisation De La Duree De Vie Dans Les Reseaux De Capteurs Sans Fil Sous Contraintes De Couvertureet De Connectivite Reseau*, 2016. (https://tel.archives-ouvertes.fr/tel-01531464/document)

[18] D. Ngom, P. Lorenz, and B. Gueye, "A distributed scheduling algorithm to improve lifetime in wireless sensor network based on geometric placement of sensors with coverage and connectivity constraints," in *SENSORCOMM 2015: The Ninth International Conference on Sensor Technologies and Applications*, pp. 57–63, 2015.

[19] D. Ngom, P. Lorenz, and B. Gueye, "Dsmac: Constraints-based coverage and connectivity for optimizing the network lifetime in wireless sensor networks," 2017.

[20] H. Shibo, C. Jiming, L. Xu, S. Xuemin, and S. Youxian, "Cost-effective barrier coverage by mobile sensor network," *Proceedings IEEE INFOCOM*, 2012. DOI: 10.1109/INFCOM.2012.6195829.

[21] 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.

[22] A. Wadaa, S. Olariu, L. Wilson, and M. Eltoweissy, "Training a wireless sensor networks," *Mobile Networks and Applications*, vol. 10, pp. 151–168, 2005.

Biography

Alain Bertrand BOMGNI is a lecturer in the department of Mathematics and Computer Science of the University of Dschang, Cameroon. He obtained his Ph.D at the University of Picardie Jules Verne (France) in 2013, his M.S. degree at the University of Yaounde I in 2006, and his B.S. degree at University of Dschang in 2002, all in Computer Science. His current research interests include parallel algoritms and architectures, distributed systems, wireless sensor network and Internet of Things.

Garrik Brel Jagho Mdemaya is a Phd student at the University of Dschang, Cameroon. He obtained his Master degree in computer science in 2016 at the University of Dschang. His current research interests include wireless communication, ad hoc networking and Internet of Things.

A Family of Pseudorandom Binary Sequences Derived from Generalized Cyclotomic Classes Modulo $p^{m+1}q^{n+1}$

Xiaolin Chen[1], Zhixiong Chen[2], and Huaning Liu[1]

(Corresponding author: Huaning Liu)

School of Mathematics, Northwest University[1]
Xi'an 710127, Shaanxi, P. R. China
Provincial Key Laboratory of Applied Mathematics, Putian University[2]
Putian 351100, Fujian, P. R. China
(Email: hnliu@nwu.edu.cn)

(Received Feb. 22, 2019; Revised and Accepted Sept. 6, 2019; First Online Feb. 29, 2020)

Abstract

Let p, q be two distinct odd primes, and let m, n be non-negative integers. We consider a family of binary sequences defined by generalized cyclotomic classes modulo $p^{m+1}q^{n+1}$. The first contribution is to determine their linear complexity, which improves certain results of Hu, Yue and Wang. The second contribution is to compute the autocorrelation values. Results obtained indicate that such sequences are 'good' from the viewpoint of cryptography.

Keywords: Autocorrelation Value; Generalized Cyclotomy; Generalized Cyclotomic Sequence; Linear Complexity; Stream Cipher

1 Introduction

The theory of cyclotomy is widely applied in cryptography. A typical application is the design of pseudorandom sequences or numbers. By defining the (generalized) cyclotomic classes modulo an integer, families of pseudorandom sequences can be designed with the desired cryptographic features. The classical examples are the Legendre sequences that derived from cyclotomic classes modulo an odd prime and the Jacobi sequences that derived from generalized cyclotomic classes modulo the product of two odd distinct primes. Attention is also paid to the generalized cyclotomic classes modulo a general number in the literature, see *e.g.*, [1–5, 9, 11, 12].

At the beginning of this decade, Hu, Yue and Wang [6] introduced families of binary sequences via defining generalized cyclotomic classes modulo N, where $N = p^{m+1}q^{n+1}$ for two distinct odd primes p and q and non-negative integers m and n. Let

$$\begin{aligned} d &= (p-1, q-1) = \left(\phi\left(p^{m+1}\right), \phi\left(q^{n+1}\right)\right), \\ e &= \frac{\phi\left(p^{m+1}\right)\phi\left(q^{n+1}\right)}{d}, \end{aligned}$$

where ϕ denotes the Euler function. Let g be a common primitive root of p^{m+1} and q^{n+1}, and let x be an integer satisfying

$$x \equiv g \pmod{p^{m+1}}, \quad x \equiv 1 \pmod{q^{n+1}}.$$

Define

$$G_i = \left\{g^s x^i : s = 0, 1, \cdots, e-1\right\}, \quad i = 0, 1, \cdots, d-1.$$

Then

$$\mathbb{Z}^*_{p^{m+1}q^{n+1}} = \bigcup_{i=0}^{d-1} G_i$$

For $0 \le a \le m+1$ and $0 \le b \le n+1$, let

$$G_i^{(a,b)} = \begin{cases} p^a q^b G_i, & \text{if } a \le m,\ b \le n,\ 0 \le i \le d-1, \\ p^a q^{n+1}\mathbb{Z}_N^*, & \text{if } a \le m,\ b = n+1,\ i = 0, \\ p^{m+1} q^b \mathbb{Z}_N^*, & \text{if } a = m+1,\ b \le n,\ i = 0, \\ \{0\}, & \text{if } a = m+1,\ b = n+1,\ i = 0. \end{cases}$$

Then Hu, Yue and Wang [6] introduced the binary sequence s^∞ of period N by setting

$$s_j = \begin{cases} 1, & \text{if } (j \bmod N) \in \Omega, \\ 0, & \text{otherwise}, \end{cases} \tag{1}$$

where Ω, usually called the *characteristic set* or *support set* of s^∞, is selected as

$$\Omega = \bigcup_{a=0}^{m+1} \bigcup_{b=0}^{n+1} \bigcup_{i \in I_{a,b}} G_i^{(a,b)},$$

for

$$I_{a,b} \subset \begin{cases} \{0,1,\cdots,d-1\}, & \text{if } a \le m,\ b \le n, \\ \{0\}, & \text{otherwise.} \end{cases} \quad (2)$$

They developed a way to compute the *linear complexity* (see the notion below) of s^∞. However, it seems difficult to determine the exact values due to the choice of $I_{a,b}$, see [6, Thm.2.5]. Motivated by this reason, we will only choose a special $I_{a,b}$ as follows and consider the linear complexity and *autocorrelation* (see the notion below) of the special binary sequence:

$$I_{a,b} = \begin{cases} \{1,3,5,\cdots,d-1\}, \\ \qquad \text{if } 0 \le a \le m \text{ and } 0 \le b \le n, \\ \emptyset, \\ \qquad \text{if } 0 \le a \le m+1 \text{ and } b = n+1, \\ \{0\}, \\ \qquad \text{if } a = m+1 \text{ and } 0 \le b \le n. \end{cases} \quad (3)$$

We remark that, results of autocorrelation of such sequences have not been reported in the literature. We organise this work as follows. In Section 2 we prove the linear complexity of sequence defined in Equation (1) with $I_{a,b}$ in Equation (3) and compute its autocorrelation values in Section 3. Finally we draw a conclusion in Section 4. We conclude this section by introducing the notions of linear complexity and autocorrelation of sequences.

The linear complexity is an important cryptographic characteristic of sequences and provides information on predictability and thus unsuitability for cryptography. Let $\mathbb{F}$ be a field. For a T-periodic sequence s^∞ over $\mathbb{F}$, the *linear complexity* $L(s^\infty)$ of the sequence s^∞ is defined to be the length of the shortest linear feedback shift register that can generate the sequence, which is the smallest nonnegative integer L satisfying

$$s_t = c_1 s_{t-1} + c_2 s_{t-2} + \cdots + c_L s_{t-L} \text{ for all } t \ge L,$$

where constants $c_1, \cdots, c_L \in \mathbb{F}$. Let

$$s(X) = s_0 + s_1 X + \cdots + s_{T-1} X^{T-1} \in \mathbb{F}[X],$$

which is called the *generating polynomial* of s^∞. Then the linear complexity over $\mathbb{F}$ of s^∞ can be computed as

$$L(s^\infty) = T - \deg\left(\gcd(X^T - 1,\ s(X))\right), \quad (4)$$

which is the degree of the *characteristic polynomial*, $\frac{X^T-1}{\gcd(X^T-1,\ s(X))}$, of the sequence. Moreover, the *autocorrelation value* $C_s(w)$ of the sequence s^∞ at shift w is defined by

$$C_s(w) = \sum_{i=0}^{T-1} (-1)^{s_{i+w}+s_i},$$

where $1 \le w \le T-1$. See, *e.g.*, [3] for details.

2 Linear Complexity

In this section, we will determine the exact values of the linear complexity of the binary sequences defined in Equation (1) with $I_{a,b}$ in Equation (3). Our result is the following.

Theorem 1. *Let s^∞ be the N-periodic binary sequence defined as in Equation (1) with $I_{a,b}$ in Equation (3) for defining Ω. Then the linear complexity of s^∞ satisfies*

$$L(s^\infty) = p^{m+1}q^{n+1} - \frac{(p^{m+1}-1)(q^{n+1}-1)}{2} - A_{p,m}(q^{n+1}-1) - A_{q,n}(p^{m+1}-1) - 1$$

if $p \equiv \pm 1 \pmod 8$, $q \equiv \pm 1 \pmod 8$ or $p \equiv \pm 3 \pmod 8$, $q \equiv \pm 3 \pmod 8$, and otherwise

$$L(s^\infty) = p^{m+1}q^{n+1} - A_{p,m}(q^{n+1}-1) - A_{q,n}(p^{m+1}-1) - 1,$$

where

$$A_{q,n} = \begin{cases} 1, & \text{if } \frac{(n+1)(q-1)}{2} \equiv 0 \pmod 2, \\ 0, & \text{if } \frac{(n+1)(q-1)}{2} \equiv 1 \pmod 2, \end{cases}$$

$$A_{p,m} = \begin{cases} 1, & \text{if } 1 + \frac{(m+1)(p-1)}{2} \equiv 0 \pmod 2, \\ 0, & \text{if } 1 + \frac{(m+1)(p-1)}{2} \equiv 1 \pmod 2. \end{cases}$$

2.1 Properties of the Generalized Cyclotomic Classes

Lemma 1. *Let α be a primitive N-th root of unity in the field $\mathbb{F}_{2^\delta}$ for $\delta = \mathrm{ord}_N(2)$. Let $(t, pq) = 1$, $0 \le u \le m+1$, $0 \le v \le n+1$.*

1) Suppose that $0 \le a \le m$ and $0 \le b \le n$. Then we have

$$\sum_{l \in G_0^{(a,b)}} \alpha^{tp^u q^v l} = \begin{cases} 0, & \text{if } u < m-a \text{ or } v < n-b, \\ \sum_{l \in G_0^{(m,n)}} \alpha^{tl}, & \text{if } u = m-a,\ v = n-b, \\ \frac{q-1}{d}, & \text{if } u = m-a,\ v > n-b, \\ \frac{p-1}{d}, & \text{if } u > m-a,\ v = n-b, \\ 0, & \text{if } u > m-a,\ v > n-b. \end{cases}$$

2) Suppose that $0 \le a \le m$ and $b = n+1$. Then we have

$$\sum_{l \in G_0^{(a,n+1)}} \alpha^{tp^u q^v l} = \begin{cases} 1, & \text{if } u = m-a, \\ 0, & \text{if } u \ne m-a. \end{cases}$$

3) Suppose that $a = m+1$ and $0 \le b \le n$. Then we have

$$\sum_{l \in G_0^{(m+1,b)}} \alpha^{tp^u q^v l} = \begin{cases} 1, & \text{if } v = n-b, \\ 0, & \text{if } v \ne n-b. \end{cases}$$

Proof. See Lemma 2.4 in [6]. □

According to [10], Whiteman's generalized cyclotomic classes of order d are defined by

$$D_i = \left\{ g^s x^i : s = 0, 1, \cdots, \frac{(p-1)(q-1)}{d} - 1 \right\},$$

where $i = 0, 1, \cdots, d-1$. Clearly,

$$\mathbb{Z}_{pq}^* = \cup_{i=0}^{d-1} D_i, \qquad D_i \cap D_j = \emptyset \text{ for } i \neq j.$$

Lemma 2. *$D_i D_j = D_{(i+j) \bmod d}$, where $i, j = 0, 1, \cdots, d-1$.*

Proof. This is Lemma 1 of [13]. □

Lemma 3. *$2 \in \bigcup_{j=0}^{\frac{d}{2}-1} D_{2j}$ if and only if $p \equiv \pm 1 \pmod 8$, $q \equiv \pm 1 \pmod 8$ or $p \equiv \pm 3 \pmod 8$, $q \equiv \pm 3 \pmod 8$; $2 \in \bigcup_{j=0}^{\frac{d}{2}-1} D_{2j+1}$ if and only if $p \equiv \pm 1 \pmod 8$, $q \equiv \pm 3 \pmod 8$ or $p \equiv \pm 3 \pmod 8$, $q \equiv \pm 1 \pmod 8$.*

Proof. See Theorem 5 in [13]. □

2.2 Proof of Theorem 1

According to Equation (4), the linear complexity of s^∞ can be computed by

$$L(s^\infty) = N - \left| \left\{ t : s(\alpha^t) = 0,\ 0 \leq t < N \right\} \right|,$$

where α is a primitive N-th root of unity in the field $\mathbb{F}_{2^\delta}$ for $\delta = \mathrm{ord}_N(2)$.

We note that $G_k^{(a,b)} = p^a q^b G_k = p^a q^b x^k G_0 = x^k G_0^{(a,b)}$ for $0 \leq a \leq m$, $0 \leq b \leq n$, and

$$\Omega = \bigcup_{a=0}^{m+1} \bigcup_{b=0}^{n+1} \bigcup_{i \in I_{a,b}} G_i^{(a,b)} = \bigcup_{b=0}^{n+1} G_0^{(m+1,b)} \bigcup_{a=0}^{m} \bigcup_{b=0}^{n} \bigcup_{i=0}^{\frac{d}{2}-1} G_{2i+1}^{(a,b)}.$$

Hence

$$\begin{aligned} s(\alpha^t) &= \sum_{j \in \Omega} \alpha^{tj} \\ &= \sum_{j \in \bigcup_{b=0}^{n} G_0^{(m+1,b)} \bigcup_{a=0}^{m} \bigcup_{b=0}^{n} \bigcup_{i=0}^{\frac{d}{2}-1} G_{2i+1}^{(a,b)}} \alpha^{tj} \\ &= \sum_{b=0}^{n} \sum_{l \in G_0^{(m+1,b)}} \alpha^{tl} + \sum_{a=0}^{m} \sum_{b=0}^{n} \sum_{i=0}^{\frac{d}{2}-1} \sum_{l \in G_0^{(a,b)}} \alpha^{x^{2i+1} tl}. \end{aligned}$$

Since

$$\mathbb{Z}_N = \bigcup_{a=0}^{m} \bigcup_{b=0}^{n} \bigcup_{k=0}^{d-1} p^a q^b G_k \bigcup_{a=0}^{m} p^a q^{n+1} \mathbb{Z}_N^* \bigcup_{b=0}^{n+1} p^{m+1} q^b \mathbb{Z}_N^*,$$

any $t \in \mathbb{Z}_N$ can be written as $t = p^u q^v x^k g^h$ for $0 \leq u \leq m+1$, $0 \leq v \leq n+1$, $0 \leq k \leq d-1$ and $0 \leq h \leq e-1$. Then by Lemma 1 we have

$$\begin{aligned} s(\alpha^t) &= \sum_{b=0}^{n} \sum_{l \in G_0^{(m+1,b)}} \alpha^{p^u q^v x^k g^h l} \\ &\quad + \sum_{a=0}^{m} \sum_{b=0}^{n} \sum_{i=0}^{\frac{d}{2}-1} \sum_{l \in G_0^{(a,b)}} \alpha^{x^{2i+1} p^u q^v x^k g^h l} \\ &= \sum_{b=0}^{n} \sum_{l \in G_0^{(m+1,b)}} \alpha^{p^u q^v l} \\ &\quad + \sum_{a=0}^{m} \sum_{b=0}^{n} \sum_{i=0}^{\frac{d}{2}-1} \sum_{l \in G_0^{(a,b)}} \alpha^{x^{2i+1+k} p^u q^v l} \\ &= \sum_{\substack{b=0 \\ b=n-v}}^{n} 1 + \sum_{\substack{a=0 \\ a=m-u}}^{m} \sum_{\substack{b=0 \\ b=n-v}}^{n} \sum_{i=0}^{\frac{d}{2}-1} \sum_{l \in G_0^{(m,n)}} \alpha^{x^{2i+1+k} l} \\ &\quad + \frac{q-1}{d} \sum_{\substack{a=0 \\ a=m-u}}^{m} \sum_{\substack{b=0 \\ b>n-v}}^{n} \sum_{i=0}^{\frac{d}{2}-1} 1 \\ &\quad + \frac{p-1}{d} \sum_{\substack{a=0 \\ a>m-u}}^{m} \sum_{\substack{b=0 \\ b=n-v}}^{n} \sum_{i=0}^{\frac{d}{2}-1} 1. \end{aligned} \tag{5}$$

Case I: For $0 \leq u \leq m$ and $0 \leq v \leq n$, from Equation (5) we have

$$\begin{aligned} s(\alpha^t) &= 1 + \sum_{i=0}^{\frac{d}{2}-1} \sum_{l \in G_0^{(m,n)}} \alpha^{x^{2i+1+k} l} + \frac{q-1}{2} \sum_{\substack{b=0 \\ b>n-v}}^{n} 1 \\ &\quad + \frac{p-1}{2} \sum_{\substack{a=0 \\ a>m-u}}^{m} 1 \\ &= 1 + \sum_{i=0}^{\frac{d}{2}-1} \sum_{r=0}^{\frac{(p-1)(q-1)}{d}-1} \alpha^{x^{2i+1+k} p^m q^n g^r} \\ &\quad + \frac{v(q-1)}{2} + \frac{u(p-1)}{2} \\ &= 1 + \sum_{i=0}^{\frac{d}{2}-1} \sum_{l \in D_{(2i+1+k) \bmod d}} \alpha^{p^m q^n l} \\ &\quad + \frac{v(q-1)}{2} + \frac{u(p-1)}{2}, \end{aligned}$$

which implies that

$$\begin{aligned} s(\alpha^t) = 0 \iff & \sum_{i=0}^{\frac{d}{2}-1} \sum_{l \in D_{(2i+1+k) \bmod d}} \alpha^{p^m q^n l} \\ & \equiv 1 + \frac{v(q-1)}{2} + \frac{u(p-1)}{2} \pmod 2. \end{aligned}$$

Hence we get

$$\left|\left\{t : s(\alpha^t) = 0,\ t \in \bigcup_{a=0}^{m}\bigcup_{b=0}^{n}\bigcup_{k=0}^{d-1} p^a q^b G_k\right\}\right| = \sum_{a=0}^{m}\sum_{b=0}^{n} A_{p,q,a,b}\frac{p^{m-a}q^{n-b}(p-1)(q-1)}{d},$$

where

$$A_{p,q,a,b} = \begin{cases} E, & \text{if } 1+\frac{b(q-1)}{2}+\frac{a(p-1)}{2} \equiv 0 \pmod 2,\\ F, & \text{if } 1+\frac{b(q-1)}{2}+\frac{a(p-1)}{2} \equiv 1 \pmod 2,\end{cases}$$

for

$$E = |\{k : \sum_{i=0}^{\frac{d}{2}-1}\sum_{l\in D_{(2i+1+k) \bmod d}} \alpha^{p^m q^n l} = 0, k = 0, \cdots, d-1\}|,$$

$$F = |\{k : \sum_{i=0}^{\frac{d}{2}-1}\sum_{l\in D_{(2i+1+k) \bmod d}} \alpha^{p^m q^n l} = 1, k = 0, \cdots, d-1\}|.$$

On the other hand, since $s(X) \in \mathbb{F}_2[X]$, it follows that $s(\alpha^t)^2 = s(\alpha^{2t})$. If $2 \in \cup_{j=0}^{\frac{d}{2}-1} D_{2j}$, then by Lemma 2 we have

$$\begin{aligned} s(\alpha^t)^2 &= s(\alpha^{2t}) \\ &= 1 + \sum_{i=0}^{\frac{d}{2}-1}\sum_{l\in D_{(2i+1+k) \bmod d}} \alpha^{2p^m q^n l} \\ &\quad + \frac{v(q-1)}{2} + \frac{u(p-1)}{2} \\ &= s(\alpha^t). \end{aligned}$$

In this case $s(\alpha^t) \in \{0, 1\}$.

Note that $\alpha^{p^m q^n}$ is a primitive pq-th root of unity in an extension field of $\mathbb{F}_2$ and

$$\sum_{i=0}^{\frac{d}{2}-1}\sum_{l\in D_{2i}} \alpha^{p^m q^n l} + \sum_{i=0}^{\frac{d}{2}-1}\sum_{l\in D_{2i+1}} \alpha^{p^m q^n l} = 1.$$

If $2 \in \bigcup_{j=0}^{\frac{d}{2}-1} D_{2j+1}$, then by Lemma 2 we have

$$\begin{aligned} s(\alpha^t)^2 &= s(\alpha^{2t}) \\ &= 1 + \sum_{i=0}^{\frac{d}{2}-1}\sum_{l\in D_{(2i+1+k) \bmod d}} \alpha^{2p^m q^n l} \\ &\quad + \frac{v(q-1)}{2} + \frac{u(p-1)}{2} \\ &= s(\alpha^t) + 1. \end{aligned}$$

Thus $s(\alpha^t) \notin \{0, 1\}$.

By Lemma 3, if $p \equiv \pm 1 \pmod 8$, $q \equiv \pm 1 \pmod 8$ or $p \equiv \pm 3 \pmod 8$, $q \equiv \pm 3 \pmod 8$, then $E = F = \frac{d}{2}$ and hence $A_{p,q,a,b} = \frac{d}{2}$ and

$$\begin{aligned} &\left|\left\{t : s(\alpha^t) = 0,\ t \in \bigcup_{a=0}^{m}\bigcup_{b=0}^{n}\bigcup_{k=0}^{d-1} p^a q^b D G_k\right\}\right| \\ &= \sum_{a=0}^{m}\sum_{b=0}^{n} \frac{p^a q^b (p-1)(q-1)}{2} \\ &= \frac{(p^{m+1}-1)(q^{n+1}-1)}{2}. \end{aligned}$$

If $p \equiv \pm 1 \pmod 8$, $q \equiv \pm 3 \pmod 8$ or $p \equiv \pm 3 \pmod 8$, $q \equiv \pm 1 \pmod 8$, then $E = F = 0$ and hence $A_{p,q,a,b} = 0$ and

$$\left|\left\{t : s(\alpha^t) = 0,\ t \in \bigcup_{a=0}^{m}\bigcup_{b=0}^{n}\bigcup_{k=0}^{d-1} p^a q^b G_k\right\}\right| = 0.$$

Case II: For $u = m+1$ and $0 \le v \le n$, from Equation (5) we have

$$s(\alpha^t) = 1 + \frac{(m+1)(p-1)}{2},$$

and

$$s(\alpha^t) = 0 \Longleftrightarrow 1 + \frac{(m+1)(p-1)}{2} \equiv 0 \pmod 2.$$

So we conclude that

$$\begin{aligned} &\left|\left\{t : s(\alpha^t) = 0,\ t \in \bigcup_{b=0}^{n} p^{m+1} q^b \mathbb{Z}_N^*\right\}\right| \\ &= A_{p,m}\sum_{b=0}^{n} q^{n-b}(q-1) = A_{p,m}(q^{n+1}-1). \end{aligned}$$

Case III: For $0 \le u \le m$ and $v = n+1$, from Equation (5) we have

$$s(\alpha^t) = \frac{(n+1)(q-1)}{2},$$

from which we obtain

$$s(\alpha^t) = 0 \Longleftrightarrow \frac{(n+1)(q-1)}{2} \equiv 0 \pmod 2.$$

Therefore

$$\begin{aligned} &\left|\left\{t : s(\alpha^t) = 0,\ t \in \bigcup_{a=0}^{m} p^a q^{n+1} \mathbb{Z}_N^*\right\}\right| \\ &= A_{q,n}\sum_{a=0}^{m} p^{m-a}(p-1) = A_{q,n}(p^{m+1}-1). \end{aligned}$$

Case IV: For $u = m+1$ and $v = n+1$, from Equation (5) we have

$$s(\alpha^t) = s(\alpha^0) = s(1) = 0.$$

Putting everything together, we complete the proof of Theorem 1. □

Remark 1. *It is not hard to show that*

$$A_{q,0} = \begin{cases} 1, & \text{if } q \equiv 1 \pmod 4, \\ 0, & \text{if } q \equiv 3 \pmod 4, \end{cases}$$

$$A_{p,0} = \begin{cases} 1, & \text{if } p \equiv 3 \pmod 4, \\ 0, & \text{if } p \equiv 1 \pmod 4. \end{cases}$$

It is obvious that our results are entirely consistent with those in [13].

3 Autocorrelations

Let $\left(\frac{\cdot}{p}\right)$ denote the Legendre symbol modulo p, and $\left(\frac{\cdot}{q}\right)$ the Legendre symbol modulo q. In this section, we determine the exact values of autocorrelation of s^∞.

Theorem 2. *Let s^∞ be the N-periodic binary sequence defined as in Equation (1) with $I_{a,b}$ in Equation (3) for defining Ω. For $1 \le w \le p^{m+1}q^{n+1} - 1$ with $\left(w, p^{m+1}q^{n+1}\right) = p^{a_0}q^{b_0}$, the autocorrelation of s^∞ satisfies*

$$C_s(w) = \begin{cases} p^m q^n + \left(1 - (-1)^{\frac{p+q}{2}}\right) \cdot \left(\frac{w}{p}\right)\left(\frac{w}{q}\right) - 2, \\ \qquad \text{if } a_0 = 0,\ b_0 = 0, \\ q^n(1-p^{m+1}) + q^{n+1} - 4, \\ \qquad \text{if } a_0 = m+1,\ b_0 = 0, \\ p^m(1-q^{n+1}) + p^{m+1}, \\ \qquad \text{if } a_0 = 0,\ b_0 = n+1, \\ p^m q^{n-b_0} + p^m q^{n-b_0+1}(1-q^{b_0}) \\ \quad + \left(1 - (-1)^{\frac{p+q}{2}}\right) \cdot \left(\frac{\frac{w}{q^{b_0}}}{p}\right)\left(\frac{\frac{w}{q^{b_0}}}{q}\right) - 2, \\ \qquad \text{if } a_0 = 0,\ 1 \le b_0 \le n, \\ p^{m-a_0} q^n + q^n p^{m-a_0+1}(1-p^{a_0}) \\ \quad + \left(1 - (-1)^{\frac{p+q}{2}}\right) \cdot \left(\frac{\frac{w}{p^{a_0}}}{p}\right)\left(\frac{\frac{w}{p^{a_0}}}{q}\right) - 2, \\ \qquad \text{if } 1 \le a_0 \le m,\ b_0 = 0, \\ q^{n-b_0}(1-p^{m+1}) \\ \quad + q^{n-b_0+1}(1-q^{b_0})(1-p^{m+1}) + q^{n+1} - 4, \\ \qquad \text{if } a_0 = m+1,\ 1 \le b_0 \le n, \\ p^{m-a_0}(1-q^{n+1}) \\ \quad + p^{m-a_0+1}(1-p^{a_0})(1-q^{n+1}) + p^{m+1}, \\ \qquad \text{if } 1 \le a_0 \le m,\ b_0 = n+1, \\ p^{m-a_0} q^{n-b_0} + p^{m-a_0} q^{n-b_0+1}(1-q^{b_0}) \\ \quad + q^{n-b_0} p^{m-a_0+1}(1-p^{a_0}) \\ \quad + p^{m-a_0+1} q^{n-b_0+1}(1-p^{a_0})(1-q^{b_0}) \\ \quad + \left(1 - (-1)^{\frac{p+q}{2}}\right) \cdot \left(\frac{\frac{w}{p^{a_0}q^{b_0}}}{p}\right)\left(\frac{\frac{w}{p^{a_0}q^{b_0}}}{q}\right) - 2, \\ \qquad \text{if } 1 \le a_0 \le m,\ 1 \le b_0 \le n. \end{cases}$$

Remark 2. *Theorem 2 shows that the autocorrelation values of s^∞ are quite good.*

3.1 Certain Identities Involving Character Sums

To prove Theorem 2, we need the following identities.

Lemma 4. *Assume that $1 \le w \le p^{m+1}q^{n+1} - 1$. Then we have*

$$\begin{aligned} &\sum_{\substack{k=0 \\ q^{n+1}\mid k \\ q^{n+1}\mid k+w}}^{p^{m+1}q^{n+1}-1} 1 - \sum_{\substack{k=1 \\ p^{m+1}\mid k \\ q^{n+1}\mid k+w}}^{p^{m+1}q^{n+1}-1} 1 \\ &- \sum_{\substack{k=0 \\ q^{n+1}\mid k \\ p^{m+1}\mid k+w \\ p^{m+1}q^{n+1}\nmid k+w}}^{p^{m+1}q^{n+1}-1} 1 + \sum_{\substack{k=1 \\ p^{m+1}\mid k \\ p^{m+1}\mid k+w \\ p^{m+1}q^{n+1}\nmid k+w}}^{p^{m+1}q^{n+1}-1} 1 \\ &= \begin{cases} q^{n+1} - 4, & \text{if } p^{m+1} \mid w, \\ p^{m+1}, & \text{if } q^{n+1} \mid w, \\ -2, & \text{if } p^{m+1} \nmid w \text{ and } q^{n+1} \nmid w. \end{cases} \end{aligned}$$

Proof. It is not hard to show that

$$\sum_{\substack{k=0 \\ q^{n+1}\mid k \\ q^{n+1}\mid k+w}}^{p^{m+1}q^{n+1}-1} 1 = \begin{cases} p^{m+1}, & q^{n+1} \mid w, \\ 0, & q^{n+1} \nmid w, \end{cases}$$

$$\sum_{\substack{k=1 \\ p^{m+1}\mid k \\ q^{n+1}\mid k+w}}^{p^{m+1}q^{n+1}-1} 1 = \begin{cases} 0, & q^{n+1} \mid w, \\ 1, & q^{n+1} \nmid w, \end{cases}$$

$$\sum_{\substack{k=0 \\ q^{n+1}\mid k \\ p^{m+1}\mid k+w \\ p^{m+1}q^{n+1}\nmid k+w}}^{p^{m+1}q^{n+1}-1} 1 = \begin{cases} 0, & q^{n+1} \mid w, \\ 1, & q^{n+1} \nmid w, \end{cases}$$

$$\sum_{\substack{k=1 \\ p^{m+1}\mid k \\ p^{m+1}\mid k+w \\ p^{m+1}q^{n+1}\nmid k+w}}^{p^{m+1}q^{n+1}-1} 1 = \begin{cases} q^{n+1} - 2, & p^{m+1} \mid w, \\ 0, & p^{m+1} \nmid w. \end{cases}$$

Lemma 4 is thus established. □

Lemma 5. *Assume that $1 \le w \le p^{m+1}q^{n+1} - 1$ with $\left(w, p^{m+1}q^{n+1}\right) = p^{a_0}q^{b_0}$. Then we have*

$$\begin{aligned} &\sum_{a=0}^{m}\sum_{b=0}^{n} \sum_{\substack{k=0 \\ (k, p^{m+1}q^{n+1}) = p^a q^b \\ q^{n+1}\mid k+w}}^{p^{m+1}q^{n+1}-1} \left(\frac{\frac{k}{p^a q^b}}{p}\right)\left(\frac{\frac{k}{p^a q^b}}{q}\right) \\ &- \sum_{a=0}^{m}\sum_{b=0}^{n} \sum_{\substack{k=0 \\ (k, p^{m+1}q^{n+1}) = p^a q^b \\ p^{m+1}\mid k+w \\ p^{m+1}q^{n+1}\nmid k+w}}^{p^{m+1}q^{n+1}-1} \left(\frac{\frac{k}{p^a q^b}}{p}\right)\left(\frac{\frac{k}{p^a q^b}}{q}\right) \end{aligned}$$

$$+\sum_{a=0}^{m}\sum_{b=0}^{n}\sum_{\substack{k=0\\ q^{n+1}|k\\ (k+w,p^{m+1}q^{n+1})=p^aq^b}}^{p^{m+1}q^{n+1}-1}\left(\frac{\frac{k+w}{p^aq^b}}{p}\right)\left(\frac{\frac{k+w}{p^aq^b}}{q}\right)$$

$$-\sum_{a=0}^{m}\sum_{b=0}^{n}\sum_{\substack{k=1\\ p^{m+1}|k\\ (k+w,p^{m+1}q^{n+1})=p^aq^b}}^{p^{m+1}q^{n+1}-1}\left(\frac{\frac{k+w}{p^aq^b}}{p}\right)\left(\frac{\frac{k+w}{p^aq^b}}{q}\right)$$

$$=\begin{cases}\left(1-(-1)^{\frac{p+q}{2}}\right)\cdot\left(\frac{\frac{w}{p^{a_0}q^{b_0}}}{p}\right)\left(\frac{\frac{w}{p^{a_0}q^{b_0}}}{q}\right), & \\ \qquad if\ \ a_0\le m,\ b_0\le n, & \\ 0, & \\ \qquad if\ \ a_0=m+1\ \ or\ \ b_0=n+1. & \end{cases}$$

Proof. By the properties of the Legendre symbols and complete residue systems we get

$$\begin{aligned}&\sum_{a=0}^{m}\sum_{b=0}^{n}\sum_{\substack{k=0\\ (k,p^{m+1}q^{n+1})=p^aq^b\\ q^{n+1}|k+w}}^{p^{m+1}q^{n+1}-1}\left(\frac{\frac{k}{p^aq^b}}{p}\right)\left(\frac{\frac{k}{p^aq^b}}{q}\right)\\ &=\sum_{a=0}^{m}\sum_{b=0}^{n}\sum_{\substack{k=0\\ (k,pq)=1\\ q^{n+1}|p^aq^bk+w}}^{p^{m+1-a}q^{n+1-b}-1}\left(\frac{k}{p}\right)\left(\frac{k}{q}\right)\\ &=\sum_{a=0}^{m}\sum_{b=0}^{n}\sum_{\substack{k=0\\ q^{n+1}|p^aq^bk+w}}^{p^{m+1-a}q^{n+1-b}-1}\left(\frac{k}{p}\right)\left(\frac{k}{q}\right)\\ &=\sum_{a=0}^{m}\sum_{b=0}^{n}\sum_{\substack{k_1=0\\ q^{n+1}|p^{m+1}q^bk_1+w}}^{q^{n+1-b}-1}\left(\frac{k_1p^{m+1-a}}{q}\right)\\ &\quad\times\sum_{k_2=0}^{p^{m+1-a}-1}\left(\frac{k_2q^{n+1-b}}{p}\right)\\ &=0,\end{aligned}\tag{6}$$

and

$$\begin{aligned}&\sum_{a=0}^{m}\sum_{b=0}^{n}\sum_{\substack{k=0\\ (k,p^{m+1}q^{n+1})=p^aq^b\\ p^{m+1}|k+w\\ p^{m+1}q^{n+1}\nmid k+w}}^{p^{m+1}q^{n+1}-1}\left(\frac{\frac{k}{p^aq^b}}{p}\right)\left(\frac{\frac{k}{p^aq^b}}{q}\right)\\ &=\sum_{a=0}^{m}\sum_{b=0}^{n}\sum_{\substack{k=0\\ (k,p^{m+1}q^{n+1})=p^aq^b\\ p^{m+1}|k+w}}^{p^{m+1}q^{n+1}-1}\left(\frac{\frac{k}{p^aq^b}}{p}\right)\left(\frac{\frac{k}{p^aq^b}}{q}\right)\\ &\quad-\sum_{a=0}^{m}\sum_{b=0}^{n}\sum_{\substack{k=0\\ (k,p^{m+1}q^{n+1})=p^aq^b\\ p^{m+1}|k+w\\ p^{m+1}q^{n+1}|k+w}}^{p^{m+1}q^{n+1}-1}\left(\frac{\frac{k}{p^aq^b}}{p}\right)\left(\frac{\frac{k}{p^aq^b}}{q}\right)\\ &=\begin{cases}\sum_{a=0}^{m}\sum_{b=0}^{n}\sum_{\substack{k=0\\ p^{m+1}|p^aq^bk+w}}^{p^{m+1-a}q^{n+1-b}-1}\left(\frac{k}{p}\right)\left(\frac{k}{q}\right)\\ \quad-\left(\frac{\frac{-w}{p^{a_0}q^{b_0}}}{p}\right)\left(\frac{\frac{-w}{p^{a_0}q^{b_0}}}{q}\right), & \\ \qquad \text{if}\ \ a_0\le m,\ b_0\le n, & \\ \sum_{a=0}^{m}\sum_{b=0}^{n}\sum_{\substack{k=0\\ p^{m+1}|p^aq^bk+w}}^{p^{m+1-a}q^{n+1-b}-1}\left(\frac{k}{p}\right)\left(\frac{k}{q}\right), & \\ \qquad \text{if}\ \ a_0=m+1\ \ \text{or}\ \ b_0=n+1, & \end{cases}\\ &=\begin{cases}-\left(\frac{\frac{-w}{p^{a_0}q^{b_0}}}{p}\right)\left(\frac{\frac{-w}{p^{a_0}q^{b_0}}}{q}\right), & \\ \qquad \text{if}\ \ a_0\le m,\ b_0\le n, & \\ 0, & \\ \qquad \text{if}\ \ a_0=m+1\ \ \text{or}\ \ b_0=n+1. & \end{cases}\end{aligned}\tag{7}$$

Similarly, we have

$$\sum_{a=0}^{m}\sum_{b=0}^{n}\sum_{\substack{k=0\\ q^{n+1}|k\\ (k+w,p^{m+1}q^{n+1})=p^aq^b}}^{p^{m+1}q^{n+1}-1}\left(\frac{\frac{k+w}{p^aq^b}}{p}\right)\left(\frac{\frac{k+w}{p^aq^b}}{q}\right)=0,\tag{8}$$

and

$$\begin{aligned}&\sum_{a=0}^{m}\sum_{b=0}^{n}\sum_{\substack{k=1\\ p^{m+1}|k\\ (k+w,p^{m+1}q^{n+1})=p^aq^b}}^{p^{m+1}q^{n+1}-1}\left(\frac{\frac{k+w}{p^aq^b}}{p}\right)\left(\frac{\frac{k+w}{p^aq^b}}{q}\right)\\ &=\begin{cases}-\left(\frac{\frac{w}{p^{a_0}q^{b_0}}}{p}\right)\left(\frac{\frac{w}{p^{a_0}q^{b_0}}}{q}\right), & \\ \qquad \text{if}\ \ a_0\le m,\ b_0\le n, & \\ 0, & \\ \qquad \text{if}\ \ a_0=m+1\ \ \text{or}\ \ b_0=n+1. & \end{cases}\end{aligned}\tag{9}$$

From Equation (6), Equation (7), Equation (8), and Equation (9), we can get the conclusion of Lemma 5 directly. □

Lemma 6. *Assume that* $1\le w\le p^{m+1}q^{n+1}-1$ *with* $\left(w,p^{m+1}q^{n+1}\right)=p^{a_0}q^{b_0}$. *Then we have*

$$\sum_{a_1=0}^{m}\sum_{b_1=0}^{n}\sum_{a_2=0}^{m}\sum_{b_2=0}^{n}\sum_{\substack{k=0\\ (k,p^{m+1}q^{n+1})=p^{a_1}q^{b_1}\\ (k+w,p^{m+1}q^{n+1})=p^{a_2}q^{b_2}}}^{p^{m+1}q^{n+1}-1}\left(\frac{\frac{k}{p^{a_1}q^{b_1}}}{p}\right)\times\left(\frac{\frac{k}{p^{a_1}q^{b_1}}}{q}\right)\left(\frac{\frac{k+w}{p^{a_2}q^{b_2}}}{p}\right)\left(\frac{\frac{k+w}{p^{a_2}q^{b_2}}}{q}\right)$$

$$= \begin{cases} p^m q^n, \\ \qquad if \;\; a_0 = 0, \;\; b_0 = 0, \\ q^n(1-p^{m+1}), \\ \qquad if \;\; a_0 = m+1, \;\; b_0 = 0, \\ p^m(1-q^{n+1}), \\ \qquad if \;\; a_0 = 0, \;\; b_0 = n+1, \\ p^m q^{n-b_0} + p^m q^{n-b_0+1}(1-q^{b_0}), \\ \qquad if \;\; a_0 = 0, \;\; 1 \le b_0 \le n, \\ p^{m-a_0} q^n + q^n p^{m-a_0+1}(1-p^{a_0}), \\ \qquad if \;\; 1 \le a_0 \le m, \;\; b_0 = 0, \\ q^{n-b_0}(1-p^{m+1}) + q^{n-b_0+1}(1-q^{b_0})(1-p^{m+1}), \\ \qquad if \;\; a_0 = m+1, \;\; 1 \le b_0 \le n, \\ p^{m-a_0}(1-q^{n+1}) + p^{m-a_0+1}(1-p^{a_0})(1-q^{n+1}), \\ \qquad if \;\; 1 \le a_0 \le m, \;\; b_0 = n+1, \\ p^{m-a_0} q^{n-b_0} + p^{m-a_0} q^{n-b_0+1}(1-q^{b_0}) \\ \quad + q^{n-b_0} p^{m-a_0+1}(1-p^{a_0}) \\ \quad + p^{m-a_0+1} q^{n-b_0+1}(1-p^{a_0})(1-q^{b_0}), \\ \qquad if \;\; 1 \le a_0 \le m, \;\; 1 \le b_0 \le n. \end{cases}$$

Proof. By the properties of character sums, greatest common divisors and complete residue systems we have

$$\sum_{\substack{a_1=0 \\ a_1>a_2}}^{m} \sum_{a_2=0}^{m} \sum_{\substack{b_1=0 \\ b_1>b_2}}^{n} \sum_{b_2=0}^{n} \sum_{\substack{k=0 \\ (k,p^{m+1}q^{n+1})=p^{a_1}q^{b_1} \\ (k+w,p^{m+1}q^{n+1})=p^{a_2}q^{b_2}}}^{p^{m+1}q^{n+1}-1} \left(\frac{\frac{k}{p^{a_1}q^{b_1}}}{p}\right) \times \left(\frac{\frac{k}{p^{a_1}q^{b_1}}}{q}\right)\left(\frac{\frac{k+w}{p^{a_2}q^{b_2}}}{p}\right)\left(\frac{\frac{k+w}{p^{a_2}q^{b_2}}}{q}\right)$$

$$= \sum_{\substack{a_1=0 \\ a_1>a_2}}^{m} \sum_{a_2=0}^{m} \sum_{\substack{b_1=0 \\ b_1>b_2}}^{n} \sum_{b_2=0}^{n} \sum_{k=0}^{p^{m+1-a_1}q^{n+1-b_1}-1} \left(\frac{k}{p}\right) \times \left(\frac{k}{q}\right)\left(\frac{\frac{p^{a_1}q^{b_1}k+w}{p^{a_2}q^{b_2}}}{p}\right)\left(\frac{\frac{p^{a_1}q^{b_1}k+w}{p^{a_2}q^{b_2}}}{q}\right)$$

$$= \sum_{\substack{a_1=0 \\ a_1>a_2}}^{m} \sum_{a_2=0}^{m} \sum_{\substack{b_1=0 \\ b_1>b_2}}^{n} \sum_{b_2=0}^{n} \sum_{k=0}^{p^{m+1-a_1}q^{n+1-b_1}-1} \left(\frac{k}{p}\right) \times \left(\frac{k}{q}\right)\left(\frac{\frac{w}{p^{a_2}q^{b_2}}}{p}\right)\left(\frac{\frac{w}{p^{a_2}q^{b_2}}}{q}\right)$$

$$= \sum_{\substack{a_1=0 \\ a_1>a_2}}^{m} \sum_{a_2=0}^{m} \sum_{\substack{b_1=0 \\ b_1>b_2}}^{n} \sum_{b_2=0}^{n} \sum_{k_1=0}^{q^{n+1-b_1}-1} \sum_{k_2=0}^{p^{m+1-a_1}} \left(\frac{k_2}{p}\right) \times \left(\frac{k_1}{q}\right)\left(\frac{\frac{w}{p^{a_2}q^{b_2}}}{p}\right)\left(\frac{\frac{w}{p^{a_2}q^{b_2}}}{q}\right) = 0.$$

In the same way we obtain

$$\sum_{\substack{a_1=0 \\ a_1>a_2}}^{m} \sum_{a_2=0}^{m} \sum_{\substack{b_1=0 \\ b_1<b_2}}^{n} \sum_{b_2=0}^{n} \sum_{\substack{k=0 \\ (k,p^{m+1}q^{n+1})=p^{a_1}q^{b_1} \\ (k+w,p^{m+1}q^{n+1})=p^{a_2}q^{b_2}}}^{p^{m+1}q^{n+1}-1} \left(\frac{\frac{k}{p^{a_1}q^{b_1}}}{p}\right) \times \left(\frac{\frac{k}{p^{a_1}q^{b_1}}}{q}\right)\left(\frac{\frac{k+w}{p^{a_2}q^{b_2}}}{p}\right)\left(\frac{\frac{k+w}{p^{a_2}q^{b_2}}}{q}\right) = 0,$$

$$\sum_{\substack{a_1=0 \\ a_1<a_2}}^{m} \sum_{a_2=0}^{m} \sum_{\substack{b_1=0 \\ b_1>b_2}}^{n} \sum_{b_2=0}^{n} \sum_{\substack{k=0 \\ (k,p^{m+1}q^{n+1})=p^{a_1}q^{b_1} \\ (k+w,p^{m+1}q^{n+1})=p^{a_2}q^{b_2}}}^{p^{m+1}q^{n+1}-1} \left(\frac{\frac{k}{p^{a_1}q^{b_1}}}{p}\right) \times \left(\frac{\frac{k}{p^{a_1}q^{b_1}}}{q}\right)\left(\frac{\frac{k+w}{p^{a_2}q^{b_2}}}{p}\right)\left(\frac{\frac{k+w}{p^{a_2}q^{b_2}}}{q}\right) = 0,$$

$$\sum_{\substack{a_1=0 \\ a_1<a_2}}^{m} \sum_{a_2=0}^{m} \sum_{\substack{b_1=0 \\ b_1<b_2}}^{n} \sum_{b_2=0}^{n} \sum_{\substack{k=0 \\ (k,p^{m+1}q^{n+1})=p^{a_1}q^{b_1} \\ (k+w,p^{m+1}q^{n+1})=p^{a_2}q^{b_2}}}^{p^{m+1}q^{n+1}-1} \left(\frac{\frac{k}{p^{a_1}q^{b_1}}}{p}\right) \times \left(\frac{\frac{k}{p^{a_1}q^{b_1}}}{q}\right)\left(\frac{\frac{k+w}{p^{a_2}q^{b_2}}}{p}\right)\left(\frac{\frac{k+w}{p^{a_2}q^{b_2}}}{q}\right) = 0,$$

$$\sum_{\substack{a_1=0 \\ a_1<a_2}}^{m} \sum_{a_2=0}^{m} \sum_{\substack{b_1=0 \\ b_1>b_2}}^{n} \sum_{b_2=0}^{n} \sum_{\substack{k=0 \\ (k,p^{m+1}q^{n+1})=p^{a_1}q^{b_1} \\ (k+w,p^{m+1}q^{n+1})=p^{a_2}q^{b_2}}}^{p^{m+1}q^{n+1}-1} \left(\frac{\frac{k}{p^{a_1}q^{b_1}}}{p}\right) \times \left(\frac{\frac{k}{p^{a_1}q^{b_1}}}{q}\right)\left(\frac{\frac{k+w}{p^{a_2}q^{b_2}}}{p}\right)\left(\frac{\frac{k+w}{p^{a_2}q^{b_2}}}{q}\right) = 0,$$

$$\sum_{\substack{a_1=0 \\ a_1=a_2}}^{m} \sum_{a_2=0}^{m} \sum_{\substack{b_1=0 \\ b_1<b_2}}^{n} \sum_{b_2=0}^{n} \sum_{\substack{k=0 \\ (k,p^{m+1}q^{n+1})=p^{a_1}q^{b_1} \\ (k+w,p^{m+1}q^{n+1})=p^{a_2}q^{b_2}}}^{p^{m+1}q^{n+1}-1} \left(\frac{\frac{k}{p^{a_1}q^{b_1}}}{p}\right) \times \left(\frac{\frac{k}{p^{a_1}q^{b_1}}}{q}\right)\left(\frac{\frac{k+w}{p^{a_2}q^{b_2}}}{p}\right)\left(\frac{\frac{k+w}{p^{a_2}q^{b_2}}}{q}\right) = 0,$$

$$\sum_{\substack{a_1=0 \\ a_1=a_2}}^{m} \sum_{a_2=0}^{m} \sum_{\substack{b_1=0 \\ b_1>b_2}}^{n} \sum_{b_2=0}^{n} \sum_{\substack{k=0 \\ (k,p^{m+1}q^{n+1})=p^{a_1}q^{b_1} \\ (k+w,p^{m+1}q^{n+1})=p^{a_2}q^{b_2}}}^{p^{m+1}q^{n+1}-1} \left(\frac{\frac{k}{p^{a_1}q^{b_1}}}{p}\right) \times \left(\frac{\frac{k}{p^{a_1}q^{b_1}}}{q}\right)\left(\frac{\frac{k+w}{p^{a_2}q^{b_2}}}{p}\right)\left(\frac{\frac{k+w}{p^{a_2}q^{b_2}}}{q}\right) = 0.$$

Summarizing the results of the eight cases we obtain

$$\sum_{a_1=0}^{m} \sum_{b_1=0}^{n} \sum_{a_2=0}^{m} \sum_{b_2=0}^{n} \sum_{\substack{k=0 \\ (k,p^{m+1}q^{n+1})=p^{a_1}q^{b_1} \\ (k+w,p^{m+1}q^{n+1})=p^{a_2}q^{b_2}}}^{p^{m+1}q^{n+1}-1} \left(\frac{\frac{k}{p^{a_1}q^{b_1}}}{p}\right)$$

$$\times\left(\frac{\frac{k}{p^{a_1}q^{b_1}}}{q}\right)\left(\frac{\frac{k+w}{p^{a_2}q^{b_2}}}{p}\right)\left(\frac{\frac{k+w}{p^{a_2}q^{b_2}}}{q}\right)$$

$$=\sum_{a_1=0}^{m}\sum_{\substack{a_2=0\\a_1=a_2}}^{m}\sum_{b_1=0}^{n}\sum_{\substack{b_2=0\\b_1=b_2}}^{n}\sum_{\substack{k=0\\(k,p^{m+1}q^{n+1})=p^{a_1}q^{b_1}\\(k+w,p^{m+1}q^{n+1})=p^{a_2}q^{b_2}}}^{p^{m+1}q^{n+1}-1}\left(\frac{\frac{k}{p^{a_1}q^{b_1}}}{p}\right)$$

$$\times\left(\frac{\frac{k}{p^{a_1}q^{b_1}}}{q}\right)\left(\frac{\frac{k+w}{p^{a_2}q^{b_2}}}{p}\right)\left(\frac{\frac{k+w}{p^{a_2}q^{b_2}}}{q}\right)$$

$$=\sum_{a=0}^{m}\sum_{b=0}^{n}\sum_{\substack{k=0\\(p^aq^bk+w,p^{m+1}q^{n+1})=p^aq^b}}^{p^{m+1-a}q^{n+1-b}-1}\left(\frac{k}{p}\right)\left(\frac{k}{q}\right)$$

$$\times\left(\frac{k+\frac{w}{p^aq^b}}{p}\right)\left(\frac{k+\frac{w}{p^aq^b}}{q}\right)$$

$$=\sum_{a=0}^{m}\sum_{b=0}^{n}\sum_{\substack{k_1=0\\q^b\|k_1p^{m+1}q^b+w}}^{q^{n+1-b}-1}\sum_{\substack{k_2=0\\p^a\|k_2p^aq^{n+1}+w}}^{p^{m+1-a}-1}\left(\frac{k_2q^{n+1-b}}{p}\right)$$

$$\times\left(\frac{k_1p^{m+1-a}}{q}\right)\left(\frac{k_2q^{n+1-b}+\frac{w}{p^aq^b}}{p}\right)$$

$$\times\left(\frac{k_1p^{m+1-a}+\frac{w}{p^aq^b}}{q}\right)$$

$$=\sum_{a=0}^{m}\sum_{b=0}^{n}\sum_{\substack{k_1=0\\p\nmid k_1p^{m+1}+\frac{w}{q^b}}}^{q^{n+1-b}-1}\sum_{\substack{k_2=0\\p\nmid k_2q^{n+1}+\frac{w}{p^a}}}^{p^{m+1-a}-1}\left(\frac{k_2q^{n+1-b}}{p}\right)$$

$$\times\left(\frac{k_1p^{m+1-a}}{q}\right)\left(\frac{k_2q^{n+1-b}+\frac{w}{p^aq^b}}{p}\right)$$

$$\times\left(\frac{k_1p^{m+1-a}+\frac{w}{p^aq^b}}{q}\right)$$

$$+\sum_{a=0}^{m}\sum_{\substack{b=0\\p^a\|w,q^{b+1}|w}}^{n}\sum_{\substack{k_1=0\\(k_1,q)=1}}^{q^{n+1-b}-1}\sum_{\substack{k_2=0\\p\nmid k_2q^{n+1}+\frac{w}{p^a}}}^{p^{m+1-a}-1}\left(\frac{k_2q^{n+1-b}}{p}\right)$$

$$\times\left(\frac{k_1p^{m+1-a}}{q}\right)\left(\frac{k_2q^{n+1-b}+\frac{w}{p^aq^b}}{p}\right)\left(\frac{k_1p^{m+1-a}}{q}\right)$$

$$+\sum_{\substack{a=0\\p^{a+1}|w,q^b\|w}}^{m}\sum_{b=0}^{n}\sum_{\substack{k_1=0\\q\nmid k_1p^{m+1}+\frac{w}{q^b}}}^{q^{n+1-b}-1}\sum_{\substack{k_2=0\\(k_2,p)=1}}^{p^{m+1-a}-1}\left(\frac{k_2q^{n+1-b}}{p}\right)$$

$$\times\left(\frac{k_1p^{m+1-a}}{q}\right)\left(\frac{k_2q^{n+1-b}}{p}\right)\left(\frac{k_1p^{m+1-a}+\frac{w}{p^aq^b}}{q}\right)$$

$$+\sum_{\substack{a=0\\p^{a+1}|w,q^{b+1}|w}}^{m}\sum_{b=0}^{n}\sum_{\substack{k_1=0\\(k_1,q)=1}}^{q^{n+1-b}-1}\sum_{\substack{k_2=0\\(k_2,p)=1}}^{p^{m+1-a}-1}\left(\frac{k_2q^{n+1-b}}{p}\right)$$

$$\times\left(\frac{k_1p^{m+1-a}}{q}\right)\left(\frac{k_2q^{n+1-b}}{p}\right)\left(\frac{k_1p^{m+1-a}}{q}\right)$$

$$=\sum_{a=0}^{m}\sum_{\substack{b=0\\p^a\|w,q^b\|w}}^{n}\sum_{k_1=0}^{q^{n+1-b}-1}\left(\frac{k_1}{q}\right)\left(\frac{k_1+\frac{w}{p^aq^b}}{q}\right)$$

$$\times\sum_{k_2=0}^{p^{m+1-a}-1}\left(\frac{k_2}{p}\right)\left(\frac{k_2+\frac{w}{p^aq^b}}{p}\right)$$

$$+\sum_{a=0}^{m}\sum_{\substack{b=0\\p^a\|w,q^{b+1}|w}}^{n}q^{n-b}(q-1)$$

$$\times\sum_{k_2=0}^{p^{m+1-a}-1}\left(\frac{k_2}{p}\right)\left(\frac{k_2+\frac{w}{p^aq^b}}{p}\right)$$

$$+\sum_{\substack{a=0\\p^{a+1}|w,q^b\|w}}^{m}\sum_{b=0}^{n}p^{m-a}(p-1)$$

$$\times\sum_{k_1=0}^{q^{n+1-b}-1}\left(\frac{k_1}{q}\right)\left(\frac{k_1+\frac{w}{p^aq^b}}{q}\right)$$

$$+\sum_{\substack{a=0\\p^{a+1}|w,q^{b+1}|w}}^{m}\sum_{b=0}^{n}p^{m-a}q^{n-b}(p-1)(q-1)$$

$$=\sum_{\substack{a=0\\p^a\|w,q^b\|w}}^{m}\sum_{b=0}^{n}(-q^{n-b})\cdot(-p^{m-a})$$

$$+\sum_{a=0}^{m}\sum_{\substack{b=0\\p^a\|w,q^{b+1}|w}}^{n}q^{n-b}(q-1)\cdot(-p^{m-a})$$

$$+\sum_{\substack{a=0\\p^{a+1}|w,q^b\|w}}^{m}\sum_{b=0}^{n}p^{m-a}(p-1)\cdot(-q^{n-b})$$

$$+\sum_{\substack{a=0\\p^{a+1}|w,q^{b+1}|w}}^{m}\sum_{b=0}^{n}p^{m-a}q^{n-b}(p-1)(q-1)$$

$$=\begin{cases}p^mq^n,\\ \qquad\text{if } a_0=0,\ b_0=0,\\ q^n(1-p^{m+1}),\\ \qquad\text{if } a_0=m+1,\ b_0=0,\\ p^m(1-q^{n+1}),\\ \qquad\text{if } a_0=0,\ b_0=n+1,\\ p^mq^{n-b_0}+p^mq^{n-b_0+1}(1-q^{b_0}),\\ \qquad\text{if } a_0=0,\ 1\le b_0\le n,\\ p^{m-a_0}q^n+q^np^{m-a_0+1}(1-p^{a_0}),\\ \qquad\text{if } 1\le a_0\le m,\ b_0=0,\\ q^{n-b_0}(1-p^{m+1})+q^{n-b_0+1}(1-q^{b_0})(1-p^{m+1}),\\ \qquad\text{if } a_0=m+1,\ 1\le b_0\le n,\\ p^{m-a_0}(1-q^{n+1})+p^{m-a_0+1}(1-p^{a_0})(1-q^{n+1}),\\ \qquad\text{if } 1\le a_0\le m,\ b_0=n+1,\\ p^{m-a_0}q^{n-b_0}+p^{m-a_0}q^{n-b_0+1}(1-q^{b_0})\\ \quad+q^{n-b_0}p^{m-a_0+1}(1-p^{a_0})\\ \quad+p^{m-a_0+1}q^{n-b_0+1}(1-p^{a_0})(1-q^{b_0}),\\ \qquad\text{if } 1\le a_0\le m,\ 1\le b_0\le n.\end{cases}$$

□

3.2 Proof of Theorem 2

For integer k, suppose that $\gcd(k,N)=p^aq^b$, $0\le a\le m$, $0\le b\le n$. Write $k=p^aq^bk'$, where $\gcd(k',N)=1$. Note that $\Omega=\bigcup_{a=0}^{m+1}\bigcup_{b=0}^{n+1}\bigcup_{i\in I_{a,b}}G_i^{(a,b)}$, we have

$$\begin{aligned}
k\in\Omega &\iff \text{there exists } i\in I_{a,b} \text{ such that } k\in p^aq^bG_i\\
&\iff \text{there exists } i\in I_{a,b} \text{ such that } k'\in G_i\\
&\iff \text{there exist } i\in I_{a,b},\ 0\le s\le e-1\\
&\qquad \text{such that } k'\equiv g^sx^i \pmod N\\
&\iff \frac{1}{\phi(N)}\sum_{i\in I_{a,b}}\sum_{s=0}^{e-1}\sum_{\chi \bmod N}\chi(k')\overline{\chi}(g^sx^i)=1\\
&\iff \frac{1}{d}\sum_{\substack{\chi \bmod N\\ \chi(g)=1}}\chi(k')\sum_{i\in I_{a,b}}\overline{\chi}(x^i)=1,
\end{aligned}$$

where $\sum_{\chi \bmod N}$ denotes the summation of all the multiplicative characters χ modulo N. Hence,

$$(-1)^{s_k}=-\frac{2}{d}\sum_{\substack{\chi \bmod N\\ \chi(g)=1\\ \chi\neq\chi_0}}\left(\sum_{i\in I_{a,b}}\overline{\chi}(x^i)\right)\chi(k'). \tag{10}$$

Every character $\chi \bmod N$ can be factored in the form $\chi=\chi_1\chi_2$, where χ_1 is a character mod p^{m+1} and χ_2 is a character mod q^{n+1}. Therefore we have

$$\begin{aligned}
&\sum_{\substack{\chi \bmod N\\ \chi(g)=1\\ \chi\neq\chi_0}}\left(\sum_{i\in I_{a,b}}\overline{\chi}(x^i)\right)\chi(k')\\
&=\sum_{\substack{\chi_1 \bmod p^{m+1}\\ \chi_1(g)\chi_2(g)=1\\ \chi_1\chi_2\neq\chi_0}}\sum_{\chi_2 \bmod q^{n+1}}\left(\sum_{i\in I_{a,b}}\overline{\chi}_1(x^i)\overline{\chi}_2(x^i)\right)\\
&\quad\times\chi_1(k')\chi_2(k')\\
&=\sum_{\substack{\chi_1 \bmod p^{m+1}\\ \chi_1(g)\chi_2(g)=1\\ \chi_1\chi_2\neq\chi_0}}\sum_{\chi_2 \bmod q^{n+1}}\left(\sum_{i\in I_{a,b}}\overline{\chi}_1(g^i)\right)\chi_1(k')\chi_2(k').
\end{aligned}$$

Write

$$\begin{aligned}
\chi_1(k') &= \begin{cases} e\left(\frac{k_1\mathrm{ind}_{g,p^{m+1}}(k')}{p^m(p-1)}\right), & (k',p)=1,\\ 0, & (k',p)>1,\end{cases}\\
\chi_2(k') &= \begin{cases} e\left(\frac{k_2\mathrm{ind}_{g,q^{n+1}}(k')}{q^n(q-1)}\right), & (k',q)=1,\\ 0, & (k',q)>1,\end{cases}
\end{aligned}$$

where $e(y)=\mathrm{e}^{2\pi iy}$, $\mathrm{ind}_{g,p^{m+1}}(k')$ is the unique integer with $k'\equiv g^{\mathrm{ind}_{g,p^{m+1}}(k')} \pmod{p^{m+1}}$, $0\le\mathrm{ind}_{g,p^{m+1}}(k')\le p^m(p-1)-1$, and $\mathrm{ind}_{g,q^{n+1}}(k')$ denotes the unique integer with $k'\equiv g^{\mathrm{ind}_{g,q^{n+1}}(k')} \pmod{q^{n+1}}$, $0\le\mathrm{ind}_{g,q^{n+1}}(k')\le q^n(q-1)-1$. Then we have

$$\begin{aligned}
&\sum_{\substack{\chi \bmod N\\ \chi(g)=1\\ \chi\neq\chi_0}}\left(\sum_{i\in I_{a,b}}\overline{\chi}(x^i)\right)\chi(k')\\
&=\sum_{\substack{k_1=0\\ e\left(\frac{k_1}{p^m(p-1)}\right)e\left(\frac{k_2}{q^n(q-1)}\right)=1\\ k_1^2+k_2^2>0}}^{p^m(p-1)-1}\sum_{k_2=0}^{q^n(q-1)-1}\left(\sum_{i\in I_{a,b}}e\left(-\frac{ik_1}{p^m(p-1)}\right)\right)\\
&\quad\times e\left(\frac{k_1\mathrm{ind}_{g,p^{m+1}}(k')}{p^m(p-1)}\right)e\left(\frac{k_2\mathrm{ind}_{g,q^{n+1}}(k')}{q^n(q-1)}\right).
\end{aligned}$$

It is not hard to show that

$$\begin{aligned}
&\mathrm{e}\left(\frac{k_1}{p^m(p-1)}\right)\mathrm{e}\left(\frac{k_2}{q^n(q-1)}\right)=1\\
&\iff \mathrm{e}\left(\frac{k_1q^n(q-1)+k_2p^m(p-1)}{p^mq^n(p-1)(q-1)}\right)=1\\
&\iff p^mq^n(p-1)(q-1)\mid k_1q^n(q-1)+k_2p^m(p-1)\\
&\iff \frac{p^mq^n(p-1)(q-1)}{d}\Big|k_1\frac{q^n(q-1)}{d}+k_2\frac{p^m(p-1)}{d}.
\end{aligned}$$

Then we deduce

$$\frac{p^m(p-1)}{d}\Big|k_1,\qquad \frac{q^n(q-1)}{d}\Big|k_2.$$

Hence,

$$\begin{aligned}
&\sum_{\substack{\chi \bmod N\\ \chi(g)=1\\ \chi\neq\chi_0}}\left(\sum_{i\in I_{a,b}}\overline{\chi}(x^i)\right)\chi(k')\\
&=\sum_{\substack{0\le t_1\le d-1\\ t_1+t_2\equiv 0 \pmod d\\ t_1^2+t_2^2>0}}\sum_{0\le t_2\le d-1}\sum_{i\in I_{a,b}}\left(e\left(-\frac{it_1}{d}\right)\right)\\
&\quad\times e\left(\frac{t_1\mathrm{ind}_{g,p^{m+1}}(k')}{d}\right)e\left(\frac{t_2\mathrm{ind}_{g,q^{n+1}}(k')}{d}\right)\\
&=\sum_{t=1}^{d-1}\left(\sum_{i\in I_{a,b}}e\left(-\frac{it}{d}\right)\right)e\left(\frac{t\mathrm{ind}_{g,p^{m+1}}(k')}{d}\right)\\
&\quad\times e\left(-\frac{t\mathrm{ind}_{g,q^{n+1}}(k')}{d}\right).
\end{aligned}\tag{11}$$

By Equation (10) and Equation (11), together with the

definition of $I_{a,b}$ we obtain

$$\begin{aligned}(-1)^{s_k} &= -\frac{2}{d}\sum_{t=1}^{d-1}\left(\sum_{i=0}^{\frac{d}{2}-1} e\left(-\frac{(2i+1)t}{d}\right)\right)\\ &\quad \times e\left(\frac{t\mathrm{ind}_{g,p^{m+1}}(k')}{d}\right) e\left(-\frac{t\mathrm{ind}_{g,q^{n+1}}(k')}{d}\right)\\ &= e\left(\frac{\mathrm{ind}_{g,p^{m+1}}(k')}{2}\right) e\left(-\frac{\mathrm{ind}_{g,q^{n+1}}(k')}{2}\right)\\ &= \left(\frac{k'}{p}\right)\left(\frac{k'}{q}\right).\end{aligned}$$

Then for $0 \le k \le p^{m+1}q^{n+1}-1$, we have

$$(-1)^{s_k} = \begin{cases} \left(\frac{k'}{p}\right)\left(\frac{k'}{q}\right), \text{ if } k = p^a q^b k', \\ \qquad 0 \le a \le m,\ 0 \le b \le n,\ (k', pq) = 1, \\ 1, \\ \qquad \text{if } q^{n+1} \mid k, \\ -1, \\ \qquad \text{if } p^{m+1} \mid k,\ k > 0. \end{cases}$$

For $1 \le w \le p^{m+1}q^{n+1}-1$ with $\left(w, p^{m+1}q^{n+1}\right) = p^{a_0}q^{b_0}$, we get

$$\begin{aligned}C_s(w) &= \sum_{k=0}^{p^{m+1}q^{n+1}-1} (-1)^{s_{k+w}+s_k}\\ &= \sum_{a_1=0}^{m}\sum_{b_1=0}^{n}\sum_{a_2=0}^{m}\sum_{b_2=0}^{n} \sum_{\substack{k=0\\(k,p^{m+1}q^{n+1})=p^{a_1}q^{b_1}\\(k+w,p^{m+1}q^{n+1})=p^{a_2}q^{b_2}}}^{p^{m+1}q^{n+1}-1} \left(\frac{\frac{k}{p^{a_1}q^{b_1}}}{p}\right)\left(\frac{\frac{k}{p^{a_1}q^{b_1}}}{q}\right)\left(\frac{\frac{k+w}{p^{a_2}q^{b_2}}}{p}\right)\left(\frac{\frac{k+w}{p^{a_2}q^{b_2}}}{q}\right)\\ &\quad + \sum_{a=0}^{m}\sum_{b=0}^{n} \sum_{\substack{k=0\\(k,p^{m+1}q^{n+1})=p^{a}q^{b}\\q^{n+1}\mid k+w}}^{p^{m+1}q^{n+1}-1} \left(\frac{\frac{k}{p^{a}q^{b}}}{p}\right)\left(\frac{\frac{k}{p^{a}q^{b}}}{q}\right)\\ &\quad - \sum_{a=0}^{m}\sum_{b=0}^{n} \sum_{\substack{k=0\\(k,p^{m+1}q^{n+1})=p^{a}q^{b}\\p^{m+1}\mid k+w\\p^{m+1}q^{n+1}\nmid k+w}}^{p^{m+1}q^{n+1}-1} \left(\frac{\frac{k}{p^{a}q^{b}}}{p}\right)\left(\frac{\frac{k}{p^{a}q^{b}}}{q}\right)\\ &\quad + \sum_{a=0}^{m}\sum_{b=0}^{n} \sum_{\substack{k=0\\q^{n+1}\mid k\\(k+w,p^{m+1}q^{n+1})=p^{a}q^{b}}}^{p^{m+1}q^{n+1}-1} \left(\frac{\frac{k+w}{p^{a}q^{b}}}{p}\right)\left(\frac{\frac{k+w}{p^{a}q^{b}}}{q}\right)\\ &\quad - \sum_{a=0}^{m}\sum_{b=0}^{n} \sum_{\substack{k=1\\p^{m+1}\mid k\\(k+w,p^{m+1}q^{n+1})=p^{a}q^{b}}}^{p^{m+1}q^{n+1}-1} \left(\frac{\frac{k+w}{p^{a}q^{b}}}{p}\right)\left(\frac{\frac{k+w}{p^{a}q^{b}}}{q}\right)\\ &\quad + \sum_{\substack{k=0\\q^{n+1}\mid k\\p^{m+1}\mid k+w\\p^{m+1}q^{n+1}\nmid k+w}}^{p^{m+1}q^{n+1}-1} 1 - \sum_{\substack{k=1\\p^{m+1}\mid k\\p^{m+1}\mid k+w}}^{p^{m+1}q^{n+1}-1} 1\\ &\quad - \sum_{\substack{k=0\\q^{n+1}\mid k\\p^{m+1}\mid k+w\\p^{m+1}q^{n+1}\nmid k+w}}^{p^{m+1}q^{n+1}-1} 1 + \sum_{\substack{k=1\\p^{m+1}\mid k\\p^{m+1}\mid k+w\\p^{m+1}q^{n+1}\nmid k+w}}^{p^{m+1}q^{n+1}-1} 1.\end{aligned}$$

The statements of Theorem 2 then follows from Lemma 4-Lemma 6.

4 Conclusions

In this paper we have proven the linear complexity and autocorrelation values of a family of generalized cyclotomic sequences of period N with any order d. The result of linear complexity improves certain statement of [6] and the result of autocorrelation is new.

In 2012 Hu, Yue and Wang [6] gave a method for computing the linear complexity of Whiteman's generalized cyclotomic sequences of period $p^{m+1}q^{n+1}$ $(m, n \ge 0)$ with any order d. The method is applied to computing the exact linear complexity of Whiteman's generalized cyclotomic sequences of period pq with order 4 and period $p^{m+1}q^{n+1}$ $(m, n \ge 0)$ with order 4, respectively. In fact, it is difficult to compute the exact value of $A_{u,v}$ for $0 \le u \le m$ and $0 \le v \le n$ in the calculation formula [6]. In this paper we determine the exact linear complexity and the exact values of autocorrelation of Whiteman's generalized cyclotomic binary sequences of any order d and period $p^{m+1}q^{n+1}$ $(m, n \ge 0)$ due to the different definitions of the support set, which makes it easier to ensure the balance of these sequences.

The autocorrelation values of generalized cyclotomic sequences with respect to p^n for any $n > 0$ are calculated in [7] by using formulas for the generalized cyclotomic numbers of order 2. We can use the proof method of Theorem 2 to calculate the autocorrelation values of these sequences.

It seems more difficult to calculate the autocorrelation values of generalized cyclotomic sequences. By making a more detailed division on $p^i q^{n+1}\mathbb{Z}_N^*$ and $p^{m+1}q^j\mathbb{Z}_N^*$, Ke, Li and Zhang [8] determined the linear complexity of a new class of generalized cyclotomic binary sequences of period $p^{m+1}q^{n+1}$ $(m, n > 0)$. However, the exact values of autocorrelation of these sequences have not been calculated by now.

We will further study the autocorrelations of quaternary cyclotomic sequences over $\mathbb{F}_4$ of length $2p^m$.

Acknowledgments

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

Z.X.C. was partially supported by the National Natural Science Foundation of China under grant No. 61772292, by the Projects of International Cooperation and Exchanges NSFC-RFBR No. 61911530130, by the Provincial Natural Science Foundation of Fujian under grant No. 2018J01425 and by the Program for Innovative Research Team in Science and Technology in Fujian Province University under grant No. 2018-49.

H.N.L was was partially supported by National Natural Science Foundation of China under Grant No. 11571277, and the Science and Technology Program of Shaanxi Province of China under Grant No. 2016GY-080 and 2016GY-077.

References

[1] N. Brandstätter and A. Winterhof, "Some notes on the two-prime generator of order 2," *IEEE Transactions on Information Theory*, vol. 51, no. 10, pp. 3654–3657, 2005.

[2] Z. X. Chen and V. Edemskiy, "Linear complexity of quaternary sequences over $\mathbb{Z}_4$ derived from generalized cyclotomic classes modulo $2p$," *International Journal of Network Security*, vol. 19, no. 4, pp. 613–622, 2017.

[3] T. W. Cusick, C. S. Ding, and A. Renvall, *Stream Ciphers and Number Theory*. Amsterdam: Elsevier, 2004.

[4] C. S. Ding, *Codes from Difference Sets*. Singapore: World Scientific, 2014.

[5] C. S. Ding and T. Helleseth, "New generalized cyclotomy and its applications," *Finite Fields and Their Applications*, vol. 4, no. 2, pp. 140–166, 1998.

[6] L. Q. Hu, Q. Yue, and M. H. Wang, "The linear complexity of Whiteman's generalized cyclotomic sequences of period $p^{m+1}q^{n+1}$," *IEEE Transactions on Information Theory*, vol. 58, no. 8, pp. 5534–5543, 2012.

[7] S.-Y. Jin, Y.-J. Kim, and H.-Y. Song, "Autocorrelation of new generalized cyclotomic sequences of period p^n," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. 93, no. 11, pp. 2345–2348, 2010.

[8] P. H. Ke, R. F. Li, and S. Y. Zhang, "The linear complexity of a new class of generalized cyclotomic binary sequences of length $p^{m+1}q^{n+1}$ (in Chinese)," *Acta Electronica Sinica*, vol. 42, no. 5, pp. 1009–1013, 2014.

[9] L. F. Liu, X. Y. Yang, X. N. Du, and B. Wei, "On the linear complexity of new generalized cyclotomic binary sequences of order two and period pqr," *Tsinghua Science and Technology*, vol. 21, no. 3, pp. 295–301, 2016.

[10] A. L. Whiteman, "A family of difference sets," *Illinois Journal of Mathematics*, vol. 6, no. 1, pp. 107–121, 1962.

[11] C. H. Wu, X. N. Du, and Z. T. Jiang, "Linear complexity of a family of pseudorandom discrete logarithm threshold sequences," *International Journal of Network Security*, vol. 18, no. 3, pp. 487–492, 2016.

[12] Z. B. Xiao, X. Y. Zeng, C. L. Li, and T. Helleseth, "New generalized cyclotomic binary sequences of period p^2," *Designs, Codes and Cryptography*, vol. 86, no. 7, pp. 1483–1497, 2018.

[13] T. J. Yan, K. Fan, X. N. Du, and G. Z. Xiao, "Linear complexity of binary Whiteman generalized cyclotomic sequences (in Chinese)," *Journal of Xidian University*, vol. 33, no. 4, 2006.

Biography

Xiaolin Chen was born in 1992. She is a Ph.D. student of Northwest University. She received the M.S. degree in Mathematics from Northwest University in 2017. Her research interests include number theory and information security.

Zhixiong Chen was born in 1972. He received the M.S. degree in Mathematics from Fujian Normal University in 1999 and Ph.D. degree in Cryptography from Xidian University in 2006, respectively. Now he is a professor of Putian University. He worked as a visiting scholar supervised by Prof. Arne Winterhof in Austrian Academy of Sciences (Linz) in 2013 and by Prof. Andrew Klapper in University of Kentucky (Lexington) during 2014-2015, respectively. His research interests include stream cipher, elliptic curve cryptography and digital signatures.

Huaning Liu was born in 1979. He received the M.S. degree in Mathematics from Northwest University in 2004 and Ph.D. degree in Mathematics from Northwest University in 2007, respectively. He was the recipient of the Zhong Jiaqing Mathematics Award in 2005. He worked as a post-doctoral fellow at School of Mathematics, Shandong University during 2007-2011. He worked as a post-doctoral fellow at Department of Pure Mathematics and Mathematical Statistics, University of Cambridge during 2012-2013. Now he is a professor of Northwest University. His research interests include number theory and information security.

Mobile Payment Security in the Context of Big Data: Certificateless Public Key Cryptography

Tianhong Yang
(Corresponding author: Tianhong Yang)
Changchun Finance College
Changchun, Jilin 130028, China
(Email: t221h7@163.com)
(Received Mar. 3, 2019; Revised and Accepted Jan. 28, 2020; First Online June 14, 2020)

Abstract

The development and popularization of wireless networks and mobile terminals dramatically facilitates people's lives. Moreover, online business activities gave birth to mobile payments and extended to offline. Mobile payment has been promoted because of its convenience, so its security has been paid great attention. This study briefly introduced the mobile payment and certificateless public key cryptography technology, applied the certificateless public key cryptography technology to the mobile payment system, and simulated the security and efficiency of the mobile payment system in the wireless LAN built in the laboratory. The results showed that users and cloud platform could encrypt and decrypt information through two-way identity authentication and some private keys and public parameters in the process of mobile payment. When attackers attacked mobile payment, whether they pretended to be users or cloud platforms, decryption would fail due to lack of identity authentication and some private keys, to ensure the security of mobile payment. Moreover, they could maintain standard processing and resist the decryption from the third party when processing a large amount of payment information. Compared with public key infrastructure (PKI) based mobile payment and wireless public key infrastructure (WPKI) based mobile payment, mobile payment had higher payment efficiency.

Keywords: Certificateless Public Key Cryptography; Internet; Mobile Payment; Payment Security

1 Introduction

With the rapid development of Internet technology and wireless communication technology, our daily life has changed dramatically in recent years. The popularity of various intelligent mobile terminals makes it easier for people to access the Internet anytime and anywhere [9]. Moreover, after combining Internet technology, the traditional service industry has gradually transformed into the online digital service industry, thus promoting the development of mobile payment [2]. With the help of mobile terminals, mobile payment can initiate transactions anytime and anywhere, and the content of transactions is diverse, which significantly facilitates users. Because of the "anytime and anywhere" feature of mobile payment, users do not need to carry a large amount of cash, which is relatively safer when they go out. Compared with money, mobile payment is more convenient and secure, but it also faces similar security problems [12].

In the process of execution, mobile payment often contains a lot of private information when sending payment information to the third-party institutions such as banks and merchants. When the popularity of the mobile network is not high, the problem is not apparent. However, with the popularization of mobile networks and the progress of computer technology, the consequences are unimaginable once the user's payment information is intercepted. Therefore, the security of mobile payment is fundamental to ensure the healthy development of the online service industry. Abughazalah *et al.* [1] provided a protocol for NFC mobile phones to meet the security requirements of mobile payment, formally analyzed the protocol using CasperFDR, and found no feasible attack. Madhoun *et al.* [7] proposed to add a new security layer to strengthen the protocol against the security vulnerability of the European MasterCard protocol (EMV), and verified the security of the improved EMV using a security verification tool named Scyther.

Thammarat *et al.* [13] proposed a new NFC mobile payment protocol, which could provide fair exchange and information security in the POS transaction process and found through experiments that the protocol had more effective protection and fairness than other protocols. This study briefly introduced the mobile payment and certificateless public key cryptography technology, applied certificateless public key cryptography technology to the mobile payment system, and carried out simulation experiments on the security and efficiency of the mobile payment system are simulated in the wireless LAN built in

the laboratory.

2 Mobile Payment

The basic framework of the mobile payment system is shown in Figure 1. The mobile payment system can be basically divided into six parts, from bottom to top.

The first part is the bearer network [10], including various forms of mobile networks, such as 3G, 4G, WiFi, *etc.*, which are usually provided by local communication companies and is the basis of the system and whose quality directly affects the stability of the system.

The second part is an access platform which provides various interfaces for connecting the bearer network for payment system to realize the unified interface of different kinds of payment services and integrate the payment function.

The third part is security authentication [15], which is the core component for the mobile payment system to ensure payment security and provides the system with functions such as data encryption and identity authentication. Security authentication covers the three platforms introduced later in the order, i.e., authentication is not an independent platform, but is integrated into the three platforms introduced then to ensure its security.

The fourth part is the management platform, which has functions of business support, management support, and system support for respectively managing business transaction content, merchant customer information, and system data.

The fifth part is a business platform, which includes a business system and payment system; the former provides various business services on the loose end, and the latter offers different forms of payment function [11].

The sixth part is the application platform, which is the top layer of the mobile payment system and undertakes the function of information interaction with users and merchants.

3 Certificateless Public Key Cryptography

The security of mobile payment system was ensured by the certificateless public key cryptography in this study. The certificateless public key cryptography [3] is based on the assumption of a difficult problem. No algorithm can solve the problem quickly in mathematics so that the security can be guaranteed. The algorithm of certificateless public key cryptography can be divided into seven steps:

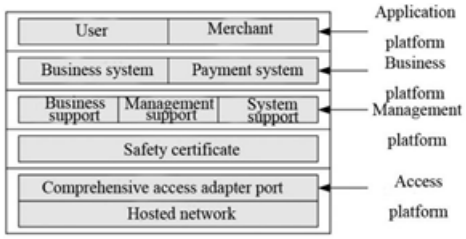


Figure 1: The basic framework of the mobile payment system

1) Firstly, through the key generation center (KGC) [5], master secrete key (msk) and public parameter [6] are obtained according to security parameter k;

2) Then msk, $params$, and ID for indicating user identity are input into KGC to obtain some private key D. D is transmitted to the corresponding user through the secure channel of the mobile network;

3) $params$ and ID are input into KGC to obtain secrete value x;

4) $params, D$ and x are combined to generate private key sk through KGC;

5) $params$ and x are combined to generate public key pk through KGC;

6) Data M which needs encryption is encrypted using $params$, sender ID, sk and receiver ID and pk, and it is transmitted to the platform of the receiver through the secure channel of mobile network;

7) After receiving the encrypted M, the receiver platform makes authentication with $params$, sender ID, sk, receiver ID and pk and decrypts it to obtain plaintext.

In the above algorithm steps of certificateless public key cryptography, Steps 1-5 are about generating encryption parameters and public and private keys about KGC, and Steps 6 and 7 are about encrypting and decrypting the plaintext information by using the encryption parameters and public and private keys. Compared with the traditional public key cryptography described above, certificateless public key cryptography avoids the management problem of public key certificate and the hosting problem of private key as the secrete key generated by KGC is only a part and will be combined with x randomly selected by user when using and the private key is stored by user.

The flow of mobile payment based on certificateless public key cryptography is shown in Figure 2. After the customer places an order on the mobile terminal, the merchant generates payment information according to the order and transmits it to the customer's mobile terminal through the secure channel of the mobile network [14].

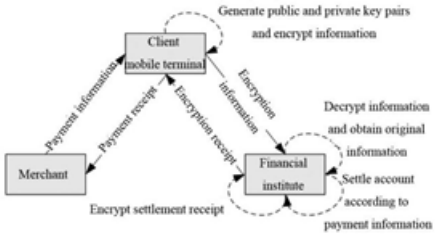


Figure 2: The mobile payment flow based on certificateless public key cryptography

Then, the public key pair is generated in the server of the mobile terminal and the payment information is encrypted.

The encrypted payment information is transmitted to the customer through the secure channel. Financial institutions decrypt the encrypted information after passing the authentication, so as to obtain the payment information, then settle the user's account according to the payment information, and generate the settlement receipt. Financial institutions also encrypt the settlement receipt by using the certificateless public key cryptography technology and transmit the encrypted receipt to the customer's mobile terminal through the secure channel. After receiving the encrypted receipt, the mobile terminal decrypts it in its server to obtain the payment receipt information and then transmits the payment receipt information to the merchant through the secure channel. After receiving the payment receipt, the merchant confirms that the payment is successful, and the whole payment process ends.

4 Simulation Experiment

4.1 Experimental Environment

In this experiment, the cloud platform was built using OpenSack [4]. The computer, mobile phone and cloud platform in the laboratory were all in the wireless LAN of the laboratory. The computer was configured with 6G memory, core i5 processor and 32-bit Windows 7 operating system; the mobile phone was configured with Android operating system, 4G memory and six-core processor; the maximum transmission speed of wireless LAN was 100 MB/s.

4.2 Experiment Setup

1) Security analysis of mobile payment based on certificateless public key cryptography.
The experimental method was to publish the payment information including user name, order information, payment amount, *etc.*, by taking computer as the merchant, then send the payment information to the mobile phone, encrypt the payment information by the certificateless public key cryptography technology and send it to the cloud platform (the cloud platform in this experiment as a third-party financial institution) to simulate the process of mobile payment. In order to verify the security of the mobile payment system in the big data environment, 10 100 pieces of payment information were processed in the WLAN at the same time, and the processing method was the same as experiment Steps 1 and 2 above. The regular payment process was used to interact 10-100 pieces of payment information in the LAN at the same time, and the irregular third-party attacker was applied for pretending to be the platform and users in LAN. In the regular payment process experiment, the number of successful payment was counted; for the irregular payment experiment, the decryption degree of the attacker to the payment information was counted. This experiment was compared with mobile payment systems based on public key cryptography and security authentication.

2) Efficiency analysis of mobile payment based on certificateless public key cryptography.
In this study, regular mobile payment was performed in the mobile payment system through mobile phone, and the whole mobile payment process was divided into five processes:

   a. The phase when mobile phones received payment information from merchants;
   b. The phase when mobile phones received information and sent it to the cloud platform after encryption;
   c. The phase when the cloud platform receives information, decrypts it, and sends the information to the mobile phone according to the information encryption payment receipt;
   d. The phase when the mobile phone receives the payment receipt and decrypts it and sends it to the merchant;
   e. The phase when the merchant receives the payment receipt.

In this study, the time point of information transmission and reception was determined by the time stamp of data transmission [8], so as to calculate the time consumption of each phase. Moreover, in order to verify the efficiency of the mobile payment system proposed in this study, it was compared with two mobile payment systems based on public key cryptography and security authentication system. The total number of payment information transmitted in the experiment was 100, and the average value was taken as the final result.

4.3 Experimental Results

There is no severe problem in the regular payment; therefore considering the limited length of paper, only the background logs of the cloud platform and attacker in the irregular payment process are displayed. It was seen from

Figure 3 that mobile phone B, as an attacker, intercepted the payment information ciphertext sent by mobile phone A at 08:32:00. At that time, mobile phone B disguised itself as a cloud platform to receive the payment ciphertext and then authenticated and decrypted the ciphertext. However, as mobile phone B was not registered in the mobile system and did not master part of the private key of mobile phone A, the authentication failed, and only messy code was obtained after decryption.

As the real payment information cannot be obtained, mobile phone B attempted to forge the payment information, encrypted the forged payment information with public parameters and forged private keys, and sent it to the cloud platform. The above process lasted about 3 s according to the records. The cloud platform received the ciphertext at 08:32:03, and the cloud platform did not know whether the ciphertext was from user A at that moment and then it verified the identity of the ciphertext sender according to the registration data stored in the database and decrypted it with part of the private key and public parameters. The final identity authentication failed, and the only messy code was obtained after decryption. The cloud platform determined the payment information as invalid and deleted it.

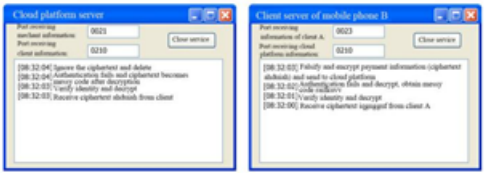


Figure 3: Background logs of the cloud platform and attacker client in the regular mobile payment process

One of the characteristics of big data environment in mobile network is the huge amount of data transferred. Therefore, in order to verify the security of this mobile payment system in big data environment, mobile payment information transferred at the same time was gradually added in wireless LAN, and the system was compared with the payment system under the other two cryptography technologies. The comparison results are shown in Figure 4.

In the regular payment process, the payment system proposed in this study could effectively handle the increase of the amount of payment information that needed to be processed in the network, without payment failure; the payment system based on public key infrastructure (PKI) and the payment system based on wireless public key infrastructure (WPKI) did not show payment failure when processing 10-30 pieces of payment information, but when processing 40 or more pieces of payment information, both payment systems had payment failure, and the WPKI based payment system failed more times.

In the irregular payment process, the decryption degree of the third party attacker to the payment ciphertext of the payment system kept at about 1% with the increase of the amount of payment information, which was regarded as the decryption failure; the decryption degree of the third party attacker to the payment ciphertext of the payment system under the other two cryptography technologies increased with the increase of the amount of payment information, in which the increase of the PKI based payment system was faster.

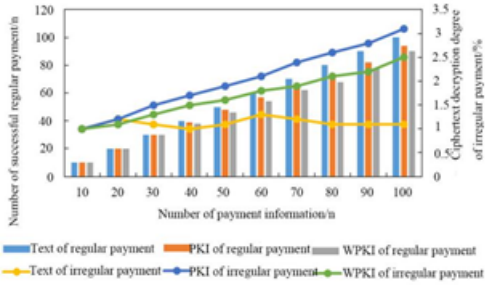


Figure 4: Security comparison of three cryptography technologies in big data environment

For mobile payment, its greatest convenience is that it can initiate transactions anytime and anywhere within the network coverage. In the existing big data environment, the amount of data transmitted in the network is very large, and the communication channel resources of the mobile network are limited, so the efficiency of mobile payment will directly affect the user's experience. The faster the payment efficiency, the lower the delay, the less the occupied channel resources, the more the mobile payment users supported. In order to verify the efficiency of the mobile payment proposed in this study, it was compared with the other two mobile payment systems, and the results are shown in Figure 5. PKI represents mobile payment based on public key cryptography, and WPKI represents mobile payment based on security authentication system. It was seen intuitively from Figure 5 that no matter what kind of mobile payment was used, the time consumed was the shortest in Steps 1 and 5, and the time consumed in Step 3 was the longest. The reason for the above phenomenon was that Steps 1 and 5 were only about the transfer of payment information and receipt information, and Step 3 not only involves encryption, but also involves decryption.

As to the comparison between the three mobile payment means at the same phase, the difference between them atSteps 1 and 5 was not large, and the reason has been described above; in Steps 2, 3, and 4, the mobile payment proposed in this study consumed the shortest time and the WPKI based mobile payment consumed the longest time. The reason for the above phenomenon was that Hash function which was used for encryption and decryption consumed more time in the computation of finite field compared to point multiplication, and the mobile payment proposed in this study used less Hash function computation in the encryption and decryption phases, and the WPKI based mobile payment needed safety cer-

tificate certification in the certificate management center in addition to Hash function computation. Finally, the average total time consumed by the mobile payment system proposed in this study was 49.28 ms, the average total time consumed by the mobile payment based on PKI was 50.72 ms, and the average total time consumed by the mobile payment based on WPKI was 51.95 ms.

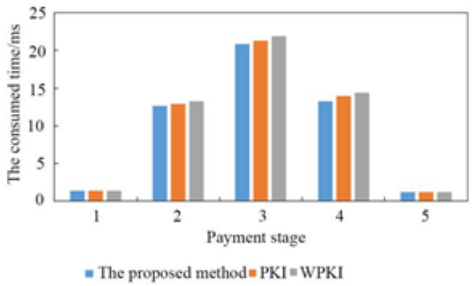


Figure 5: Comparison of mobile payment efficiency under three kinds of cryptography

5 Conclusion

This paper introduced the mobile payment and certificateless public key cryptography, applied certificateless public key cryptography to the mobile payment system, and carried out the simulation experiment on the security and efficiency of the mobile payment system in the wireless LAN built in the laboratory. The experimental results are as follows:

1) No matter another mobile terminal with the same configuration pretended to be the cloud platform to receive information or pretended to be the regular user to send information when attacking the mobile payment, it failed to pass authentication and decryption;

2) With the increase of payment information in the network, the payment system proposed in this study could ensure the success of payment in the process of regular payment and maintained a very low degree of decryption under the attack of the third party; times of failure of the other two payment systems increased with the increase of payment information to be processed, and the degree of decryption under attack also increased;

3) The average total time consumed by the mobile payment system proposed in this study was 49.28 ms, the average total time consumed by the mobile payment based on PKI was 50.72 ms, and the average total time consumed by the mobile payment based on WPKI was 51.95 ms; the phase of encryption and decryption consumed the longest time in the whole payment process.

Acknowledgment

The study is supported by Changchun Finance College "Research on Security Problems and Countermeasures of the Environment for Mobile Payments" (JZSJPT2019004).

References

[1] S. Abughazalah, K. Markantonakis, and K. Mayes, "Secure mobile payment on NFC-enabled mobile phones formally analysed using casperFDR," in *IEEE 13th International Conference on Trust, Security and Privacy in Computing and Communications*, 2014.

[2] T. Dahlberg, J. Guo, and J. Ondrus, "A critical review of mobile payment research," *Electronic Commerce Research and Applications*, vol. 14, no. 5, pp. 265-284, 2015.

[3] S. K. H. Islam, and A. Singh, "Provably secure one-round certificateless authenticated group key agreement protocol for secure communications," *Wireless Personal Communications*, vol. 85, no. 3, pp. 879-898, 2015.

[4] B. Klugah-Brown, J. B. A. K. Ansuura, and Q. Xia, "A Signcryption Scheme from Certificateless to Identity-based Environment for WSNs into IoT," *International Journal of Computer Applications*, vol. 120, no. 9, pp. 16-23, 2015.

[5] 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.

[6] M. M. Ma, D. B. He, and N. Kumar, "Certificateless searchable public key encryption scheme for industrial internet of things," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 2, pp. 759-767, 2017.

[7] N. E. Madhoun, and G. Pujolle, "Security enhancements in EMV protocol for NFC mobile payment," in *IEEE Trustcom/BigDataSE/ISPA*, 2016.

[8] S. Mathi, R. Nivetha, B. Priyadharshini, and S. Padma, "A certificateless public key encryption based return routability protocol for next-generation IP mobility to enhance signalling security and reduce latency," *Sadhana*, vol. 42, no. 12, pp. 1987-1996, 2017.

[9] T. Oliveira, M. Thomas, G. Baptista, and F. Campos, "Mobile payment," *Computers in Human Behavior*, vol. 61(C), pp. 404-414, 2016.

[10] J. Ondrus, A. Gannamaneni, and K. Lyytinen, "The impact of openness on the market potential of multi-sided platforms: A case study of mobile payment platforms," *Journal of Information Technology*, vol. 30, no. 3, pp. 260-275, 2015.

[11] C. Phonthanukitithaworn, C. Sellitto, and M. W. L. Fong, "An investigation of mobile payment (m-payment) services in Thailand," *Asia-Pacific Journal*

of Business Administration, vol. 8, no. 1, pp. 37-54, 2016.

[12] E. Taylor, "Mobile payment technologies in retail: a review of potential benefits and risks," *International Journal of Retail & Distribution Management*, vol. 44, no. 2, pp. 159-177, 2016.

[13] C. Thammarat, W. Kurutach, and S. Phoomvuthisarn, "A secure lightweight and fair exchange protocol for NFC mobile payment based on limited-use of session keys," in *17th International Symposium on Communications and Information Technologies (ISCIT'17)*, 2017.

[14] T. Tsai, Y. Tseng, and S. Huang, "Efficient revocable certificateless public key encryption with a delegated revocation authority," *Security and Communication Networks*, vol. 8, no. 18, pp. 3713-3725, 2016.

[15] M. H. Zhong, and H. Wu, "Research on the development of mobile payment industry chain in China," *Journal of Computational and Theoretical Nanoscience*, vol. 14, no. 1, pp. 221-224, 2017.

Biography

Tianhong Yang, born in 1989, has gained the master's degree. She is now a lecturer in Changchun Finance College. She is interested in electronic business.

A Novel Identity-based Authentication Scheme for IoV Security

Changguang Wang, Zimeng Dai, Dongmei Zhao, and Fangwei Wang
(Corresponding author: Fangwei Wang)

Key Lab of Network and Information Security of Hebei Province
No.20, South ErHuan Road, YuHua District, Shijiazhuang 050024, China
College of Computer and Cyber Security, Hebei Normal University
No.20, South ErHuan Road, YuHua District, Shijiazhuang 050024, China
(E-mail: fw_wang@hebtu.edu.cn)
(Received Feb. 10, 2019; Revised and Accepted Aug. 3, 2019; First Online Sept. 8, 2019)

Abstract

In order to enhance the security of the IoV (Internet of Vehicles), a novel bi-directional authentication scheme is presented in this paper. By use of the elliptic curve encryption algorithm and the bilinear pair mapping theory, this scheme is designed to store the main system information in the RSU (Road Side Unit). During the process of communication, the shared key, identity ID, and handshake principle are used to perform mutual security authentication between the RSU and OBU (On Board Unit), thus ensuring the legitimacy of the communication nodes. Simulation experiments show that the computational complexity is reduced by about 10%, the efficiency and security of the scheme are improved compared with the existing schemes while meeting the security requirements.

Keywords: Authentication; Certificate Authority; On-Board Unit; Road Side Unit; Security

1 Introduction

The Internet of Vehicle (IoV) is an application of the IoT in road traffic and is also an important part of the Intelligent Transport System (ITS). Through advanced information and communication technology [7, 9, 22], such as GPS, sensing technology, network technology, and image identification technology, it can inform and help drivers to avoid accidents and even take control measures in case of emergency.

Consisting of interconnected entities on the road, IoV is created spontaneously and used to exchange data, perceive the traffic conditions, monitor the running state of the car, improve road traffic effectively and provide comfort for drivers and passengers. Vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) are two types of communication between entities on the road. In order to achieve the above functions, when vehicles pass through V2V or V2I, they will communicate and exchange information with their neighbor nodes. They send information automatically at set intervals to find and judge their neighbors and share the state information (position, speed, acceleration and direction) so that dangerous accidents can be avoided. On the other hand, emergency information can be transmitted to inform all nearby neighbors in the event of an emergency notice.

There are several available schemes of authentication, such as the authentication scheme based on an anonymous certificate, the authentication scheme based on a group signature, and the authentication scheme based on an RSU, and so on [4, 11].

The authentication scheme based on the anonymous certificates was proposed by Raya and Hubaux in 2007 [20]. The main idea is that vehicles need a large number of anonymous certificates, which are issued by CA (Certificate Authority) and stored in the OBUs of vehicles. When a vehicle needs communication, it will randomly select an anonymous certificate to sign the message which needs to be broadcasted and discard the certificate after signing, to ensure the security and privacy of the message and then hide the identity information of the vehicle. Accordingly, it can make the communication process untraceable. Besides, in order to track the vehicle in the event of an accident, CA will keep the correspondence between the real identity information and the anonymous certificate of the vehicle during registration, so that it can achieve the traceability of the malicious behaviors. Although this kind of scheme can ensure the anonymity of the message, there are still some deficiencies, among which the cancellation process is the largest weakness. When a vehicle is revoked, the cancellation information needs to be broadcasted across the network. A Certificate Revocation List (CRL) will load a lot of certificates, thus reducing the efficiency of message authentication. Besides, this kind of scheme makes a high request for the

OBU's storage capacity.

The authentication scheme based on a group signature was first proposed by Chaum and Van Heyst in 1991. In 2001, Dan Boneh proposed a signature scheme based on the elliptic curve and the ultra-short of the elliptic curve, whose signature length was half of the DSA. In 2007, Lin *et al.* proposed a GSIS scheme [13], which combined the group signature and the identity authentication technology. In 2010 Wasef *et al.* proposed a group signature scheme [26], which supported the vehicle batch validation for the IoV security. In 2011 Chim *et al.* proposed a group signature authentication scheme based on software [5], which used Bloom filter and binary search technology. In 2013, Zhu also proposed a group signature authentication scheme [31] which used hash value for CRL verification. This kind of scheme reduces the storage requirements on the OBU, but introduces the role of the group administrator which can be fatal to the entire IoV when being attacked. Moreover, it is difficult to balance the network scale. If the number of vehicle nodes in the group is large, the growth of CRL will be rapid, leading to the decline of node authentication efficiency. If there are few vehicle nodes in the group, for example, only one vehicle enters the group area, it can easily attack the network through the group identity information.

In IoV, RSUs are fixed units with large capacity and high transmission rate and are generally deployed at crossroads. Lu *et al.* proposed an ECPP protocol [15], which was characterized by generating a dynamic short-time anonymous key between the OBU and the RSU. The LPA protocol [28] proposed by Xue *et al.* introduced the concept of RSU neighbor set. The characteristics of the scheme are using RSU/OBU for online authentication and providing the certificate updates for OBU by RSU. This scheme can quickly generate anonymous keys, fast perform anonymous authentication and track privacy between the OBU and RSU while minimizing the storage for the anonymous key. Therefore, it can reduce the overhead and the complexity of the certificate management, and also provide good security and high efficiency for vehicle communication. In the RSU-based scheme, the amount of computation and the storage of OBU are far less than that of other schemes, but the signature and verification signature are largely dependent on RSU. Accordingly, V2V communication is not supported and all the communications are dependent on RSUs.

Lee *et al.* proposed an improved Identity-Oriented batch authentication scheme [10], but it could not resist a replay attack and could not satisfy traceability. Bayat *et al.* put forward another improved scheme [2] to improve the safety performance, but the scheme is designed only based on bilinear pairs, which is inefficient and cannot meet the time performance requirements of vehicle networking.

At present, the convergent signature authentication algorithm and the certificate free cryptosystem [8, 27] have been studied, but they are not suitable for IoV communication for the high computational cost and not able to resist some attacks such as identity forgery.

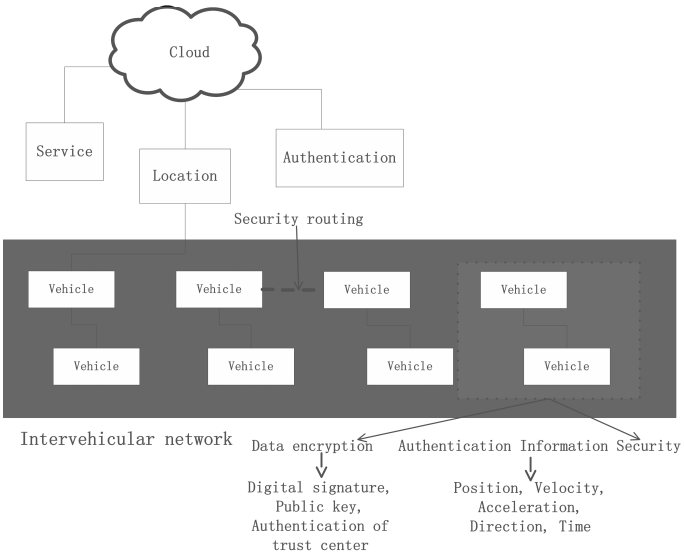


Figure 1: Security architecture of IoV

In this paper, we propose a novel identity-based bi-directional authentication scheme between OBU and RSU for IoV security based on considering the advantages and disadvantages of the existing schemes. Our scheme is designed by combining elliptic curve encryption with bilinear mapping theory, and can effectively make use of the characteristics of RSU to make the information transmission process more direct, fast and secure.

2 The Security Authentication Model

2.1 Network Model

The IoV security problems increase with the continuous development of its various applications. And the inherent characteristics of it, such as short distance, fast topology change, and openness, make the security threats even worst. The security architecture of it is shown in Figure 1.

Security requirements of the IoV mainly include non-forgery, privacy protection, traceability, revocability [14, 16, 30]. More and more secure authentication and privacy protection schemes are proposed to make the IoV nodes able to communicate with each other securely. In IoV, authentication is the core security requirement, which provides the integrity of information and avoids the manipulation of the exchanged information [25]. Fundamentally, all applications in IoV need to be authenticated [6, 19].

2.2 Identity Authentication Algorithm

Node identity authentication plays an important role in IoV researches. The most popular identity-based authentication algorithm is shown as follows [3]:

Algorithm 1 Identity-based authentication algorithm

1: **if** A is the identity **then**
2: Check whether VerA (A) is "valid"
3: Store A and its digest
4: **else if** A is the digest **then**
5: Verify whether A has been stored and whether it is valid
6: **end if**
7: Verify whether A exists in the RCL

2.3 Identity Authentication Model

In our scheme, there are mainly three entity parts which are CA, RSU, and OBU, respectively.

1) CA: The trust center mainly generates system parameters for the scheme and announce them to the public.

2) RSU: The roadside unit, the core part of the scheme, is used to store the main system information.

3) OBU: The onboard unit enables the vehicles to communicate with RSU or other vehicles.

Vehicle nodes can apply to CA for vehicle-related information when registering in the vehicle management office. The specific message construction is shown in Table 1.

Table 1: Message construction

Real ID of OBU
Private key of OBU
Certificate of OBU
Signature of OBU
Timestamp

Our scheme is designed based on the elliptic curve cryptography algorithm [12, 24, 29] and the bilinear pair mapping theory [23]. Here the hash function [18] is introduced to reduce the computational complexity. The scheme stores the main parameters of the system in the RSU instead of the OBU. When the vehicle arrives in the area covered by RSU, RSU and OBU need to perform mutual authentication. RSU judges whether to send the shared key for the OBU. OBU judges whether RSU is legal or not. If RSU is legal, OBU joins in the group. The main processes of the scheme include system initialization, RSU registration, OBU generating pseudonym identity, RSU and OBU mutual authentication, the RSU generating a temporary key, the OBU signing and transmitting the message, and revoking the identity. The specific model description is shown in Figure 2.

3 Scheme Flow

This paper proposes an identity-based mutual authentication scheme for IoV security, the specific process is shown in Figure 3.

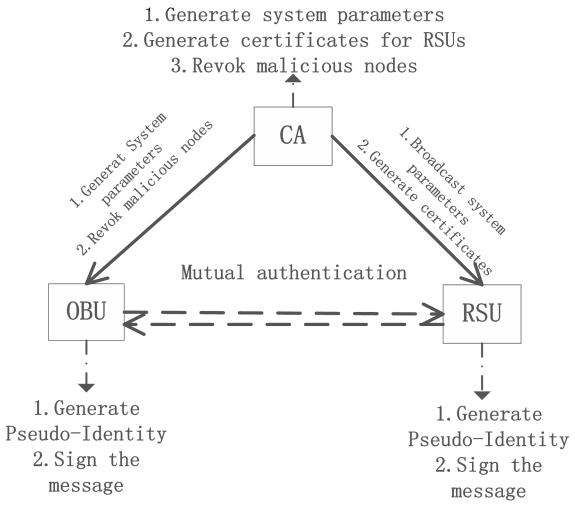


Figure 2: Schematic model

3.1 System Initialization

CA establishes system parameters for the scheme, which mainly includes the following aspects:

1) Define the finite field Z_q^* ={*0,1,2,3,. . . ,q-1*}, select the large prime number q, and select the elliptic curve $y^2 = x^3 + ax + b$, where $4a^3 + 27b^2 \neq 0 (\bmod q)$; define the cyclic additive group G_1 and the cyclic multiplicative group G_2. G_1 and G_2 have the same order q (q is a large prime number in the finite field), where $\hat{e} : G_1 \times G_2 \to G_2$ is the bilinear pairing principle. According to this principle, *CA* generates system parameters *(G_1, G_2,e,P,q)*, where P is the generator of the elliptic curve.

2) Define hash functions: $H1 : \{0,1\}^* \to G_1$, $H2 : \{0,1\}^* \times G_2 \to Z_q^*$.

3) In the finite field, *CA* selects a random number s as the system private key and then calculates $p = s \times P$ as the system public key.

4) *CA* selects the encryption and decryption functions *Ex(.)* and *Dx(.)* according to the elliptic curve encryption algorithm.

5) *CA* announces the parameters {G_1, G_2, q, P, s, p, $\hat{e}$, *H1(.), H2(.), Ex(.), Dx(.)*} to public, and stores them in RSU and OBU, respectively.

3.2 RSU Registration

1) *CA* selects a random number S_{RSU_j} in the finite field as the private key for each RSU and calculates $P_{RSU_j} = S_{RSU_j} \times P$ as the public key of RSU_j.

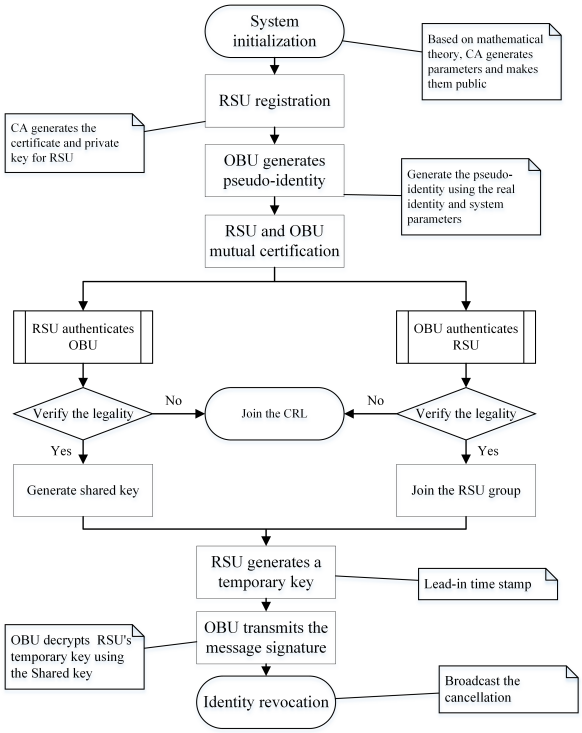


Figure 3: Flow chart of the scheme

2) *CA* uses the Schnorr scheme [17] to sign for each RSU. In the process of signature, ID_{RSU_j}, the identity of RSU_j, is used to select a random number k and a large prime number Q, where

$$\begin{aligned} m &= P^k \bmod Q, \\ e &= h(ID_{RSU_j} \| m), \\ d &= S_{RSU_j} e + k \bmod q. \end{aligned}$$

$Sign_{RSU_j}$ is *(d,e)*, and the generated certificate is $Cert_{RSU_j} = (P_{RSU_j}, Sign_{RSU_j})$

3) The generated certificate and private key are sent to each RSU through the security channel and stored in the RSU.

The process diagram is shown in Figure 4.

3.3 OBU Generates Pseudonym Identity

To ensure the traceability of vehicle nodes, sign the nodes based on OBU, and generate certificates $Cert_{OBU_i}$. The signature method is similar to RSU.

Here, each OBU uses real identity and public parameters to generate a pseudonym identity, thus can securely authenticate the OBU. The identity generation process is as follows:

$\mathbf{OBU}_i$ selects a random parameter r from the finite field

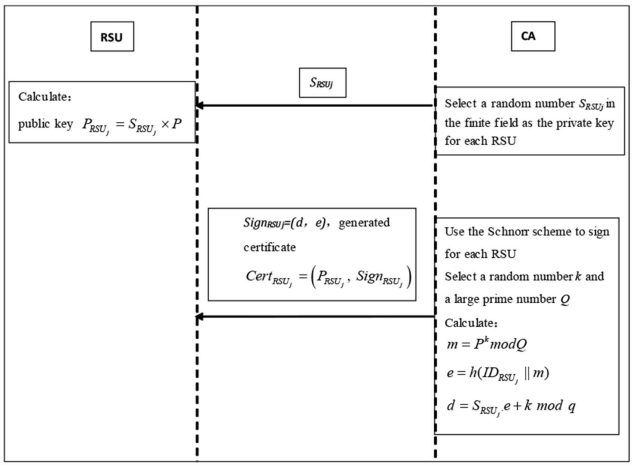


Figure 4: RSU registration process diagram

and then calculates

$$\begin{aligned} ID_i^{'} &= rP, \\ ID_i^{''} &= ID_i^R \oplus H2(rp), \\ ID_i &= \langle ID_i^{'}, ID_i^{''} \rangle. \end{aligned} \quad (1)$$

$\mathbf{ID}_i^R$ is the real ID of OBU_i, and ID_i is the pseudonym ID of OBU_i.

3.4 RSU and OBU Authenticate Mutually

In order to make it legal for the nodes of the communication process, each RSU and OBU perform mutual authentication to prevent malicious nodes from faking identity for attacks, thus guaranteeing the security of the system to the utmost extent.

1) RSU authenticates OBU
Before sending a message, OBU_i sends the pseudo-identity ID obtained in Step 3.3 to RSU_j through the secure channel, and RSU_j calculates the real identity ID of OBU_i through the public key p of the system.

$$ID_i^R = ID_i^{''} \oplus H2(sID_i^{'}).$$

After obtaining the real identity ID, RSU_j checks its own CRL to confirm whether OBU_i is legal, then selects a random integer n to calculate

$$S_{share} = H2((nP)_{S_{share}}).$$

It acts as a shared key between OBU and RSU, then it is sent to OBU_i.

2) OBU and RSU authenticate each other
OBU_i randomly selects *t1* from the finite field and then sends *t1* to RSU_j. RSU_j randomly selects *d, t2* from the finite field to calculate $N1 = Sign_{RSU_j} \times$

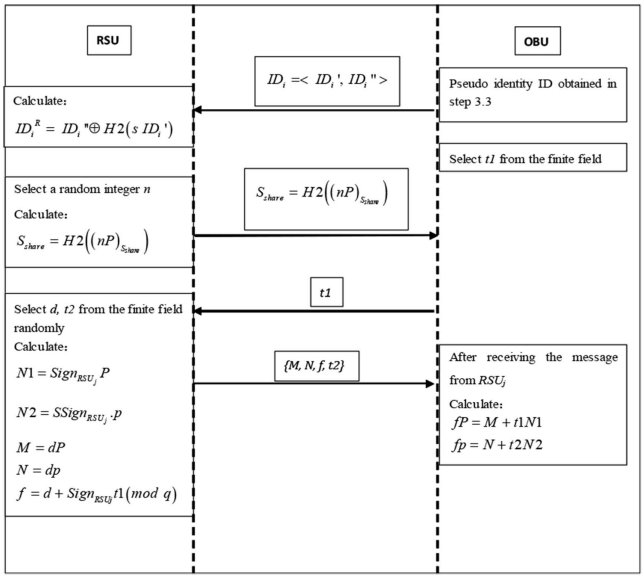


Figure 5: RSU and OBU authenticate mutually process diagram

$P, N2 = Sign_{RSU_j} \times p$ announces $N1$ and $N2$, calculates

$$\begin{aligned} M &= dP,\ N = dp \\ f &= d + Sign_{RSU_j} t1 (\text{mod } q) \end{aligned}$$

and then send $\{M, N, f, t2\}$ to OBU_i.

After receiving the message from RSU_j, OBU_i performs the following calculations:

$$fP = M + t1N1,\ fp = N + t2N2.$$

If the above equation is true, RSU_j is authenticated by OBU_i, and OBU_i will join the group of RSU_j to prepare for subsequent communication.

After RSU_j and OBU_i mutual authentication are completed successfully, OBU_i decrypts the shared key. The process diagram is shown in Figure 5.

3.5 RSU Generates a Temporary Key

RSU_j uses its primary private key to generate the temporary private key at the timestamp T_s, computes $S_{RSU_j}^{T_s} = H2(S_{RSU_j} \| T_s)$, generates the corresponding public key $P_{RSU_j}^{T_s} = S_{RSU_j}^{T_s} P$, and then broadcasts the public key. According to the elliptic curve encryption algorithm $E_{S_{share}}(S_{RSU_j}^{T_s}, Cert_{RSU_j}^{T_s})$, the private key is encrypted and the $\{nID_i^{'}, E_{S_{share}}(S_{RSU_j}^{T_s}, Cert_{RSU_j}^{T_s})\}$ is sent to OBU_i.

3.6 OBU Signs and Transmits Messages

Considering the time validity problem, the timestamp T_s is introduced, at which the temporary key of RSU_j is generated. OBU_i uses the temporary key of RSU_j to generate the pseudo identity and its corresponding key at the timestamp, and then uses the generated identity and key to sign the message for transmission.

1) OBU_i decrypts the shared key S_{share}
After receiving the message from RSU_j, OBU_i calculates the $S_{share} = H2((nID_i^{'} r^{-1})_{S_{share}})$ to get the shared key.

2) OBU_i decrypts the temporary key of RSU_j
After OBU_i obtains the shared key, it uses $D_{S_{share}}(S_{RSU_j}^{T_s}, Cert_{RSU_j}^{T_s})$ to decrypt and obtain the temporary private key $S_{RSU_j}^{T_s}$ of RSU_j.

3) OBU_i calculates the temporary pseudo-identity
OBU_i selects a random integer g from the finite field and calculates

$$\begin{aligned} ID_{T_s}^{'} &= gP, \\ ID_{T_s}^{''} &= ID_i^R \oplus H2(gP_{RSU_j}^{T_s}), \\ ID_i^{T_s} &= \langle ID_{T_s}^{'}, ID_{T_s}^{''} \rangle, \end{aligned} \quad (2)$$

where $ID_i^{T_s}$ is the temporary pseudo identity.

4) Calculate the temporary private key of OBU_i
The temporary private key of OBU_i is calculated as follows:

$$\begin{aligned} S_{OBU_i}^{T_s} &= \langle S_{OBU_i^{'}}^{T_s}, S_{OBU_i^{''}}^{T_s} \rangle \\ S_{OBU_i^{'}}^{T_s} &= S_{RSU_j}^{T_s} ID_{T_s}^{'} \\ S_{OBU_i^{''}}^{T_s} &= S_{RSU_j}^{T_s} H1(ID_{T_s}^{'} \| ID_{T_s}^{''} \| T_s). \end{aligned} \quad (3)$$

5) Sign the message M:

$$C = S_{OBU_i^{'}}^{T_s} + H2(M) S_{OBU_i^{''}}^{T_s}. \quad (4)$$

$(ID_{T_s}^{'}, C, M, ID_i^{T_s})$ will be sent to the recipient. The process diagram is shown in Figure 6.

3.7 Identity Revocation

According to the previous, the tracking and revocation of the malicious node can be performed by the information in the trust center and RSU_j. The trust center can judge the real identity of the malicious node by sending the master key in the RSU_j message, as follows:

$$ID_{T_s}^{''} \oplus H1(S_{RSU_j}^{T_s} ID_{T_s}^{'}) = ID_i^R \quad (5)$$

The trust center adds ID_i^R to the CRL and then broadcasts the CRL among the RSUs, thus the malicious node can never be authenticated or communicate with other nodes.

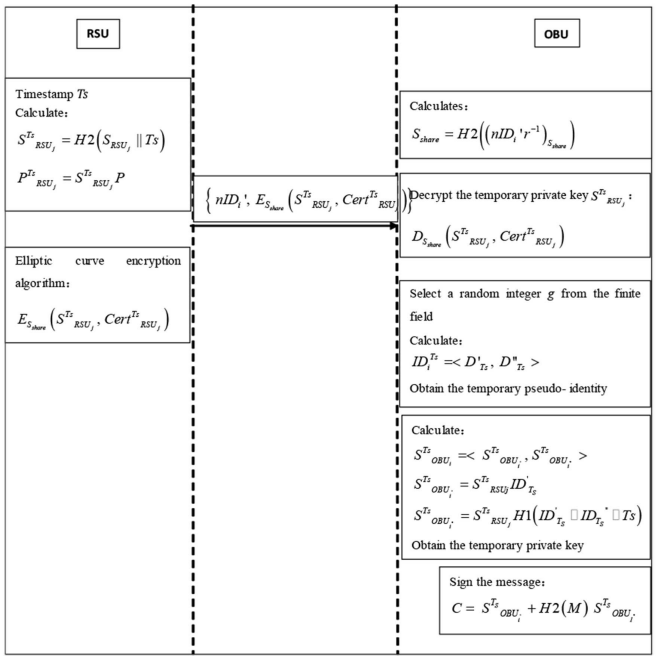


Figure 6: Sign and transmit messages process diagram

4 Performance Analysis

In this section, we will analyze the security of the scheme by means of formal analysis and simulation analysis whose simulation software is OMNET++.

4.1 Security Analysis

First of all, we analyze how the scheme meets the security requirements of IoV.

Non-forgery:

1) According to Equation (4), if we want to generate a valid signature, it is necessary for us to know the key S^{Ts}_{OBUi} which is generated by the temporary private key of the legal node RSU_j at timestamp T_s. This temporary key is encrypted with the Shared key and sent to the appropriate OBU along with the certificate. If we want to get S^{Ts}_{OBUi}, we need to know the shared secret S_{share} which is calculated by the random integer n. According to the mathematical difficult problems ECDLP, it is very difficult to calculate n. Because OBU and RSU have authenticated each other mutually before sending the message, the reliability between the communication nodes is high and the attacker cannot forge a signature.
2) If P, nID_i', rP are intercepted, according to Equation (1), the attacker needs to know r to calculate the pseudo-identity of OBU, so it can pass the RSU authentication, otherwise, it will be added to the revocation list. Considering the difficulty of ECDLP, if the user is malicious, the timestamp is invalid, which guarantees the unforgeability of the message.

Privacy protection:

1) According to the elliptic curve encryption algorithm and bilinear mapping theory, there are three basic difficult problems (ECDLP, BCDH, BDDH) that can guarantee the irreversibility of group operations, which makes it impossible for attackers to obtain the relevant certificate and key information through reverse engineering. When an OBU wants to get the temporary secret key of an RSU and join its communication group, it should use the hash functions to generate a temporary pseudonym identity ID according to the identity information and parameters of the trust center. Therefore, it can prevent the attacker from tracking the OBU as it moves between different RSUs.
2) RSU and OBU use the temporary keys in communication, while the generation of temporary keys uses the shared key generated in Section 3.4. According to Equations (2) and (3), the generation of temporary pseudo-identity uses the temporary keys S^{Ts}_{OBUi}, random integer g, and real identity ID_i^R . Moreover, as shown in Equations (3) and (4), the signature of a message uses different keys, and no node except the trust center and RSU can establish a relation between OBU_i*'s* pseudo-ID and the signature. Due to the mathematical difficult problems, the group operation is not reversible, so the privacy protection of the scheme is guaranteed.

Then we prove that our scheme satisfies security notions in the random oracle model.
Setting both sides of the game as attackers and challengers, the attacker is algorithm A running in polynomial time, the challenger is algorithm B, giving B a key exchange protocol input *(P, nP, bP, H)*, algorithm B uses A to solve the key agreement problem. The main steps are as follows:

System initialization: After selecting input *(P, nP, bP, H)*, algorithm B sends it to algorithm A.

Selection process: Algorithms A select ID_0^R and ID_1^R to Algorithms B.

Challenging process: Algorithms B randomly sets up a S_{share}, and then calculates the pseudo-identity of OBU:

$$\begin{aligned} ID_i^{'} &= nP, \\ ID_i^{''} &= ID_i^R \oplus H2(nbP), \\ ID_i &= \langle ID_i^{'}, ID_i^{''} \rangle. \end{aligned}$$

Guessing inquiry: A sends a guess about S_{share} to B. If the guess is right, then algorithm B solves the key agreement problem.

If $H = nbP$it can be calculated:

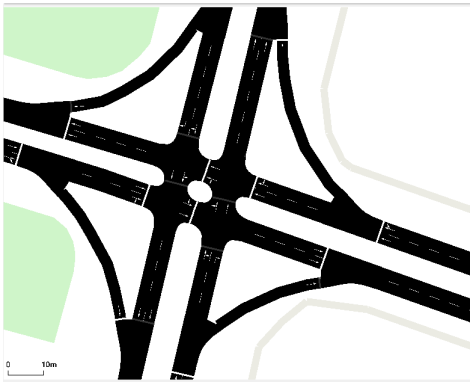
$$ID_x^{''} = ID_x^R \oplus H2(nbP) = ID_x^R \oplus H2(bID_x^{'}).$$

The probability of A guesses S_{share} is $1/2+\varepsilon$, then the probability of B's success is $1/2+\varepsilon$. Because H is randomly selected, ID_i^R it cannot be obtained, thus the probability of B solving the key agreement problem is $(1/2+\ \varepsilon\ +1/2)*1/2$, because ε it can be ignored, B can solve the problem. This is contrary to the assumption of key agreement difficulty and elliptic curve difficulty, so the scheme satisfies the privacy protection characteristics.

Traceability: Since all vehicles have been registered in the trust center and CA has the real identity of each vehicle, once a vehicle is attacked, the trust center can obtain the real identity information of the malicious node according to Equation (5), which ensures the traceability of the scheme.

Revocability: According to Section 3.7, when a vehicle is attacked and becomes malicious, the trust center can obtain its true identity and add it to the revocation list. When the malicious node initiates a communication again, RSUs will find that it is in the CRL and then refuse its communication request.

(a) Real view of the intersection



(b) Intersection Simulation

Figure 7: Map

4.2 Simulation Results

Our scheme is simulated by the use of Veins framework [21] which adopts SUMO as the transportation network platform and OMNET++ as the network simulation platform, respectively. Both platforms are based on C++. An RSU consists of three modules: appl, nic, and mobility. The nic module is based on the IEEE 802.11p protocol.

Due to the characteristics of large traffic flow and complex vehicle conditions, it is more meaningful for us to study the IoV authentication scheme at the crossroads. Therefore, we select a certain crossroad in Shijiazhuang urban area as the location of the simulation experiment. Figure 7(a) is a real scene map, and Figure 7(b) is a screenshot of the experimental environment. Taking this as an example, we analyze the delay, packet loss rate and signature efficiency of the scheme.

In this scheme, ECC is chosen as the main body of the cryptosystem. The security of ECC is based on the difficult problem of determining s with given sP and P, that is, the elliptic curve logarithm problem. The ECC key in the existing cryptosystem is short, the computation is small, the efficiency is high and the reliability is good.

1) Average delay: Because of the particularity of the IoV, messages are required to be transmitted as fast as possible. Thus, the time delay is an important index to measure the scheme performance. In our scheme, the time cost complexity is related to the time difference between message entering and quitting the Mac layer as follows:

$$D = \frac{\sum_{i=1} Sum(T_{out} - T_{in})}{Sum}$$

where Sum is the vehicle density, T_{out} and T_{in} are message entry and exit time, respectively. In this paper, the delay comparison is made between our scheme and ECPP scheme [15], Bayat's scheme [2] and Lee's scheme [10]. The experiment results are shown in Figure 8.

It can be seen from Figure 8 that the scheme performances have little difference when the vehicle density is small. With the increase of the vehicle density, the time delay of the authentication process increases accordingly and the performance differences of each scheme also increase gradually. In our scheme, on account of the higher efficiency of the adopted hash functions, the algorithm complexity is reduced. Therefore, the delay is lower and the performance is better compared with other schemes.

2) Packet loss rate: A too high packet loss rate will seriously affect data transmission. In this paper, the packet loss rate L of ECPP scheme [15], Bayat's

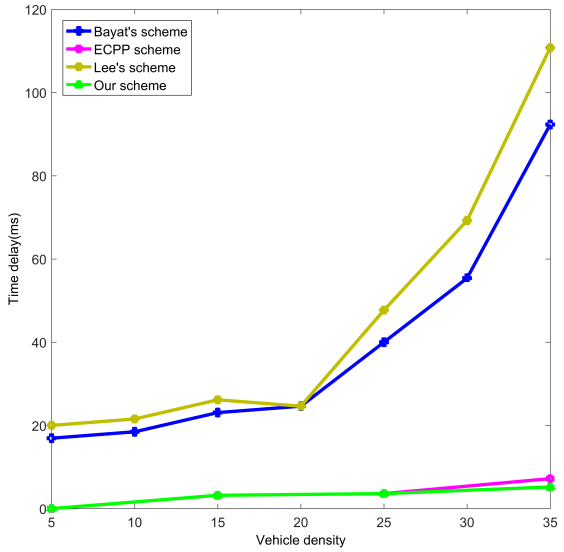


Figure 8: Relationship between delay and vehicle density

scheme [2] and Lee's scheme [10] are compared with that of our scheme. The calculation formula of the packet loss rate L is as follows:

$$L = \frac{\sum_{i=1}(M_r/M_l)}{Sum}$$

where Sum represents the density of communication vehicles, M_r is the total number of the received messages, and M_l is the total number of lost messages. Simulation results are shown in Figure 9.

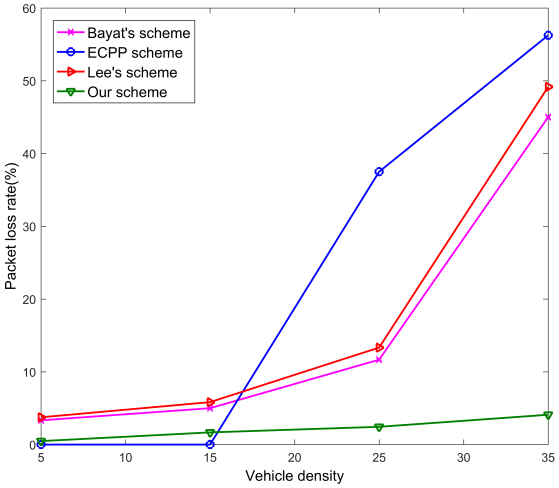
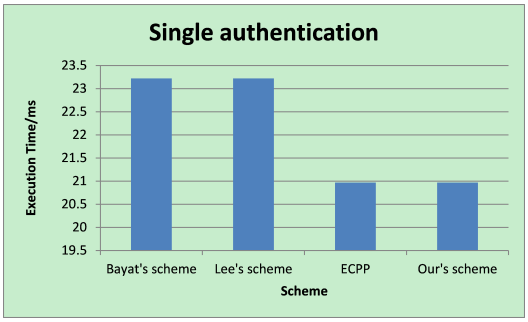


Figure 9: Packet loss rate

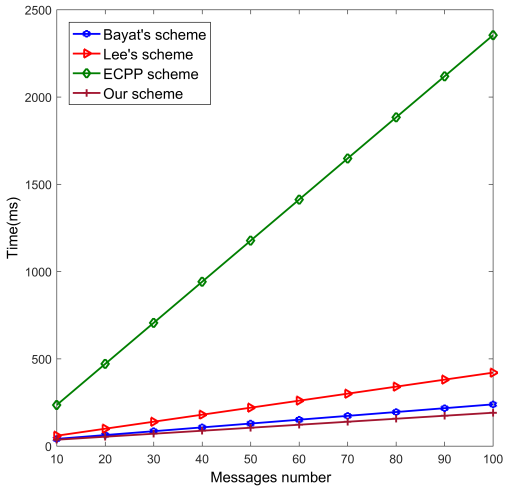
It can be seen from the figure that the performance of the scheme is similar when the vehicle density is small. With the increase of the vehicle density and the increase of communication load, ECPP and Bayat and Lee schemes have increased significantly. However, the packet loss rate of our scheme is still at a low level, mainly because the scheme reduces the computational complexity, reduces the delay and improves the efficiency of message processing.

3) Computational complexity: In this paper, the computational complexity of the scheme is compared with the computational complexity of the existing scheme. The computational formula is shown in Table 2.

The simulation results are shown in Figure 10. Figure 10(a) is the computational complexity of authenticating a single message and Figure 10(b) is the computational complexity of batch authentication.



(a) Single authentication



(b) Batch authentication

Figure 10: Authentication

The notations description in the formula is shown in Table 3 and the execution time is shown in Figure 11.

In this paper, the computational complexity of the scheme is compared with that of the existing scheme, as shown in Table 1, in which the description and execution time of each notation are shown in Table 2. According to the results, the complexity of this scheme is the same as ECPP, but lowers than Bayat's scheme and Lee's scheme; it is about 90% of these two schemes. With the increase of batch authentication messages, the scheme coefficient in this paper is 1.7177, which is less than that in other schemes, and the advantages are gradually obvious.

Table 2: Computational complexity comparison

Scheme	Single authentication	Batch authentication
Bayat's scheme	$3T_b+T_{bm}+T_H+T_h$	$3T_b+nT_{bsm}+3(n-1)T_{ba}+nT_H+nT_h$
Lee's scheme	$3T_b+T_{mul}+T_H$	$3T_b+nT_{bm}+3(n-1)T_{ba}+nT_H+nT_h$
ECPP scheme	$3T_b + T_{mul} + T_H$	$3nT_b+11nT_{mul}$
Our scheme	$3T_b+T_{mul}+T_H$	$3T_b+nT_{mul}+nT_H$

Table 3: Notation description

Notations	Description
T_b	Bilinear pairing operation
T_{bm}	Bilinear pair scalar multiplication
T_H	Map-To-Point hash function operation
T_h	One-way hash function operation
T_{bsm}	Bilinear pair small factor multiplication
T_{ba}	Additive operation
T_{mul}	Scalar Point Multiplication on Elliptic Curves

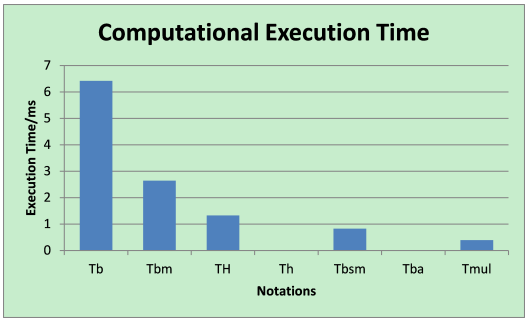


Figure 11: Computational execution time

The main reason is that the elliptic curve encryption and bilinear pairing used in our scheme has a shorter key and less work of computation. Besides, the Schnorr scheme is used to generate the certificate, and the main information of the system is stored in the RSU which has a large space. Therefore, our scheme is more efficient and stable than other schemes.

In this scheme, $Sign_{OBU_i}$ and $Sign_{RSU_j}$ are used to represent signatures. According to Schnoor signature algorithm, two parameters in group G_1 are used. In order to meet the security requirements of the system, the large prime q in the finite field is 170 bits and the element length in group G_1 is 171 bits, so the length of signature field is 171*2+170=512 bits, therefore the length of message is 173B, which is better than the existing scheme, which reduces communication overhead and improves authentication efficiency.

5 Conclusions

A novel identity-based authentication scheme is proposed in this paper. It is designed based on bilinear mapping theory, the elliptic curve encryption, and hash functions. At the beginning of communication, RSU and OBU authenticate each other through different algorithms to prevent malicious nodes from forging identities and to maximize the legality of the IoV nodes. Storing the main system information in RSU can be helpful to improve the efficiency of information exchange. Theoretical analysis proves that our scheme meets the requirements of non-forgery, privacy protection, traceability, and revocability. Simulation results show that our scheme has a smaller time delay, a lower packet loss rate, a lower computational complexity, and higher authentication efficiency compared with other schemes. Thus, it is more suitable for IoV applications [1].

In the following work, we will deeply study the impact of vehicle movement speed on message authentication, and further improve the privacy protection performance of the IOV.

Acknowledgments

The authors would like to thank the anonymous reviewers for their valuable suggestions given to improve the quality of the manuscript significantly. This work was supported by the National Natural Science Foundation of China under Grants No. 61572170 and No. 61672206, Program for Hundreds of Outstanding Innovative Talents in Higher Education Institutions of Hebei Province (III) under Grant No. SLRC2017042, Natural Science Foundation of Hebei Province of China under Grant No.F2018205162 and No.F2019205163, and Natural Sci-

ence Foundation of Hebei Normal University under Grant No.L072018Z10.

References

[1] T. Alam and B. Rababah, "Convergence of manet in communication among smart devices in IoT," *International Journal of Wireless and Microwave Technologies (IJWMT'19)*, vol. 9, no. 2, pp. 1–10, 2019.

[2] M. Bayat, M. Barmshoory, M. Rahimi, and M. R. Aref, "A secure authentication scheme for VANETs with batch verification," *Wireless Networks*, vol. 21, no. 5, pp. 1733–1743, 2015.

[3] M. Boban and A. Festag, "Service-actuated multichannel operation for vehicular communications," *Computer Communications*, vol. 93, pp. 17–26, 2016.

[4] E. F. Cahyadi, C. Damarjati, M. S. Hwang, "Research on identity-based batch verification schemes for security and privacy in VANETs", *Journal of Electronic Science and Technology*, vol. 18, 2020.

[5] T. Chim, S. Yiu, L. Hui, Z. Jiang, and V. O. K. Li, "Specs: Secure and privacy enhancing communications schemes for VANETs," *Ad Hoc Networks*, vol. 9, no. 2, pp. 189–203, 2011.

[6] J. Cui, J. Zhang, H. Zhong, R. Shi, and Y. Xu, "An efficient certificateless aggregate signature without pairings for vehicular ad hoc networks," *Information Sciences*, vol. 451-452, pp. 1–15, 2018.

[7] S. Hammad, R. A. Rehman, and B. S. Kim, "Services and security threats in sdn based VANETs: A survey," *Wireless Communications and Mobile Computing*, vol. 2018, no. 3, pp. 1–14, 2018.

[8] S. Horng, S. Tzeng, P. Huang, X. Wang, T. Li, and M. K. Khan, "An efficient certificateless aggregate signature with conditional privacy-preserving for vehicular sensor networks," *Information Sciences*, vol. 317, pp. 48–66, 2015.

[9] M. Inam, Z. Li, A. Ali, and A. Zahoor, "A novel protocol for vehicle cluster formation and vehicle head selection in vehicular ad-hoc networks," *International Journal of Electronics and Information Engineering*, vol. 10, no. 2, pp. 103–119, 2019.

[10] C. Lee and Y. M. Lai, "Toward a secure batch verification with group testing for VANET," *Wireless Networks*, vol. 19, no. 6, pp. 1441–1449, 2013.

[11] C. T. Li, M. S. Hwang, 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, July 2008.

[12] J. Li, Y. Lin, R. Li, S. Zhou, and S. Wang, "Secure anonymous authentication scheme based on elliptic curve and zero-knowledge proof in VANET," *Journal on Communications*, vol. 34, no. 5, pp. 52–61, 2013.

[13] X. Lin, X. Sun, P. Ho, and X. S. Shen, "Gsis: A secure and privacy-preserving protocol for vehicular communications," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 6, pp. 3442–3456, 2007.

[14] P. Liu, B. Liu, Y. Sun, B. Zhao, and I. You, "Mitigating dos attacks against pseudonymous authentication through puzzle-based co-authentication in 5G-VANET," *IEEE Access*, vol. 6, no. 99, pp. 20795–20806, 2018.

[15] R. Lu, X. Lin, H. Zhu, P. Ho, and X. Shen, "ECPP: Efficient conditional privacy preservation protocol for secure vehicular communications," in *IEEE Computer and Communications Societies*, pp. 1229–1237, April 2008.

[16] S. Manvi and S. Tangade, "A survey on authentication schemes in VANETs for secured communication," *Vehicular Communications*, vol. 9, pp. 19–30, 2017.

[17] H. Morita, J. C. N. Schuldt, T. Matsuda, G. Hanaoka, and T. Iwata, "On the security of schnorr signatures, dsa, and elgamal signatures against related-key attacks," *IEICE Transactions on Fundamentals of Electronics Communications and Computer Sciences*, vol. 100, no. 1, pp. 73–90, 2017.

[18] W. Ng, X. Zhou, X. Tian, X. Wang, and D. Yeung, "Bagging-boosting- based semi-supervised multi-hashing with query- adaptive re-ranking," *Neurocomputing*, vol. 275, pp. 916–923, 2017.

[19] Q. Pei, B. Kang, L. Zhang, K. R. Choo, Y. Zhang, and Y. Sun, "Secure and privacy-preserving 3d vehicle positioning schemes for vehicular ad hoc network," *EURASIP Journal on Wireless Communications and Networking*, vol. 2018, no. 1, p. 271, 2018.

[20] M. Raya and J. P. Hubaux, "Securing vehicular ad hoc networks," *Journal of Computer Security*, vol. 15, no. 1, pp. 39–68, 2007.

[21] C. Sommer, I. Dietrich, and F. Dressler, "Simulation of ad hoc routing protocols using omnet++," *Mobile Networks and Applications*, vol. 15, no. 6, pp. 786–801, 2010.

[22] C. Song, G. Tan, and N. Ding, "Rsu-coordinated multichannel mac protocol in vehicular ad hoc network (in chinese)," *Journal on Communications*, vol. 39, no. 11, pp. 10–22, 2018.

[23] C. Song, M. Zhang, W. Peng, Z. Jia, Z. Liu, and X. Yan, "Research on batch anonymous authentication scheme for VANET based on bilinear pairing," *Journal on Communications*, vol. 38, no. 11, pp. 35–43, 2017.

[24] A. Studer, F. Bai, B. Bellur, and A. Perrig, "Flexible, extensible, and efficient VANET authentication," *Journal of Communications and Networks*, vol. 11, no. 6, pp. 574–588, 2008.

[25] K. Verma, H. Hasbullah, and A. Kumar, "Prevention of dos attacks in VANET," *Wireless Personal Communications*, vol. 73, no. 1, pp. 95–126, 2013.

[26] A. Wasef and X. M. Shen, "Efficient group signature scheme supporting batch verification for securing vehicular networks," in *IEEE International Conference on Communications*, pp. 1–5, May 2010.

[27] H. Xiong, Z. Guan, Z. Chen, and F. Li, "An efficient certificateless aggregate signature with constant pair-

ing computations," *Information Sciences*, vol. 219, no. 10, pp. 225–235, 2013.

[28] X. Xue and J. Ding, "Lpa: A new location-based privacy-preserving authentication protocol in VANET," *Security and Communication Networks*, vol. 5, no. 1, pp. 69–78, 2012.

[29] L. Zhe and S. Hwajeong, "Iot-nums: Evaluating nums elliptic curve cryptography for iot platforms," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 3, pp. 720–729, 2019.

[30] H. Zhong, S. Han, and J. Cui, "Privacy-preserving authentication scheme with full aggregation in VANET," *Information Sciences*, vol. 476, pp. 211–221, 2019.

[31] X. Zhu, S. Jiang, L.Wang, L. Hui, and L. Zan, "Privacy-preserving authentication based on group signature for VANETs," in *IEEE Globecom Workshops*, pp. 4609–4614, Dec. 2013.

Biography

Changguang Wang is currently a professor in the College of Computer and Cyber Security of Hebei Normal University. His research interests include network and information security, wireless network security, IoV, *etc.*

Zimeng Dai is currently a Master degree student in the college of Computer and Cyber Security of Hebei Normal University. Her research interests include network and information security, sensor networks and IoV.

Dongmei Zhao is a professor at Hebei Normal University, Shijiazhuang, China. Her research interests include network and information security, network situation assessment, AI, *etc.*

Fangwei Wang is a professor at Hebei Normal University, Shijiazhuang, China. His research interests include network and information security, network worms,*etc.*

A Hybrid Framework for Security in Cloud Computing Based on Different Algorithms

Jannatul Ferdous[1], Md. Fuad Newaz Khan[1], Karim Mohammed Rezaul[2], Maruf Ahmed Tamal[1], Md. Abdul Aziz[1], and Pabel Miah[1]

(Corresponding author: Karim Mohammed Rezaul)

Department of Computer Science & Engineering, Daffodil International University, Bangladesh[1]
Faculty of Arts, Science & Technology, Wrexham Glyndŵr University[2]
Mold Rd, Wrexham LL11 2AW, United Kingdom
(Email: rezababu@gmail.com)

(Received Sept. 12, 2019; Revised and Accepted Jan. 3, 2020; First Online May 2, 2020)

Abstract

Cloud computing is the concept used to decode Daily Computing Issues. It is essentially a virtual pool of resources and also provides these tools to customers through the Internet. It is the net-based advancement and utilized in computer technology. The widespread problem connected with cloud computing is information privacy, protection, anonymity, and dependability, etc. However, the main issue involving them is safety and how the cloud supplier guarantees it. Securing the cloud means to secure the treatments (calculations) and storage (databases hosted by the Cloud provider). The paper reviews concurrent articles on security in cloud computing. By conducting research, we have managed to identify and analyze different security issues associated with the cloud as well as various cryptographic algorithms adaptable to better security for the cloud, and based on those algorithms, we have proposed a hybrid framework for security in cloud computing.

Keywords: AES; Cloud Coputing; DES; Security Attack; Steganography

1 Introduction

Cloud computing is a paradigm of information technology that provides scalable on-demand computing services such as computing, storage, network, software, and much more on the Internet [2] this enables businesses and organizations to concentrate their efforts on their key company or activity by outsourcing their IT resources [4, 27, 28]. This new technology offers many benefits such as cost efficiency, enhanced storage ability, backup and recovery, ongoing accessibility of resources and independence of location [9, 12]. Even though Cloud computing (CC) isn't entirely new, traction between organizations and individual users is still profiting. Transition into the cloud surroundings, however, isn't straightforward, and lots of operational and security problems exist.

The usage of a hypervisor and Virtual Machine (VM) technology is also a security problem since these and VM technology is vulnerable to VM level attacks. These programs have quite a few onsite computer organizations that might have a massive number of hardware and software systems. Vulnerabilities in VM infrastructure could be exploited by attackers to exfiltrate data or conduct attacks like DDoS (Distributed Denial of Services) [11,24]. This is a result of the inherent flaws in the TCP / IP stack. Additionally, several new strikes have emerged lately which use polymorphism and Metamorphosis to violate detection.

Attackers can inject kernel scripts into the server operating system (OS), and as all guest OS runs their OS with this kernel, attackers can command all VMs. Moreover, by successfully exploiting known or zero-day vulnerabilities in the hosted VM, the attackers may then obtain access to the host's VMs because the hypervisor shares the hardware and applications in the common virtual environment. Some hypervisors supply APIs which leave the VM facility entirely observable to Traffic however, these APIs provide additional paths for attackers to see and exploit network communication. Additionally, there are other strikes Such as information intrusion, information availability and data integrity targeting CC. Whenever consumers depend on these sides, information that is stored in the cloud is not reliable. Thus, nowadays users themselves are engaged in the process of encrypting their sensitive data before sending it for storage to the cloud [1, 18].

To demonstrate this, Gartner [10] predicts that 92% of workloads will be processed by cloud information centers by 2020. Cloud workloads are expected to rise 3.2 times over the same time, Cisco predicts [6]. However, the main drawback of this technology comes with the loss of control over the cloud infrastructure. Individuals, businesses and organizations, therefore, resist the adoption of pub-

lic clouds because of security and privacy concerns [5,19]. Recent cloud attacks, like the one in 2014 when Dropbox's 50 million user accounts were hacked which makes data security a hot topic.

2 Related Works

Cryptography is the main technique to secure data on clouds so that no one can steal our data and use it somewhere else or abuse it. Cloud computing confronts today's most common major issues with data integrity and confidentiality [25]. Cloud users store their information in different storage systems that cloud vendors provide. The problem, however, is that the user does not comprehend where the information is stored and has no control over it. Data acquisition is a technique or process of acquiring information from different hardware.

Cloud consumers and service suppliers should be acquainted with the information flow and Peer to Peer operations [20], how and where we access the data. For customers to collect their personal or confidential information in the cloud, data confidentiality is crucial. It's one of the major issues on the cloud. Cloud information is stored in remote places and cloud infrastructure used by providers to store information such as VM machine (image), copy and track logs or servers [20]. To divide the information and program, the client uses the shared storage. Due to attack, malicious action, and system failure, confidentiality problems sometimes arise. Hence, we want great safety processes and methods for fastening delicate information, unsecured transmission or storage. Ethics and Authenticity are another cloud security problem.

Integrity of information implies the provision of information from unauthorized deletion, production or modification. In standalone programs and databases, data integrity is easy; however, in cloud cases, it is hard because cloud suppliers work with countless databases, software, servers, and networks [26]. Authenticity refers to the practice of monitoring data and information accessibility. Only those consumers access the information that the supplier approves. Cloud is a data supply accessible, so many consumers have been facing the problem of permission and accessibility of information for a while. On the other hand, multi-tenancy indicates where computing tools, storage, services, and network shared by cloud technologies. It is a cost saving and provides better resource utilization. However, due to its data confidentiality due to shared resources, it is harmful. Many malicious activities ruin servers and community instruments, so it is not hard to control the flow of information or data (leakage). One of the following problems using multi-tenancy is a digital machine strike. Enhancing technologies and using networks provide people with a lot of equipment. It also improves many security issues, though; cyber-attack is just one of these.

Cyber-attacks use malicious code to alter computer code or information, resulting in damaging impacts that can undermine data and lead to cyber-crimes, including information and identity theft. Vulnerabilities of Shared Techniques make the cloud so intriguing is also a safety criticality point. As Navati et al. [23] showed that, attackers in the hypervisor could exploit vulnerabilities and achieve access to the physical host where other nearby virtual machines (VM) are located. Data from users can suffer from both accidental data loss and intrusive malicious behavior. Data loss is beyond the scope of this job, as we only consider data breaches here, i.e. stealing sensitive data (such as private or credit card details) [7]. A user can lose control over their own accounts. This enables the intruder to get into critical areas of a deployed service and possibly compromise the confidentiality, integrity, and availability of those services [7].

Denial of Service (DoS) is one of the most alarming scenarios is when the cloud infrastructure is made unavailable (just think that an outage costs Amazon 66 K dollars per minute). DoS in a cloud context is even more dangerous than in a traditional one since when the workload increases concerning a specific service, the cloud environment provides additional computational power to that service. This means that on the one hand, the cloud system counters the effects of the attack, but on the other hand, it supports the attacker in his evil activity, by providing him with more resources [8]. Malicious Insiders raise the list of top threats from the cloud. The chance of an insider being malicious – e.g. a worker possibly trying to take advantage of his privileged situation to access delicate data is becoming increasingly concrete and worrying [16].

3 Proposed Hybrid Model for the Security of Data

The algorithms emphasized in this research are: DSA (Digital Signature Algorithm), AES (Advanced Encryption Standard), and Steganography (hiding data behind an audio file or image).

DSA. The U.S. launched the Digital Signature Algorithm (DSA) in 1994. Digital signatures are extremely essential in the contemporary world to check the sender's identity. A digital signature is a digital signature that's used for confirmation and authentication of information. An electronic signature is represented as a series of binary digits in the pc system. The touch includes a set of parameters and rules (algorithm) like the identity of the individual signing the document in addition to the creativity of this information could be confirmed. The signature is created with the support of a personal key. A private key is known only to the sender. The signature is verified by the recipient using a public key that corresponds to the private key. A digital signature may be used with any sort of data, whether it's encrypted or not. Digital signatures are utilized to

detect unauthorized alterations of information by the third party. Additionally, the recipients of a digitally signed record assure that the record assigned by the person who it is promised to be signed up by. That is known as non-repudiation since the person who signed the record can't repudiate the signature afterward. Digital signature algorithms may be utilized in e-mails, electronic funds transfer, software distribution, data storage which guarantee the integrity, validity, and creativity of information. A hash function is being used in the signature generation process to get a condensed version of information, called a message digest [13].

AES. The Advanced Encryption Standard (AES) is the U.S. established electronic information encryption specification. National Standards and Technology Institute (NIST) in 2001 [3]. AES is based on a design principle known as a replacement-permutation network, combining both replacement and permutation, and is quick in both software and hardware [17]. Once DES was used as an encryption standard for over 20 years and it had been able to be deciphered in a relative short Quantity of time, NIST (United State National Institute of Standard and Technology) Chose a new Benchmark, the Advanced Encryption Standard (AES), had to be put into Position. AES is based on Rijndael cipher. AES was embraced by the US government and is popular nowadays. This decision was announced in January 1997, along with a petition for AES candidates had been created. The AES was to be a symmetric block cipher algorithm supporting keys sizes of 128-, 192-, and 256-bit keys. AES is based on substitution and permutation networks. It doesn't utilize the Feistel network. It's more secure than DES and difficult to crack. AES is much more complicated than DES, but it's quick and very effective. It operates with 128-bit fix block size plain text and version key sizes.

Steganography. Steganography is the science of concealing messages in this manner that nobody aside from the intended recipient knows of the presence of the message. Steganography is the practice of concealing just one medium of communication (text, audio or picture) in another. The term steganography comes from the Greek Steganos (covered or secret) and graphy (drawing or writing), and therefore it means, coated writing. Steganography is the tradition of encoding secret data in a way like the very existence of the info is hidden under the picture or picture in which it's hidden. Throughout history, many steganography techniques are recorded, for example, usage of cleverly-chosen words, invisible ink composed between traces, modulation of word or line spacing, and microdots. Normally the secret information is hidden by using an innocuous cover to arouse no distress to anybody. Edge of steganography over cryptography is that the key message doesn't draw attention to itself because the message could be hidden under picture file, video file, etc. [15].

There are various techniques available for steganography which are as follows:

1) Data hiding within wax pill;
2) Data hiding within noisy picture;
3) Hidden messages beneath sterile portion of some other message;
4) Data concealing inside a sound file;
5) Data concealing beneath video file.

4 The Proposed Scheme

4.1 Design of Proposed Framework

Figure 1 represents our framework's working process. At first, it will use DSA to create a digital signature, then it will use AES to encrypt the data of the user and to increase the security we have used steganography.

4.2 Elaboration of Overall System

There are some steps of the proposed scheme:

Step 1. Applying DSA for Generating Digital Signature:
A digital signature is a mathematical method used to validate a message, software or digital document's authenticity and integrity. A digital signature offers much more inherent security as the digital equivalent of a handwritten signature or stamped seal and is intended to solve the problem of manipulation and impersonation in digital communications.

The algorithm used for creating this signature is given Algorithm 1 [14].

Algorithm 1 DSA for Signature Creation

1: Input: Domain parameters (p, q, g); signer's private key a; message-to-be-signed M, with message digest $h = Hash(M)$.
2: Output: Signature (r, s).
3: Choose a random k in the range $[1, q-1]$.
4: Compute $X = g^k \bmod p$ and $r = X \bmod q$.
5: **if** r = 0 (unlikely) **then**
6: go to Step 3.
7: **end if**
8: Compute $k^{-1} \bmod q$.
9: Compute $h = Hash(M)$.
10: Compute $s = k^{-1}(h + ar) \bmod q$.
11: **if** $s = 0$ (unlikely) **then**
12: go to Step 3.
13: **end if**
14: Return (r, s).

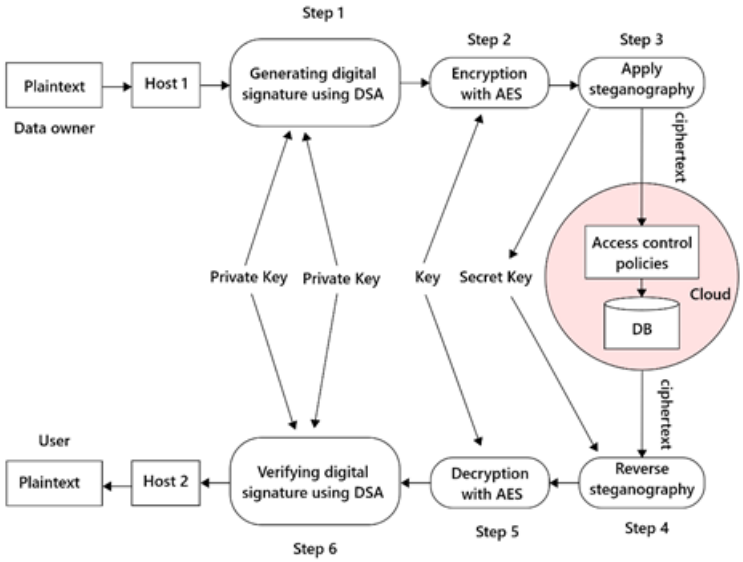


Figure 1: A hybrid framework for security in cloud computing

Step 2. Applying AES for Encryption:
AES is not a Feistel cipher, but an iterative one. It is based on the 'network of substitution – permutation'. It includes a series of linked operations, some of which involve replacing inputs with specific outputs (substitutions) and others involve shuffling bits (permutations) around them. Interestingly, AES performs all its computations on bytes rather than bits. AES therefore treats the 128 parts of the plaintext block as 16 bytes. These 16 bytes are arranged for matrix processing in four columns and four rows. The number of rounds in AES, unlike DES, is variable and depends on the key length. For 128-bit keys, AES utilizes 10 rounds, for 192-bit keys 12 rounds and 256-bit keys 14 rounds. Each round utilizes another 128-bit round key calculated from the initial AES key.

Step 3. Applying Steganography for Encryption:
Figure 2 shows a general structure for encryption using steganography. It is presumed that the sender wants to send a signal to a receiver through Steganographic transmission. The sender begins with a cover message, in which the integrated message is concealed, which is an input to the stego system. The message concealed is called the message integrated. A Steganographic algorithm combines the cover massage with the e1mbedded message, which is something to hide in the cover. The algorithm may or may not use a Steganographic key (stego key), which is additional secret data that may be needed in the hidden process [22].

Step 4. Applying Steganography for Decryption:
Usually, the same (or linked) key is required to retrieve the integrated massage. The Steganographic

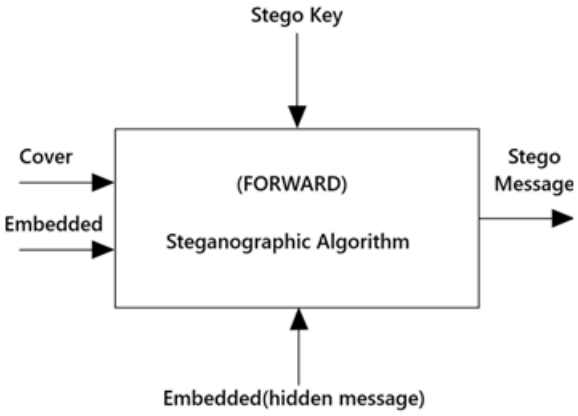


Figure 2: Encryption using steganography [21]

algorithm's output is the message of stego. The cover massage and stego text must be of the same sort of information, but another sort of information may be the integrated message. To extract the embedded signal, the receiver reverses the embedding process [22]. Figure 3 shows a general process of reverse steganography.

Step 5. Applying AES for Decryption:
The cipher text we got from Step 2 will have to be decrypted in this step by using AES again.

Step 6. Applying DSA for Verifying the Digital Signature:
The signature we created in Step 1 will have to be verified by using DSA in this step. Otherwise, the data will be lost.

The algorithm used for verifying this signature is

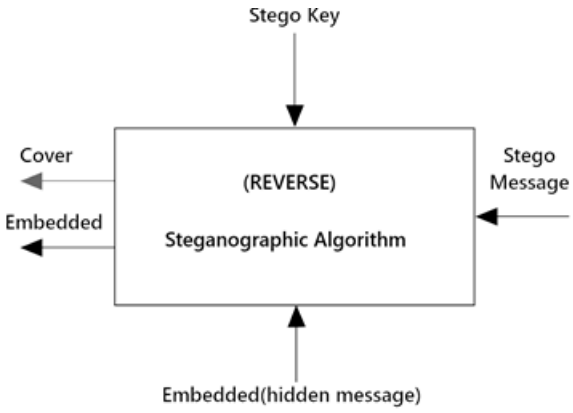


Figure 3: Decryption using steganography [21]

given in Algorithm 2 [14].

Algorithm 2 Reverse DSA

1: Input: Domain parameters (p, q, g); signer's public key A; signed message M; message digest $h = Hash(M)$; signature (r, s).
2: Output: Accept or Reject.
3: **if** $r, s \notin [1, q-1]$ **then**
4: return "Reject"
5: stop.
6: **end if**
7: Compute $w = s^{-1} \bmod q$.
8: Compute $h = Hash(M)$.
9: Compute $u_1 = hw \bmod q$ and $u_2 = rw \bmod q$.
10: Compute $X = g^{u_1} A^{u_2}$.
11: **if** $v = r$ **then**
12: return "Accept"
13: **else**
14: return "Reject".
15: **end if**

4.3 How the Data will be Secured in Our System

In our hybrid framework, in the 1st step, we will create a digital signature using DSA to verify the owner of the data, then in the 2nd step that data will be encrypted by AES algorithm which is the best cryptographic algorithm since time, and then in the 3rd step, that encrypted data will be again encrypted using STEGANOGRAPHY for an extra layer of security.

From the 4th step, our system will start to reverse the whole process to get the original data. In the 4th step, it will reverse the data through reverse STEGANOGRAPHY to get the encrypted file from Step 2 and then it will decrypt that data using AES again in Step 5 to get the original data afterward in the 6th step or in the final step, it will verify the data using reverse DSA again to verify the owner of that data.

5 Future Work and Conclusion

Data security is the most important issue of cloud computing in the IT industry. Future work includes implementing Digital Signature Algorithm (DSA), Advanced Encryption Standard (AES) and STEGANOGRAPHY to provide maximum security in cloud computing. By implementing these three algorithms, it is possible to provide authenticity, security and data integrity to data. We hope this work helps secure data from outsiders or hackers, who try to access and destroy the important data. We have located that the Time complexity is high because it is a one by one process, but in the future, this time complexity could be reduced. We will carry on this research in order to improve the functionalities of these algorithms in terms of robustness, reducing time complexity, hiding capacity and use other security algorithms or methods to protect information (data) on the cloud.

For many Internet facilities, cloud computing provides a flexible and cost-effective alternative. The contemporary sector of IT is focused entirely on online service or Internet facilities. This article outlines security problems along with cloud computing techniques and how they can be averted. Here we use techniques of cryptography and steganography together to protect information. Algorithms for DSA and AES are somewhat more secure than other algorithms. In order to give more data protection, we integrate AES and DSA with steganography. We get an encoded image in the steganography operation in which the human eye looks exactly the same as the initial image. While studying the binary picture codes, we could see the differences. Otherwise, the original picture cannot be spotted. The suggested hybrid framework for cloud computing security can facilitate to create a strong data security structure in the cloud computing region or the Internet.

References

[1] D. I. G. Amalarethinam, H. M. Leena, "Enhanced RSA Algorithm with varying Key Sizes for Data Security in Cloud," in *World Congress on Computing and Communication Technologies (WCCCT'17)*, pp. 172-175, 2017.

[2] D. R. Bharadwaj, A. Bhattacharya, M. Chakkaravarthy, "Cloud threat defense – A threat protection and security compliance solution," in *IEEE International Conference on Cloud Computing in Emerging Markets (CCEM'18)*, pp. 95-99, 2018.

[3] B. M. Belkaid, L. Mourad, C. Mehdi, "Meteosat Images Encryption based on AES and RSA Algorithms," *International Journal of Advanced Computer Science and Applications*, vol. 6, no. 6, pp. 203-208, 2015.

[4] C. A. B. D. Carvalho, R. M. D. C. Andrade, M. F. D. Castro, E. F. Coutinho, N. Agoulmine, "State of the art and challenges of security SLA for cloud

computing," *Computers and Electrical Engineering*, vol. 59, pp. 141-152, 2017.

[5] V. Casola, A. D. Benedictis, M. Rak, "On the adoption of security slas in the cloud," *Lecture Notes in Computer Science*, vol. 8937, pp. 45-62, 2015.

[6] Cisco, *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update*, Dec. 17, 2019. (https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white-paper-c11-738429.html)0

[7] L. Coppolino, S. D'Antonio, G. Mazzeo, L. Romano, "Cloud security: Emerging threats and current solutions," *Computers and Electrical Engineering*, vol. 59, pp. 126-140, 2017.

[8] R. V. Deshmukh, K. K. Devadkar, "Understanding DDos attack & its effect in cloud environment," *Procedia Computer Science*, vol. 49, pp. 202– 210, 2015.

[9] M. A. Fera, C. Manikandaprabhu, I. Natarajan, K. Brinda, R. Darathiprincy, "Enhancing security in Cloud using trusted monitoring framework," *Procedia Computer Science*, vol. 48, pp. 198-203.

[10] Gartner, *Gartner Forecasts Worldwide Public Cloud Revenue to Grow 17% in 2020, Nov. 13, 2019.* (https://www.gartner.com/en/newsroom/press-releases/2019-11-13-gartner-forecasts-worldwide-public-cloud-revenue-to-grow-17-percent-in-2020)

[11] Y. Gilad, A. Herzberg, M. Sudkovitch, M. Goberman, "CDN-on-Demand: An affordable DDoS defense via untrusted clouds," in *Proceedings 2016 Network and Distributed System Security Symposium*, 2016.

[12] W. F. Hsien, C. C. Yang and M. S. Hwang, "A survey of public auditing for secure data storage in cloud computing," *International Journal of Network Security*, vol. 18, no. 1, pp. 133-142, 2016.

[13] M. S. Hwang, C. C. Lee, J. L. Lu, "Cryptanalysis of the Batch Verifying Multiple DSA-type Digital Signatures", *Pakistan Journal of Applied Sciences*, vol. 1, no. 3, pp. 287-288, 2001.

[14] D. Ireland and DI Management Services Pty Limited, *Public key cryptography using discrete logarithms. Part 4: Digital Signature Algorithm (DSA)*, Aug. 22, 2019. (https://www.di-mgt.com.au/public-key-crypto-discrete-logs-4-dsa.html)

[15] S. S. Iyer, K. Lakhtaria, "New robust and secure alphabet pairing text Steganography Algorithm," *International Journal of Current Trends in Engineering & Research*, vol. 2, no. 7, pp. 15–21, 2016.

[16] M. Kandias, N. Virvilis, D. Gritzalis, "The insider threat in cloud computing," *Lecture Notes in Computer Science*, vol. 6983, pp. 93–103, 2013.

[17] S. Kulkarni, "Study of Modern Cryptographic Algorithms," *International Journal of Advanced Research in Computer Science*, vol. 8, no. 3, pp. 97-103, 2017.

[18] 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.

[19] J. Luna, N. Suri, M. Iorga and A. Karmel, "Leveraging the Potential of Cloud Security Service-Level Agreements through Standards," *IEEE Cloud Computing*, vol. 2, no. 3, pp. 32-40, 2015.

[20] M. V. Malakooti and N. Mansourzadeh, "A Two Level-Security Model for Cloud Computing based on the Biometric Features and Multi-Level Encryption," in *International Conference on Digital Information Processing, Data Mining, and Wireless Communications(DIPDMWC'15)*, pp. 100-111, 2015.

[21] A. W. Naji, S. A. Hameed, B. B. Zaidan, W. F. Al-Khateeb, O. O. Khalifa, A. A. Zaidan and T. S. Gunawan, "Novel framework for hidden data in the image page within executable file using computation between advanced encryption standard and distortion techniques," *International Journal of Computer Science and Information Security*, vol. 3, no. 1, 2009.

[22] A. W. Naji, A. A. Zaidan, B. B. Zaidan, S. A. Hameed and O. O. Khalifa, " Novel Approach of Hidden Data in the Unused Area 2 within EXE File Using Computation Between Cryptography and Steganography," *International Journal of Computer Science and Network Security*, vol.9, no.5, pp. 294-300, 2010.

[23] M. Nanavati, P. Colp, B. Aiello, A. Warfield, "Cloud security: a gathering storm," *Communications of the ACM*, vol. 57, no. 5, pp. 70–79, 2014.

[24] O. Osanaiye, K. K. R. Choo, M. Dlodlo, "Distributed denial of service (Ddos) resilience in cloud: Review and conceptual cloud Ddos mitigation framework," *Journal of Network and Computer Applications*, vol. 67, pp.147-165, 2016.

[25] V. K. Pant, J. Prakash, A. Asthana, "Three step data security model for cloud computing based on RSA and steganography techniques," in *International Conference on Green Computing and Internet of Things (ICGCIoT'15)*, pp. 490-494, 2015.

[26] V. K. Pant, Mr. A. Saurabh, "Cloud security issues, challenges and their optimal solutions," *International Journal of Engineering Research & Management Technology*, vol. 2, no. 3, pp. 41-50, 2015.

[27] S. Rezaei, M. A. Doostari, M. Bayat, "A lightweight and efficient data sharing scheme for cloud computing," *International Journal of Electronics and Information Engineering*, vol. 9, no. 2, pp. 115–131, 2018.

[28] C. Yang, Q. Chen, Y. Liu, "Fine-grained outsourced data deletion scheme in cloud computing," *International Journal of Electronics and Information Engineering*, vol. 11, no. 2, pp. 81–98, 2019.

Biography

Jannatul Ferdous was born in Chattogram, Bangladesh in 1997. He received the B.Sc. degree in Computer Science and Engineering department from Daffodil Interna-

tional University, in 2019.His research interests include information security, cloud security, data analysis based on the machine learning algorithm.

Md. Fuad Newaz Khan was born in Dhaka, Bangladesh in 1996. He received the B.Sc. degree in Computer Science and Engineering department from Daffodil International University, in 2019.His research interests include information security, cloud security, data analysis based on the machine learning algorithm.

Karim Mohammed Rezaul was awarded a PhD degree in Computing and Communications Technology from North East Wales Institute (NEWI) of Higher Education (presently Glyndŵr University), University of Wales, UK in October 2007. He received his BSc. degree in the field of Naval Architecture and Marine Engineering from Bangladesh University of Engineering and Technology (BUET), Dhaka in 1998 and, MSc. degree in Marine Technology from Norwegian University of Science and Technology (NTNU), Trondheim, Norway in 2001. In February 2002, Dr. Karim was appointed as a visiting lecturer in the department of Computing, Communications Technology and Mathematics at London Metropolitan University, and continued until June 2005. Presently, he is a Visiting Professor of Computing & Communications Technology at Wrexham Glyndwr University, UK and Adjunct Professor in Management at IPE Management School Paris, France. Since 2002, Prof. Karim has been working as an Academic advisor and Programme director of various International colleges in UK. Prof. Karim is a member of the Institute of Electrical and Electronics Engineers (IEEE), Association for Computing Machinery (ACM), Centre for Applied Internet Research (CAIR, UK), and a fellow of the Institution of Engineers Bangladesh (IEB, Bangladesh). He is the founder and director of Applied Research Centre for Business and Information Technology (ARCBIT) UK, Global Academy of Professionals (GAP) UK, Centre for Applied Research in Software and IT (CARSIT) Bangladesh, and Centre for Applied Research in Business, IT & Engineering (CARBITE) Bangladesh.

Prof. Karim is an author of a numerous Scientific and Business articles (Scholarly & Refereed publications) which include book, book chapters, journals and International conference papers. He is an editor of several international journals and member of the Technical Program Committee (TPC) of multiple International conferences. His research interests include IS Design and Development; ICT-based Pedagogy; Internet of Things (IoT); Artificial Intelligence (AI); Fractals and Nanotechnology; Data Science; Networking - Traffic Engineering, Quality of Service (QoS) Control, Traffic modelling & simulation etc.; Distributed DBMS; Information Security; Business Intelligence; E-Business/E-commerce; ICT Project Management; Computing.

Maruf Ahmed Tamal was born in Barishal, Bangladesh in 1996. He received the B.Sc. degree in Computer Science and Engineering department from Daffodil International University, in 2019.His research interests include Machine Learning, Data mining and Pedagogy.

Md. Abdul Aziz was born in Rajshahi, Bangladesh in 1995. He received the B.Sc. degree in Computer Science and Engineering department from Daffodil International University, in 2019.His research interests include information security, cloud security, data analysis based on the machine learning algorithm.

Pabel Miah was born in Tangail, Bangladesh in 1997. He received the B.Sc. degree in Computer Science and Engineering department from Daffodil International University, in 2019.His research interests include information security, cloud security, data analysis based on the machine learning algorithm.

A Revocable Certificateless Aggregate Signature Scheme with Enhanced Security

Fuxiao Zhou[1], Yanping Li[1], and Changlu Lin[2]
(Corresponding author: Yanping Li)
School of Mathematics and information Science, Shaanxi Normal University, Xi'an, Shaanxi 710119, China[1]
Fujian Normal University, Fuzhou, Fujian 350007, China[2]
(Email: lyp@snnu.edu.cn)
(Received Mar. 14, 2019; Revised and Accepted Sept. 3, 2019; First Online Sept. 21, 2019)

Abstract

In certificateless public key cryptosystem, a tough problem is how to revoke a user when the user's private key is compromised or expired. So the revocable certificateless schemes come into being. Certificateless aggregate signature (CLAS) is an efficient way to verify a large amount of signatures from different users simultaneously. However, none of CLAS schemes considers the user revocation currently. In this paper, we firstly demonstrate that an efficient certificateless aggregate signature (abbreviated to ECLAS) scheme proposed by Kang *et al.* is vulnerable to forged signature attack from the type II adversary by a concrete example, although they claimed that their scheme is existentially unforgeable against the adaptively chosen-message attacks. Furthermore, based on the ECLAS scheme and the revocable idea, we proposed a revocable certificateless aggregate signature scheme, which was proved to be existentially unforgeable against adaptive chosen-messages attacks under the hardness assumption of computational Diffie-Hellman problem. As far as we know, this is the first revocable CLAS scheme. Finally, numerical analyses and performance comparisons show our scheme saves computational cost, communication bandwidth and storage space than some related schemes.

Keywords: Certificateless Aggregate Signature; Cryptography; Existentially Unforgeable; Revocable

1 Introduction

In traditional public key infrastructure (PKI), a trusted entity called certificate authority (CA) often issues certificates to users by binding users true identities with their public keys. However, certificate management and authentication are quite complicated and expensive, which bring a heavy burden to CA in real-life. In 1984, Shamir first proposed the identity-based public key cryptosystem (ID-PKC) [13] to overcome the heavy certificate management and deep dependence on CA in PKI. The motive of this proposal was to choose the unique identity information such as social security number, telephone number of each party as user's public key. However, it needs a trusted third party named private key generator (PKG) to generate the private key for user. Hence the PKG possesses the private key of each user and can sign messages on behalf of any user at will, which makes the key escrow being the biggest criticism of the ID-PKC system. In order to eliminate the above problems in PKI and ID-PKC, Al-Riyami and Paterson introduced a new paradigm named certificateless public key cryptosystem (CL-PKC) in 2003 [1]. In a CL-PKC, a user's private key consists of two components: a partial private key issued by key generation center (KGC) and a secret value selected by the user. Since the KGC has no access to the user's entire private key, CL-PKC is not subject to the key escrow problem [14]. Additionally, CL-PKC also does not need certificates to authenticate public keys. Therefore, the CL-PKC is currently recognized as a promising public key cryptosystem.

Unfortunately, CL-PKC has the user revocation problem. It is well known that to revoke a user in PKC when the user's private key is compromised or expired is very cumbersome [2, 20]. The same problem inevitably exists in the CL-PKC and it gets more complex because the user's ID (*i.e.*, the public key) cannot change frequently. A previous revocation solution in CL-PKC was to use an on-line mediator called security mediator (SEM) [22]. In this kind of mechanism, the KGC divides a user's partial private key into two parts: One is delivered to the user and the other is delivered to the SEM. All these communications are conducted via secure channels, which greatly increase the communication costs. Later, Shen *et al.* [15] and Tsai *et al.* [17] successively presented two revocable certificateless encryption schemes. In both schemes, a user's private key consists of three parts: an initial partial private key, a time key and a secret value. The KGC controls the revocation of users by updating of the time key. It is noteworthy that the time key is renewed periodically over public channels by the KGC, which reduces

the need for secure channels and saves communication costs. Inspired by the idea used in [18], Sun *et al.* proposed the first revocable certificateless signature (RCLS) scheme [16], and soon after Zhang *et al.* put forward another improved RCLS scheme [24]. In both above RCLS schemes, the KGC generates a partial private key and a time key, where the time key is updated periodically. And the KGC just stops issuing the new time update key to revoke a user. Without the update time key, the user cannot sign a valid signature.

The notion of aggregate signature was introduced by Boneh *et al.* in 2003 [3]. Its primary focus is to aggregate n signatures on n messages from n users into a short signature, so a verifier can convince that the validity of n signatures by verifying the correctness of aggregate signature. Therefore, aggregate signatures greatly reduce the storage space, communication bandwidth and computational cost in verification and become a research hot spot. Combined with the prominent advantages of CL-PKC and aggregate signatures, a large number of certificateless aggregate signatures (CLAS) are put forward for various application scenarios [4–7, 10–12, 19, 21, 23, 25]. Gong *et al.* proposed two CLAS schemes to realize the aggregate signature scheme in CL-PKC [6]. However, Zhang *et al.* pointed out their schemes are insecure and proposed a new scheme and refined the security models [23]. In 2013, Xiong *et al.* put forward a CLAS scheme in [21] and claimed the scheme is more efficient than others. However, it was pointed out that an adversary could forge a legal signature for any message [7]. Li *et al.* proposed a novel and provably secure certificateless aggregate signature scheme in [11] and Nie *et al.* put forward a novel and efficient CLAS scheme [12]. Unfortunately, Nie's scheme was later proved that an adversary could forge any signer's signature on any message by obtaining a pair of message and its corresponding signature. Cui *et al.* proposed a CLAS scheme without pairings based on the elliptic curve cryptosystem [5]. Zhou *et al.* put forward a practical and compact certificateless aggregate signature with share extraction [25]. However, Chen *et al.* showed their scheme is in fact insecure against a type I adversary [4]. Wu *et al.* pointed out the CLAS scheme in [10] is vulnerable to signature forgery and proposed a new CLAS to fix the security flaws [19]. Recently, Kang *et al.* proposed an efficient CLAS scheme (ECLAS for short) and claimed their ECLAS scheme is existentially unforgeable against the adaptively chosen-message attacks [9]. In this article, we prove that the ECLAS scheme in [9] cannot satisfy the security they claimed by presenting a concrete example. As far as we know, there are no aggregation signature schemes with users' revocation at present. Therefore, we try to propose a revocable certificateless aggregate signature (RCLAS) scheme in this paper just in order to provide a secure revocation mechanism for CL-PKC-based aggregation signatures.

Our Contributions: In this paper, we propose a revocable certificateless aggregate signature (RCLAS) scheme. The contributions are summarized as follows:

1) Firstly, we demonstrate that the ECLAS scheme in [9] is not secure since it cannot resist the type II adversary. Specifically speaking, any type II adversary A_2 could forge any signer's signature on any message based on a valid signature, so that A_2 can forge a valid aggregate signature. At the same time, we analyze the reasons why the scheme is vulnerable to such attack and give the design principle of resisting this kind of attack;

2) Secondly, an improved scheme, namely, a revocable certificateless aggregate signature (RCLAS) scheme is proposed, which can revoke the user flexibly to meet the actual scenarios by using the time key. Then our RCLAS scheme is proven to be secure in the random oracle model under the hardness assumption of computational Diffie-Hellman problem (CDHP);

3) Thirdly, numerical analyses and performance comparisons demonstrate that our scheme has better performance than some existing schemes in [9, 16, 19]. Specifically, the length of aggregate signature in our RCLAS scheme only consists of two elements in G_1 which is far shorter than the aggregate signature in [19] and is independent of the number of signatures being aggregated. Additionally, the verification costs in the RCLAS scheme are relatively small.

The rest of this article is arranged as follows. In Section 2, some essential preliminaries are given. In Section 3, the ECLAS scheme is briefly reviewed and a specific attack on the ECLAS scheme is given. Our improved RCLAS scheme and its security proof are presented in Section 4 and Section 5, respectively. In Section 6, the performance of our scheme compares with some existing schemes. Finally, Section 7 concludes our paper.

2 Preliminaries

In this section, we introduce some necessary knowledge required in this paper.

2.1 Bilinear Pairing

Let G_1 be a cyclic additive group of prime order q and G_2 be a cyclic multiplicative group of the same order, P be a generator of G_1, $e : G_1 \times G_1 \to G_2$ be a bilinear map if it satisfies the following properties [8]:

1) Blinearity: $e(aP, bQ) = e(P, Q)^{ab}$,where $P, Q \in G_1$ and $a, b \in Z_q^*$;

2) Non-degeneracy: There exists $P \in G_1$, such that $e(P, P) \neq 1$;

3) Computability: It is efficient to compute $e(P, Q)$ for all $P, Q \in G_1$.

Definition 1. *Computational Diffie-Hellman problem (CDHP): Let G_1 be a cyclic additive group of prime order*

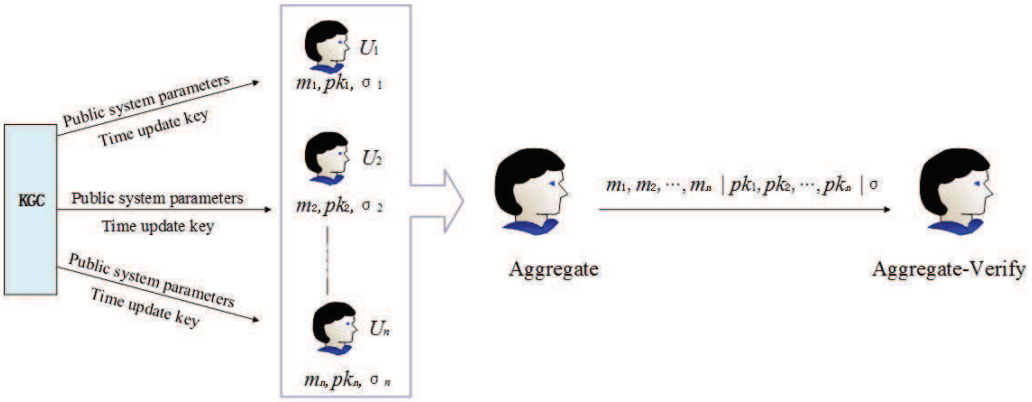


Figure 1: The proposed scheme

q and P be a generator of G_1. Given the elements P, aP and bP for the unknown $a, b \in Z_q^*$, it is hard to find abP.

2.2 Framework of a RCLAS Scheme

Generally, there are a KGC, n users and a signature aggregator in a RCLAS scheme, which consists of eight algorithms: Setup, Public-Key-Extract, Partial-Private-Key-Extract, Time-Key-Update, Private-Key-Extract, Sign, Aggregate, Aggregate Verify. The details of these algorithms will be described in Section 4 and not be repeated here because of the limit length. In the following, the system architecture of our RCLAS is given in Figure 1.

2.3 Security Model

In traditional RCLS schemes, three types of adversaries are considered. The type I adversary A_1 cannot obtain the master secret key msk, but can replace any user's public key, which describes an external adversary who did not know the msk.The type II adversary A_2 who has access to the msk but is unable to replace the user's public key, which depicts an internal adversary, such as the dishonest KGC. The type III adversary A_3 is used to describe the revoked malicious signers, who holds his/her partial private key and can replace other user's public key, but A_3 have no access to the msk and will no longer be issued the current time update key. Up to now, none of the existing revocable certificateless signature schemes can resist the collusion attack of KGC and revoked users. Hence, such attack is not considered in this article.

Definition 2. *The security model for the RCLAS scheme is defined by the following three games (**Game 1**, **Game 2** and **Game 3**) between a challenger C and three types of adversaries, respectively. The game details are given as follows.*

Game 1: A type I adversary A_1 interacts with the challenger C in this game. There are three phases in the game: **Setup**, **Queries**, **Forgery**.

Setup: C performs the setup algorithm that takes a security parameter l as input to obtain the master secret key msk and the system parameters $params$. Then C sends $params$ to A_1 while holds msk secret.

Queries: A_1 can perform a polynomially bounded number of the following types of queries in an adaptive way as follows:

- Hash queries: A_1 can request the hash values of any messages, C returns the corresponding results to A_1.
- Partial-Private-Key-Extract queries: When A_1 submits a private key query on an identity ID_i of a user U_i, C returns the corresponding private key D_i to A_1 by running the Partial-Key-Extract algorithm.
- Time-Key-Update queries: When A_1 submits a private key query, C runs the Time-Key-Extract algorithm to generate user's time key T_i and sends it to A_1.
- Public-key-Extract queries: When A_1 requests the public key of a user U_i with identity ID_i, C returns the corresponding public key pk_i by running the Public-key-Extract algorithm.
- Secret-Value-Extract queries: When A_1 requests the secret value of a user U_i with identity ID_i, C returns the corresponding secret value x_i by running Secret-Value-Extract algorithm. But note that A_1 is not allowed to ask for the secret value of a replaced public key.
- Public-key-Replacement queries: For any user U_i with identity ID_i, A_1 can select a new public key for the user U_i. C will record this replacement.
- Sign queries: When A_1 requests a user's signature query on a message m_i, C responds with the corresponding signature by running the Sign algorithm.

Forgery: The adversary A_1 outputs a tuple $(m^*, ID^*, t^*, w, \sigma^*)$ in which $t^* = (t_1^*, t_2^*, \cdots, t_n^*)$ was the expiration times, $m^* = (m_1^*, m_2^*, \cdots, m_n^*)$, $ID^* =$

$(ID_1^*, ID_2^*, \cdots, ID_n^*)$, w is a state information and σ^* is an aggregate signature. We say that A_1 wins **Game 1** if and only if:

1) σ^* is a valid aggregate signature on messages m^*.
2) At least one of the identities, without loss of generality, say $ID_1^* \in ID^*$ has never submitted during the Partial-Private-Key-Extract queries.
3) (m^*, ID^*, t^*, w) has never been submitted to the Sign queries.

Game 2: A type II adversary A_2 interacts with the challenger C in this game. There are three phases in the game: **Setup**, **Queries**, **Forgery**.

Setup: C performs the setup algorithm that takes a security parameter l as input to obtain the master secret key msk and the system parameters $params$. Then C sends the $params$ and msk to adversary A_2.

Queries: The adversary A_2 can perform a polynomially bounded number of queries as in **Game 1** in an adaptive way. Note that A_2 can make the Hash queries, Public-key-Extract queries, Secret-Value-Extract queries and Sign queries. But A_2 has no need to request the Partial-Private-Key-Extract queries and Time-Key-Update queries since the internal adversary A_2 who has access to the master secret key msk.

Forgery: The adversary A_2 outputs a tuple $(m^*, ID^*, t^*, w, \sigma^*)$ in which $t^* = (t_1^*, t_2^*, \cdots, t_n^*)$ was the expiration times,$m^* = (m_1^*, m_2^*, \cdots, m_n^*)$, $ID^* = (ID_1^*, ID_2^*, \cdots, ID_n^*)$, w is a state information and σ^* is an aggregate signature. We say that A_2 wins **Game 2** if and only if:

1) σ^* is a valid aggregate signature on messages m^*.
2) At least one of the identities, without loss of generality, say $ID_1^* \in ID^*$ has never submitted during the Secret-Value-Extract queries.
3) (m^*, ID^*, t^*, w) has never been submitted to the Sign queries.

Game 3: A type III adversary A_3 interacts with the challenger C in this game. There are three phases in the game: **Setup**, **Queries**, **Forgery**. It is worth noting that **Game 3** is very similar to **Game 1**, except that the conditions for the adversary to win the game are different. Details are given in the following.

Forgery: The adversary A_3 outputs a tuple $(m^*, ID^*, t^*, w, \sigma^*)$ in which $t^* = (t_1^*, t_2^*, \cdots, t_n^*)$ was the expiration times, $m^* = (m_1^*, m_2^*, \cdots, m_n^*)$, $ID^* = (ID_1^*, ID_2^*, \cdots, ID_n^*)$, w is a state information and σ^* is an aggregate signature. We say that A_3 wins **Game 3** if and only if:

1) σ^* is a valid aggregate signature on messages m^*.
2) At least one of the identities, without loss of generality, say $(ID_1^*, t_1^*) \in (ID^*, t^*)$ has never submitted during the Time-Key-Update queries.
3) (m^*, ID^*, t^*, w) has never been submitted to the Sign queries.

Definition 3. *A revocable certificateless aggregate signature scheme is said to be existential unforgeable against adaptive chosen-message attacks if no a probabilistic polynomial-time (PPT) adversary has non-negligible advantage in the above games (**Game 1**, **Game 2** and **Game 3**).*

3 Review and Security Analysis of the ECLAS Scheme

In this section, we briefly introduce the ECLAS scheme in [9] and give a specific attack.

3.1 Simple Review of the ECLAS Scheme

Setup: Given a security parameter l, the KGC picks two groups G_1 and G_2 with prime order q where G_1 is an additive cyclic group and G_2 is a multiplicative cyclic group, generates a generator P of G_1 and a bilinear map $e : G_1 \times G_1 \to G_2$, randomly chooses $s \in Z_q^*$ as a master secret key and calculates the system public key as $P_{pub} = sP$, chooses four cryptographically secure hash functions: H_1, H_3 and H_4: $\{0,1\}^* \to G_1$, H_2: $\{0,1\}^* \times \{0,1\}^* \times G_1 \times G_1 \to Z_q^*$. Finally the KGC keeps s secret and makes the $params = \{G_1, G_2, e, q, P, P_{pub}, H_1, H_2, H_3, H_4\}$ public.

Partial-Private-Key-Extract: The KGC calculates $Q_i = H_1(ID_i)$, $D_i = sQ_i$ and outputs D_i as the partial private key of user U_i with identity ID_i.

User-Key-Generate: By performing the following steps, a user U_i randomly selects $x_i \in Z_q^*$ as secret value and computes $P_i = x_i P$ as the public key.

Sign: Given a message m_i and a state information w, a user U_i with identity ID_i executes the following procedures to generate the signature:

1) Select randomly $r_i \in Z_q^*$ to compute $R_i = r_i P$.
2) Compute $T_i = h_i D_i + x_i Z + r_i F$, where $h_i = H_2(m_i, ID_i, P_i, R_i)$, $Z = H_3(w)$ and $F = H_4(w)$.
3) Output the signature $\sigma_i = (R_i, T_i)$ on the message m_i.

Aggregate: When receiving n message-signature pairs $\{(m_1, \sigma_1), (m_2, \sigma_2), \cdots, (m_n, \sigma_n)\}$ from n users $(U_1, U_2, \cdots, U_n)$, a signature aggregator calculates

$T = \sum_{i=1}^{n} T_i$ and outputs $\sigma = (R_1, R_2, \cdots, R_n, T)$ as an aggregate signature on the message $(m_1, m_2, \cdots, m_n)$.

Aggregate Verify: Given n users $(U_1, U_2, \cdots, U_n)$ with identities $(ID_1, ID_2, \cdots, ID_n)$, n corresponding public keys $(pk_1, pk_2, \cdots, pk_n)$, the state information w and the aggregate signature σ on the messages $(m_1, m_2, \cdots, m_n)$, the verifier takes the following steps:

1) Calculate $Q_i = H_1(ID_i)$, $h_i = H_2(m_i, ID_i, P_i, R_i)$ for all $i(1 \le i \le n)$, $Z = H_3(w)$ and $F = H_4(w)$.

2) Check whether the following equation holds or not. The aggregate signature is accepted if the equation holds, otherwise it will be invalid and refused.

$$e(T, P) = e(P_{pub}, \sum_{i=1}^{n} h_i Q_i) e(Z, \sum_{i=1}^{n} P_i) e(F, \sum_{i=1}^{n} R_i).$$

3.2 Security Analysis of ECLAS Scheme

The authors in [9] claimed that the ECLAS scheme is existentially unforgeable under adaptive chosen-message attacks against the two types of adversaries. In this subsection, we will prove the ECLAS scheme is insecure against the type II adversary by a concrete attack.

In **Game 2**, the type II adversary A_2 acts as a malicious KGC, it has access to the master secret key but cannot replace the public key of any user. Next, we show that how A_2 initiates an attack to forge a valid signature. The detailed steps are shown as follows.

First, suppose that A_2 intercepts a legal message-signature pair $(m_i, \sigma_i = (R_i, T_i))$, where $R_i = r_i P$, $T_i = h_i D_i + x_i Z + r_i F$, where r_i is a random value of Z_q^* and unknown to A_2.

Second, A_2 can compute $T_i' = T_i - h_i D_i$, where $h_i = H_2(m_i, ID_i, P_i, R_i)$. Since knowing the master secret key s, A_2 can compute the partial private key $D_i = sQ_i$ of the user with identity ID_i, where $Q_i = H_1(ID_i)$. Then A_2 computes $h_i' = H_2(m_i', ID_i, P_i, R_i')$, where $R_i' = R_i$, finally calculates $T_i'' = T_i' + h_i' D_i$.

Third, for a message $m_i'(m_i' \ne m_i)$, A_2 outputs the forged signature $\sigma_i' = (R_i, T_i'')$ on m_i'.

Obviously, the forged signature σ_i' is a valid signature on the message m_i' because it satisfies the equation $e(T_i'', P) = e(P_{pub}, h_i' Q_i) e(Z, P_i) e(F, R_i)$.

$$\begin{aligned} e(T_i'', P) &= e(T_i - h_i D_i + h_i' D_i, P) \\ &= e(T_i, P) e(h_i D_i, P)^{-1} e(h_i' D_i, P) \\ &= e(P_{pub}, h_i Q_i) e(Z, P_i) e(F, R_i) \\ &\quad e(h_i' D_i, P) e(h_i D_i, P)^{-1} \\ &= e(h_i D_i, P) e(Z, P_i) e(F, R_i) \\ &\quad e(h_i' D_i, P) e(h_i D_i, P)^{-1} \\ &= e(Z, P_i) e(F, R_i) e(h_i' D_i, P) \\ &= e(h_i' s Q_i, P) e(Z, P_i) e(F, R_i) \\ &= e(P_{pub}, h_i' Q_i) e(Z, P_i) e(F, R_i). \end{aligned}$$

Once intercepting n valid messages-signature pairs $(m_i, \sigma_i = (R_i, T_i))_{i=1}^n$, A_2 performs above attacks and forges n valid message-signature pairs $(m_i', \sigma_i' = (R_i, T_i'))_{i=1}^n$, where $(m_i' \ne m_i)_{i=1}^n$. Then A_2 outputs $T' = \sum_{i=1}^{n} T_i''$ and the forged aggregate signature $\sigma' = (R_1, R_2, \cdots, R_n, T')$. Obviously, since the individual equation $e(T_i'', P) = e(P_{pub}, h_i' Q_i) e(Z, P_i) e(F, R_i)$ holds, it is easily verified that the forged aggregate signature σ_i' is a legal signature by the following equation:

$$\begin{aligned} e(T_i', P) &= e((T_1'', T_2'', \cdots, T_n''), P) \\ &= e(T_1'', P) e(T_2'', P), \cdots, e(T_n'', P) \\ &= e(P_{pub}, h_1' Q_1) e(Z, P_1) e(F, R_1), \cdots, \\ &\quad e(P_{pub}, h_n' Q_n) e(Z, P_n) e(F, R_n) \\ &= e(P_{pub}, h_i' Q_i) e(Z, \sum_{i=1}^{n} P_i) e(F, \sum_{i=1}^{n} R_i). \end{aligned}$$

In conclusion, the ECLAS scheme is not secure as the authors claimed. In fact, A_2 is the most difficult to deal with in CL-PKC schemes since it knows the master secret key and can compute the partial private key D_i for any ID_i. By intercepting a legal signature $\sigma_i = (R_i, T_i)$ on message m_i, A_2 can create a new valid signature $\sigma_i = (R_i, T_i'')$ on the message $m_i'(m_i' \ne m_i)$ without changing R_i. The main reason is that a lot of variables in the linear expression of T_i are known or easy to compute for A_2. This is a taboo that must be avoided in designing an aggregate signature scheme.

4 A Revocable Certificateless Aggregate Signature Scheme

To resist the drawback of ECLAS and address the compromise or expiration of signing key, we put forward a revocable certificateless aggregate signature (RCLAS) scheme in this section. The RCLAS mainly consists of the following eight algorithms.

4.1 Setup

Input a security parameter l, the algorithm outputs two groups G_1 and G_2 with prime order q where G_1 is an additive cyclic group and G_2 is a multiplicative cyclic group, a generator P of G_1, a bilinear map $e : G_1 \times G_1 \to G_2$ and five cryptographically secure hash functions, where H_1, H_2, H_3 and $H_4 : \{0,1\}^* \to G_1$, $H_5 : \{0,1\}^* \to Z_q^*$. Next it randomly chooses $s \in Z_q^*$ as the master secret key and computes the system public key $P_{pub} = sP$. Finally, the KGC makes

$params = \{G_1, G_2, e, q, P, P_{pub}, H_1, H_2, H_3, H_4, H_5\}$ public while keeps s secret.

4.2 Public-Key-Extract

Without losing generality, assume U_i has identity ID_i. The user U_i selects a random value $x_i \in Z_q^*$ as secret value and takes x_i as input to compute $pk_i = x_i P$ as the public key.

4.3 Partial-Private-Key-Extract

The KGC generates the partial private key D_i for each user U_i with the corresponding public key pk_i by the following steps:

1) Calculate $Q_i = H_1(ID_i, pk_i)$.
2) Output $D_i = sQ_i$ and send D_i to the user by a secure channel.

4.4 Time-Key-Update

Given the identity ID_i of U_i, the corresponding public key pk_i and an expiration time t_i, KGC executes the following operations:

1) Compute $V_i = H_2(ID_i, pk_i, t_i)$ and the user's time update key $T_i = sV_i$.
2) Send T_i to the user and make (ID_i, t_i, T_i) public.

The reason to make (ID_i, t_i, T_i) public is that anyone can easily compute V_i and verify that T_i is actually a time update key on the identity ID_i and the time period t_i by checking whether the equation $e(P, T_i) = (V_i, P_{pub})$ holds. When the verification equation does not hold, it means that the user does not update his/her time key T_i in time.

4.5 Private-Key-Extract

A U_i generates his/her private key by taking D_i, T_i and x_i as inputs, calculates private key $sk_i = (D_i + T_i, x_i)$. The sk_i will update accordingly to the change of expiration time t_i.

4.6 Sign

Given a message m_i, a state information w (w can be current time, system parameters or arbitrary strings, which is selected and broadcasted to each signer by the aggregator, like the roadside unit RSU periodically broadcasting information in vehicular networks.), a non-revoked user U_i with private/public key sk_i/pk_i to execute the following procedures to generate a signature:

1) Randomly select $r_i \in Z_q^*$ to compute $U_i = r_i P$.
2) Compute $h_i = H_5(m_i, ID_i, U_i, pk_i, t_i, w)$, $F = H_3(w)$, $W = H_4(w)$, $R_i = h_i U_i$.
3) Compute $S_i = D_i + T_i + x_i F + h_i r_i W$.
4) Output the signature $\sigma_i = (R_i, S_i)$ on the message m_i.

4.7 Aggregate

When receiving n message-signatures pairs $\{(m_1, \sigma_1), (m_2, \sigma_2), \cdots, (m_n, \sigma_n)\}$ from n distinct non-revoked users $(U_1, U_2, \cdots, U_n)$ under the same state information w with the expiration times $(t_1, t_2, \cdots, t_n)$, a signature aggregator can calculate $R = \sum_{i=1}^{n} R_i$, $S = \sum_{i=1}^{n} S_i$ and output $\sigma = (R, S)$ as an aggregate signature on message $(m_1, m_2, \cdots, m_n)$.

4.8 Aggregate Verify

Given n users $(U_1, U_2, \cdots, U_n)$ with identities $(ID_1, ID_2, \cdots, ID_n)$, n public keys $(pk_1, pk_2, \cdots, pk_n)$, n expiration times $(t_1, t_2, \cdots, t_n)$, the state information w and the aggregate signature $\sigma = (R, S)$ on message $(m_1, m_2, \cdots, m_n)$, any verifier takes the following steps:

1) Calculate $Q_i = H_1(ID_i, pk_i)$, $V_i = H_2(ID_i, pk_i, t_i)$, $F = H_3(w)$, $W = H_4(w)$.
2) Check whether the following equation (1) holds or not. If the equation holds, the aggregated signature σ is regarded as valid, otherwise, $\sigma = (R, S)$ is considered as an invalid signature.

$$e(S, P) = e(\sum_{i=1}^{n}(Q_i + V_i), P_{pub}) e(\sum_{i=1}^{n} pk_i, F) e(R, W). \quad (1)$$

5 Security Proof

In this section, the security (including correctness and unforgeability) of our RCLAS scheme will be proven.

5.1 Correctness

$$\begin{aligned} e(S, P) &= e(\sum_{i=1}^{n}(D_i + T_i + x_i F + h_i r_i W), P) \\ &= e(\sum_{i=1}^{n}(D_i + T_i), P) e(\sum_{i=1}^{n} x_i P, F) e(\sum_{i=1}^{n} h_i r_i P, W) \\ &= e(\sum_{i=1}^{n} s(Q_i + V_i), P) e(\sum_{i=1}^{n} pk_i, F) e(\sum_{i=1}^{n} R_i, W) \\ &= e(\sum_{i=1}^{n}(Q_i + V_i), P_{pub}) e(\sum_{i=1}^{n} pk_i, F) e(R, W). \end{aligned}$$

5.2 Unforgeability

In this subsection, the security proof of our RCLAS scheme is proved under the hardness assumption of

CDHP. The **Theorems 1**, **2** and **3** show that the RCLAS scheme is secure against three types of adversaries in **Game 1**, **Game 2** and **Game 3**, respectively. Among the three types of adversaries, A_2 simulates an adversary who has known the master key s. Generally, A_2 has the strongest attack force and is the most difficult adversary to resist, so we mainly take **Theorem 2** as an example to show how our RCLAS scheme can achieve security under A_2 attacks. In the following, t_m represents the time for computing a scalar multiplication in G_1 and n is the size of the aggregate set.

Theorem 1. *In the random oracle model, if there is a type I adversary A_1 who has a non-negligible advantage ε in forging a valid aggregate signature of the RCLAS scheme in an attack model of **Game 1** within a time span t after making at most q_i times queries to the random oracles $H_i(1 \leq i \leq 5)$, q_{ppk} times Partial-Private-Key-Extract queries, q_{tk} Time-Key-Update queries, q_{pk} times Public-key-Extract queries, q_{rep} times Public-key-Replacement queries and q_{sig} times Sign queries, then the CDHP can be solved within time $t' \leq t + O[(2q_1 + q_2 + q_3 + q_4 + 2q_{ppk} + q_{tk} + q_{pk} + 5q_{sig} + n + 2)t_m]$ and with non-negligible probability $\varepsilon' \geq \frac{\varepsilon}{e(q_{ppk}+n)}$.*

Proof. The proof process is very similar to **Theorem 2**. The details of the proof process will be omitted here duo to the limit length. □

Theorem 2. *In the random oracle model, if there is a type II adversary A_2 who has a non-negligible advantage ε in forging a valid aggregate signature of our RCLAS scheme in an attack model of **Game 2** within a time span t after making at most q_i times queries to the random oracles $H_i(3 \leq i \leq 5)$, q_{pk} times Public-key-Extract queries, q_s times Secret-Value queries and q_{sig} times Sign queries, then the CDHP can be solved within time $t' \leq t + O[(q_3 + q_4 + q_{pk} + q_s + 5q_{sig} + n + 1)t_m]$ and with non-negligible probability $\varepsilon' \geq \frac{\varepsilon}{e(q_s+n)}$.*

Proof. Let the challenge C receives a random CDHP instance (P, aP, bP) in G_1, here P is a generator of G_1. A type II adversary A_2 interacts with C as modeled in **Game 2**, and we show that how C can use A_2 as a subroutine to find the solution abP to the CDHP instance. □

Setup: C firstly chooses a master secret key s and compute $P_{pub} = sP$. C selects system parameters $params = \{G_1, G_2, e, q, P, P_{pub}, H_1, H_2, H_3, H_4, H_5\}$, then sends the master secret key s and $params$ to A_2.

Queries: A_2 can perform a polynomially bounded number of the following types of queries in an adaptive manner. Hash functions H_1, H_2, H_3, H_4 and H_5 are considered as random oracles. All inquiries-responses will be kept in the corresponding lists. Since A_2 knows the master secret key s, it can compute all partial private keys and all time keys, so A_2 has no need to request the H_1 queries, H_2 queries, Partial-Private-Key-Extract queries and Time-Key-Update queries.

- H_3 queries: C maintains an initially empty list L_3 with structure (w, γ_i, F_i). When A_2 issues a query $H_3(w)$, the same answer will be given if the query has been asked before. Otherwise, C selects randomly $\gamma_i \in Z_q^*$, sets $F_i = \gamma_i aP$, adds (w, γ_i, F_i) to L_3 and returns F_i to A_2.

- H_4 queries: C maintains an initially empty list L_4 with structure (w, δ_i, W_i). When A_2 submits a query $H_4(w)$, the same response will be given if the query has been asked before. Otherwise, C picks randomly $\delta_i \in Z_q^*$, sets $W_i = \delta_i P$, adds (w, δ_i, W_i) to L_4 and returns W_i to A_2.

- H_5 queries: C maintains an initially empty list L_5 with structure $(m_i, ID_i, U_i, pk_i, t_i, w, h_i)$. When A_2 issues a query $(m_i, ID_i, U_i, pk_i, t_i, w)$ to H_5, the same answer will be given if the query has been asked before. Otherwise, C selects randomly $h_i \in Z_q^*$, adds $(m_i, ID_i, U_i, pk_i, t_i, w, h_i)$ to L_5 and returns the answer h_i to A_2.

- Public-Key-Extract queries: C keeps an initially empty list L_{pk} with structure (ID_i, x_i, pk_i, c_i). When A_2 performs a query with the identity ID_i to this random oracle, the same answer will be given if the query has been asked before. Otherwise, C first chooses a random value $x_i \in Z_q^*$ as the secret value, and then flips a coin $c_i \in \{0, 1\}$ that yields 0 with probability θ and 1 with probability $1 - \theta$. If $c_i = 0$, C computes $pk_i = x_i bP$ and adds $(ID_i, \perp, pk_i, c_i)$ to L_{pk}. If $c_i = 1$, C computes $pk_i = x_i P$, and adds (ID_i, x_i, pk_i, c_i) to L_{pk} and returns pk_i to A_2.

- Secret-Value queries: When A_2 performs a Secret-Value query on ID_i, C first makes a Public-Key-Extract query and finds (ID_i, x_i, pk_i, c_i) in L_{pk}. If $c_i = 1$, C computes $pk_i = x_i P$ and returns x_i to A_2. Or else, C returns $\perp$.

- Sign queries: When A_2 performs a Sign query on the tuple $(m_i, ID_i, w, pk_i, t_i)$, C executes the following operations to generate a valid signature:

1) If $c_i = 0$, C selects $r_i \in Z_q^*$ at random, sets $h_i = \delta_i^{-1}$ and $W_i = \delta_i aP$, computes $U_i = r_i P - \gamma_i pk_i$, $R_i = h_i U_i$ and $S_i = D_i + T_i + r_i aP$, finally returns the signature $\sigma_i = (R_i, S_i)$.

2) If $c_i = 1$, C runs the Sign algorithm normally to get a regular signature $\sigma_i = (R_i, S_i)$.

Forgery: In the end, suppose A_2 can output a tuple $(m^*, ID^*, t^*, w, \sigma^*)$ in which w is a state information, $m^* = (m_1^*, m_2^*, \cdots, m_n^*)$, $ID^* = (ID_1^*, ID_2^*, \cdots, ID_n^*)$, $t^* = (t_1^*, t_2^*, \cdots, t_n^*)$ and $\sigma^* = (R^*, S^*)$ is a valid forged aggregate signature. For

$1 \le i \le n$, C finds tuples of (w, γ_i, F_i), (w, δ_i, W_i) and $(m_i, ID_i, U_i, pk_i, t_i, w, h_i)$ from L_3, L_4 and L_5, respectively. C proceeds only if $c_1^* = 0, c_i^* = 1(2 \le i \le n)$. Otherwise, C aborts. If the forged signature $\sigma^* = (R^*, S^*)$ meets the above conditions, then satisfies Equation (1), we have

$$e(pk_1^*, F^*) = e(S^*, P)e(\sum_{i=1}^{n}(Q_i + V_i), P_{pub})^{-1} e(\sum_{i=2}^{n} pk_i^*, F^*)^{-1} e(R^*, W^*).$$

Where $pk_1^* = x_1^* bP, F^* = \gamma^* aP, W^* = \delta^* P$ and $pk_i^* = x_i^* P(2 \le i \le n)$. So it is easy for C to obtain the solution to the given CDHP instance:

Now, we analyze the probability to solve a CDHP by type II adversary A_2 in the polynomial bounded time. We analyze the three events for C to succeed:

E_1: C does not abort all the queries of Secret-Value-Extract queries.

E_2: A_2 generates a valid and nontrivial aggregate signature forgery.

E_3: E_2 occurs, $c_1^* = 0, c_i^* = 1(2 \le i \le n)$.

C succeeds if the above events happen, so $\varepsilon' = Pr[E_1 \wedge E_2 \wedge E_3]$. We can know that $Pr[E_1] \ge (1-\theta)^{q_s}, Pr[E_2|E_1] \ge \varepsilon, Pr[E_3|E_1 \wedge E_2] \ge \theta(1-\theta)^{n-1}$, thus $\varepsilon' = Pr[E_1 \wedge E_2 \wedge E_3] \ge (1-\theta)^{q_s}\varepsilon\theta(1-\theta)^{n-1} = \theta(1-\theta)^{q_s+n-1}\varepsilon$

When $\theta = \frac{1}{q_s+n}$, $\theta(1-\theta)^{q_s+n-1}$ is maximized at $\frac{1}{q_s+n}(1-\frac{1}{q_s+n})^{q_s+n-1}$. When q_s is sufficient large, this probability approaches $\frac{\varepsilon}{e(q_s+n)}$. So we can get $\varepsilon' \ge \frac{\varepsilon}{e(q_s+n)}$.

The running time for C is the sum of $A_2's$ running time, the time for C to response the queries and the time for C to compute the CDHP instance. During H_3 queries, H_4 queries, Public-key-Extract queries, Secret-Value queries and Sign queries, it needs 1, 1, 1, 1, 5 scalar multiplications, respectively. And during C computing the CDHP instance, it needs $n+1$ scalar multiplication, so $t' \le t + O[(q_3 + q_4 + q_{pk} + q_s + 5q_{sig} + n + 1)t_m]$.

From all of the above, C can solve the CDHP instance with non-negligible probability that contradicts to the intractability assumption of CDHP.

Theorem 3. *In the random oracle model, if there is a type III adversary A_3 who has a non-negligible advantage ε in forging a valid aggregate signature of the RCLAS scheme in an attack model of **Game 3** within a time span t after making at most q_i times queries to the random oracles $H_i(1 \le i \le 5)$, q_{ppk} times Partial-Private-Key-Extract queries, q_{tk} Time-Key-Update queries, q_{pk} times Public-key-Extract queries and q_{sig} times Sign queries, then the CDHP can be solved with non-negligible probability $\varepsilon' \ge \frac{\varepsilon}{e(q_{tk}+n)}$ and within time $t' \le t + O[(2q_1 + 2q_2 + q_3 + q_4 + 2q_{ppk} + 2q_{tk} + q_{pk} + 5q_{sig} + n + 2)t_m]$.*

Proof. The proof process is very similar to **Theorem 2**. The details of the proof process will be omitted here because of the limit length.

According to **Theorem 1**, **Theorem 2** and **Theorem 3**, we can conclude that there is no PPT adversary of any type can forge a valid aggregate signature of the proposed RCLAS scheme with a non-negligible advantage in polynomial time. Hence, our scheme is secure under the hardness assumption of CDHP.

6 Performance Comparisons

In this section, we make performance comparisons between our RCLAS scheme and the schemes in [9, 16, 19]. Due to the limited knowledge of the authors, no revocable certificateless aggregate signature scheme has been found so far. Therefore, this paper compares the revocable certificateless signature (RCLS) scheme [16], the ECLAS scheme [9] which has been analyzed in our Section 3, and the latest new certificateless aggregate signature (NCLAS) scheme [19] as the comparison schemes. In comparison, we omit the computations which take little time such as Hash for simplicity.

From Table 1, we can see that our RCLAS scheme has relatively little computation and shorter length of aggregate signature than other schemes while realizing the function of user revocation. Compared with [16], our RCLAS scheme adds the property of signature aggregation which can greatly improve verification efficiency and may enjoy better practicality. As for the length of the aggregate signature, our RCLAS scheme only consists of two elements in G_1, which is far shorter than the schemes in [9, 19] and greatly saves the communication costs and storage space. In addition, our RCLAS scheme can realize user's revocation flexibly by time update key for practical scenarios while the schemes in [9, 19] cannot. In general, our RCLAS scheme has better comprehensive performance (Note: In Table 1, Sign and A-V cost denote the computational cost of generation and verification of aggregate signature, respectively; A-S size represents the size of an aggregate signature; s and p mean the computational cost of scalar multiplication and a bilinear pairing operation, respectively; $|G_1|$ represents the bit length of an element in G_1; "$\surd$" means "support"; "$\times$" means "not support"; "—" means "not mentioned").

Here, we give a more intuitively quantitative analyses for schemes in [9,19] and our scheme. We adopt the experiment in [10], which observes processing time for the Tate pairing on a 159-bit subgroup of an MNT curve with an implanting degree 6 at an 80-bit security level, running on an Intel i7 3.07 GHz machine. Thus the time consumed by various operations is as follows: P is 3.21ms and S is 0.39ms. Suppose that n=100 in the Aggregate Verify phase, the comparisons of computational cost are shown in Figure 2. From Figure 2, we can see that the computational cost of the three schemes is equal in Sign phase, yet in the Aggregate-Verify phase, the computational cost of

Table 1: Comprehensive comparisons between related schemes

Scheme	Sign cost	A-V cost	A-S size	Revocation
RCLS in [16]	3s	—	—	√
NCLAS in [19]	4s	$3p + 2ns$	$(n+1)\|G_1\|$	×
ECLAS in [9]	4s	$4p + ns$	$(n+1)\|G_1\|$	×
RCLAS	4s	$4p$	$2\|G_1\|$	√

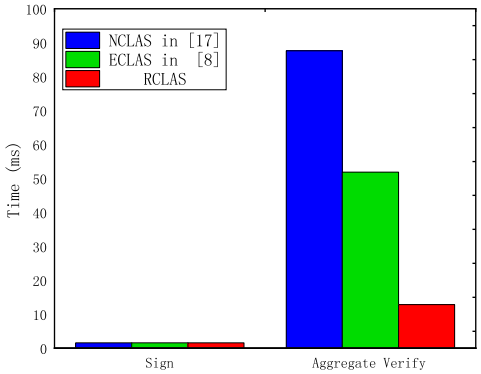


Figure 2: Computational cost comparisons

our scheme is much lower than other two schemes. Thus, the total computational cost of our scheme is reduced by 85.3, 75.2 percentage compared with those of schemes in [19] and [9], respectively.

7 Conclusion

In this paper, we first analyze the security of an efficient certificateless aggregate signature scheme (ECLAS) proposed in [9] and then give a specific attack. More specifically, any type II adversary A_2 can forge a valid aggregate signature on any set of messages as long as A_2 intercepts some legal message-signature pairs. In order to overcome this security flaw, we put forward an improved revocable certificateless aggregate signature (RCLAS) scheme, which not only can keep the advantages of aggregate signature, but also can flexibly deal with the problem of user's private key being compromised or expired in CL-PKC. The length of the aggregate signature in our RCLAS scheme only consists of two points in G_1 which is far shorter and greatly saves the communication cost and storage space. Finally, we show that our RCLAS scheme is proved to be existential unforgeable against adaptive chosen-message attacks under the hardness assumption of CDHP. And performance analyses show our RCLAS has better comprehensive performance while maintaining high computation and storage efficiency than some existing schemes.

Acknowledgements

This work was partly supported by the National Natural Science Foundation of China under Grant 61802243, U1705264; the Key R-D Program in industry

eld of Shaanxi Province under Grant 2019JY-013; the Fundamental Research Funds for the Central Universities under Grant2019CSLY002, GK201803005; the Natural Science Foundation of Fujian Province under Grant 2019J01275. The authors gratefully acknowledge the anonymous reviewers for their valuable comments.

References

[1] S. S. Al-Riyami and K. G. Paterson, "Certificateless public key cryptography," *Asiacrypt, Lecture Notes in Computer Science*, vol. 2894, no. 2, pp. 452–473, 2003.

[2] D. Boneh, X. Ding, G. Tsudik, and C. M. Wong, "A method for fast revocation of public key certificates and security capabilities," in *Conference on 10th Usenix Security Symposium*, 2001. (`https://ink.library.smu.edu.sg/cgi/viewcontent.cgi?article=2045&context=sis_research`)

[3] D. Boneh, C. Gentry, B. Lynn, and H. Shacham, "Aggregate and verifiably encrypted signatures from bilinear maps," in *International Conference on the Theory and Applications of Cryptographic Techniques*, pp. 416–432, 2003.

[4] C. Chen, H. Chien, and G. Horng, "Cryptanalysis of a compact certificateless aggregate signature scheme," *International Journal of Network Security*, vol. 18, no. 4, pp. 793–797, 2016.

[5] J. Cui, J. Zhang, H. Zhong, R. Shi, and Y. Xu, "An efficient certificateless aggregate signature without pairings for vehicular ad hoc networks," *Information Sciences*, vol. 451, pp. 1–15, 2018.

[6] Z. Gong, Y. Long, X. Hong, and K. F. Chen, "Two certificateless aggregate signatures from bilinear maps," in *The 8th ACIS International Conference on SPND*, vol. 3, pp. 183–193, 2007.

[7] D. He, M. Tian, and J. Chen, "Insecurity of an efficient certificateless aggregate signature with constant pairing computations," *Information Sciences*, vol. 268, pp. 458–462, 2014.

[8] 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.

[9] B. Kang, M. Wang, and D. Jing, "An efficient certificateless aggregate signature scheme," *Wuhan University Journal of Natural Sciences*, vol. 22, no. 2, pp. 165–170, 2017.

[10] P. Kumar, S. Kumari, V. Sharma, A. Sangaiah, J. Wei, and X. Li, "A certificateless aggregate signature scheme for healthcare wireless sensor network," *Sustainable Computing: Informatics and Systems*, vol. 18, no. 1, pp. 80–89, 2017.

[11] Y. Li, H. Nie, Y. Zhou, and B. Yang, "A novel and provably secure certificateless aggregate signature scheme," *Journal of Cryptologic Research*, vol. 2656, no. 7, pp. 526–535, 2015.

[12] H. Nie, Y. Li, W. Chen, and Y. Ding, "Nclas: A novel and efficient certificateless aggregate signature scheme," *Security and Communication Networks*, vol. 9, no. 6, pp. 3141–3151, 2016.

[13] A. Shamir, "Identity-based cryptosystems and signature schemes," *Workshop on the Theory and Application of Cryptographic Techniques*, vol. 196, pp. 47–53, 1984.

[14] S. Shan, "An efficient certificateless signcryption scheme without random oracles," *International Journal of Electronics and Information Engineering*, vol. 11, no. 1, pp. 9-15, 2019.

[15] L. Shen, F. Zhang, and Y. Sun, "Efficient revocable certificateless encryption secure in the standard model," *Computer Journal*, vol. 57, no. 4, pp. 592–601, 2014.

[16] Y. Sun, F. Zhang, and L. Shen, "A revocable certificateless signature scheme," *Journal of Computers*, vol. 9, no. 8, pp. 355–364, 2014.

[17] T.T Tsai and Y.M Tseng, "Revocable certificateless public key encryption," *IEEE Systems Journa*, vol. 9, no. 3, pp. 824–833, 2015.

[18] Y. M. Tseng and T. T. Tsai, "Efficient revocable id-based encryption with a public channel," *The Computer Journal*, vol. 55, no. 4, pp. 475–486, 2012.

[19] L. Wu, Z. Xu, D. He, and X. Wang, "New certificateless aggregate signature scheme for healthcare multimedia social network on cloud environment," *Security and Communication Networks*, vol. 2018, pp. 1–13, 2018.

[20] T. Y. Wu, T. T. Tsai, and Y. M. Tseng, "Revocable id-based signature scheme with batch verifications," in *The 8th International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pp. 49–54, 2012.

[21] H. Xiong, Z. Guan, Z. Chen, and F. Li, "An efficient certificateless aggregate signature with constant pairing computations," *Information Sciences*, vol. 219, no. 10, pp. 225–235, 2013.

[22] W. S. Yap, S. S. M. Chow, S. H. Heng, and B. M. Goi, "Security mediated certificateless signatures," *Lecture Notes in Computer Science*, vol. 4521, pp. 459–477, 2007.

[23] L. Zhang and F. Zhang, "A new certificateless aggregate signature scheme," *Computer Communications*, vol. 32, no. 6, pp. 1079–1085, 2009.

[24] Y. Zhang, C. Li, D. Zhou, and C. Wang, "Efficient revocable certificateless signature scheme," *Computer Engineering*, vol. 41, no. 7, pp. 157–162, 2015.

[25] M. Zhou, M. Zhang, C. Wang, and B. Yang, "Cclas: A practical and compact certificateless aggregate signature with share extraction," *International Journal of Network Security*, vol. 16, no. 3, pp. 174–181, 2014.

Biography

Fuxiao Zhou received her B.S. degree from Henan Normal University, Xinxiang, China, in 2017. She now is a M.S. degree candidate in Applied Mathematics with the School of Mathematics and Information Science, Shaanxi Normal University, Xi'an, China. Her research interests include certificateless signature and its applications.

Yanping Li received her M. S. degree from Shaanxi Normal University in 2004 and Ph. D degree from Xidian University in 2009, Xi'an, China. She now is an associate professor with the School of Mathematics and Information Science, Shaanxi Normal University. Her research interests include applied cryptography and its applications.

Changlu Lin received the B.S. and M.S. degrees in mathematics from Fujian Normal University, China, in 2002 and 2005, respectively, and the Ph.D. degree in information security from the State Key Laboratory of Information Security, Graduate University of Chinese Academy of Sciences, China, in 2010. He currently works with the College of Mathematics and Informatics, and the Fujian Provincial Key Laboratory of Network Security and Cryptology, Fujian Normal University. He is interested in cryptography and network security. He has conducted research in diverse areas, including secret sharing, multiparty computation, public key cryptography and their applications.

Forgery Node Detection Algorithm Based on Dynamic Reputation Value in the Internet of Vehicles

Peng-Shou Xie, Guo-Qiang Ma, Tao Feng, Yan Yan, and Xue-Ming Han
(Corresponding author: Guo-Qiang Ma)

School of Computer and Communications, Lanzhou University of Technology
No. 287 Lan gong ping road, Lanzhou, Gansu 730050, China
(Email: magq1514@163.com)

(Received Sept. 23, 2019; Revised and Accepted Jan. 15, 2020; First Online Feb. 5, 2020)

Abstract

Abnormal traffic messages may be sent by the internal forgery nodes to influence the normal behavior of other nodes in the Internet of Vehicles. However, the detection efficiency of the forgery nodes causing such attacks is generally low, and the accuracy of the detection algorithm is not high. Aiming at the above problems, the traffic messages published, forwarded and received by nodes are defined, and the effective features are extracted. On this basis, the forgery node detection model based on traffic messages is constructed, and the detection algorithm based on dynamic reputation value is designed. Finally, simulation experiments and performance analysis are completed. The results show that the time overhead of the detection algorithm is reduced, and the accurate detection rate of the detection algorithm is improved. It achieves the effect of quickly and accurately detecting the forgery nodes, and enhances the security of the Internet of Vehicles.

Keywords: Detection Algorithm; Detection Model; Dynamic Reputation Value; Forgery Node; Internet of Vehicles

1 Introduction

A special mobile ad hoc network is the Internet of Vehicles (IoV). Each vehicle is used as the message source to establish an information system that uses vehicles as nodes and communicates between people, vehicles and roads [19]. Its communication methods are mainly vehicles to vehicles (V2V) or vehicles to infrastructure (V2I). Road information (such as road congestion, collision accidents, *etc.*) is sent to nearby vehicles to realize timely sharing of traffic messages, in order to avoid potential accidents and enhance the safety of traffic roads [5]. Therefore, the Internet of Vehicles is widely used in the field of intelligent transportation. However, because of the wireless multi-hop communication, high mobility and the operation of vehicle nodes is limited, the security problem of the IoV is becoming more and more serious [11]. Among many security problems, abnormal and unreliable traffic messages by attackers are sent to surrounding vehicles to falsify traffic scenes, they damage the benefits brought by the application of the Internet of Vehicles, and even lead to more serious traffic accidents. This type of attack is called internal forgery node attack [21]. Then how to avoid the internal forgery node attack, ensure that the node can receive normal and reliable road traffic messages and select legal node to complete the service, which is one of the key issues in the research of the IoV.

In view of the safety of the IoV, many scholars at home and abroad have done a lot of research on this. At present, two types of detection schemes are proposed for internal attacks in the Internet of Vehicles.

1) Entity-based detection scheme. It is the judgment of the legal nodes through the communication between the node and other nodes. There are mainly identity-based authentication, trust-based evaluation and dynamic game-based schemes [14, 17, 20]. The advantages of such schemes are simple detection methods and low computational power requirements for processors. However, these schemes can only exert better detection performance when the number of normal nodes is more than the number of malicious nodes, otherwise its false detection rate is high.

2) Message-based detection scheme. It detects the abnormal message through the effective feature. There are mainly message authentication, deductive-based trust models, and message-based encryption schemes [2, 4, 15]. The advantage of such scheme is that it can avoid attacks caused by abnormal messages published or forwarded by the node. However, such these schemes can not detect and cull the nodes that send or forward abnormal messages, avoid continuing attacks in the future, and also has high time

overhead.

Therefore, in view of the shortcomings of the forgery node detection algorithm in the Internet of Vehicles, with the advantages of the existing two types of detection schemes, a forgey node detection algorithm based on dynamic reputation value (FNDA-IoV) is designed. Firstly, the effective features of the traffic messages of the Internet of Vehicles are described. The forgery node detection model of Internet of Vehicles is constructed. Secondly, the forgery node detection process of the Internet of Vehicles is extracted. Finally, the forgery node detection algorithm is designed to detect the internal forgery node of the Internet of Vehicles.

2 The Effective Features of Traffic Messages in the Internet of Vehicles are Described

The Internet of Vehicles plays an important role in traffic safety through the sharing and timely publishing of traffic messages. However, an open network environment, the complexity of road traffic, the numerous vehicle nodes and the fact that each node publishes or forwards a large number of various types of traffic messages at all times influence the security of the Internet of Vehicles [13]. In order to explain the problem more clearly, we define as follows:

Definition 1. *The traffic message type set is* $E=(e_1, e_2, e_3,..., e_n)$*, and* e_i *represents a certain type of traffic message published or forwarded by each node, such as emergency electronic brake lights (EEBL), post crash notification (PCN), road congestion notifyation (RCN),* etc.

Definition 2. *Traffic messages in the IoV can be divided into two types, namely* $\Theta = \{0, 1\}$*, where the "0" represents normal traffic messages and the "1" represents abnormal traffic messages. The normal traffic message refers to an instructive traffic message is published or forwarded by the legal node, and the abnormal traffic message refers to a malicious traffic message that is forged, falsified, published or forwarded by the forgery node. Here, it is assumed that the traffic messages published by the RSU are normal and trusted.*

Definition 3. *The set of vehicle nodes is* $V = (v_1, v_2,..., v_n)$*, vehicle nodes in the Internet of Vehicles broadcast traffic messages with digital signatures and public key certificates to other vehicle nodes in the process of traveling.*

Forgery node broadcasts abnormal messages means that the attacker changes the behavior of other nodes by publishing abnormal messages, tampering with real messages or injecting invalid messages [6,8,10]. For example, when a legitimate node receives a false alarm message, it may change its driving route, *etc.* As shown in Figure 1, the forgery node broadcasts an abnormal traffic message: the forgery node V_1 publishes or forwards an abnormal traffic message to the neighbor node V_2 to deceive the node V_2, and attempt to change the traveling path of the neighbor node V_2.

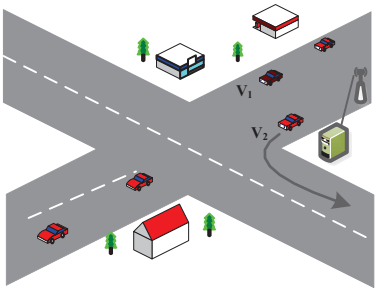


Figure 1: Forgery node broadcasts abnormal traffic messages

Since the features of traffic messages in the Internet of Vehicles have multiple dimensions, effective features (EF) in traffic messages are expressed in the form of a column vector, namely $EF = [x_1; x_2; x_3;...; x_n]$, then its corresponding data set (DS) can be expressed as $DS = \{(x_1, y_1), (x_2, y_2),..., (x_n, y_n)\}$, where $y_i \in \Theta$ represents the corresponding output result for the effective feature x_i, and n is the effective feature number. Table 1 lists the effective features in the traffic messages.

Therefore, the effective feature vector EF of the traffic messages in the Internet of Vehicles can be expressed as Equation (1).

$$EF = [e; d; v; a; t_0; s]. \quad (1)$$

3 Forgery node detection model in the Internet of Vehicles

The future behavior of vehicle nodes is uncertain, but the behavior trend of vehicle nodes can be predicted according to the historical behavior data of vehicle nodes. For this reason, the concept of trust is proposed in the nodes detection of Internet of Vehicles [9]. In human society, the trust is one of the most common concepts. Earlier, Mui *et al.* defined trust as follows: trust depends on experience and changes over time. When two people meet, their attitude towards each other is directly understood by the subject; The other is the recommender, neighbor node give recommendations based on own knowledge [1]. However, the evaluation of trust in social relations can also be carried out in the following ways: First, the subject directly determines the attitude toward the object according to the behavior of the object, and then feeds back its attitude to the third party who manages the subject and the object, and allows the third party to determine whether it continues to trust and whether it continues to exist in social relationships.

Therefore, referring to the trust evaluation method described above, the forgery node detection model as shown

Table 1: Effective features in traffic messages

number	feature	meaning
1	Sender(s)	The sender's identity type, including RSU (0) and vehicle node (1)
2	Time (t_0)	Timestamp of the sent traffic message
3	Direction(d)	The sender's direction of traveling
4	Vehicle (v)	The speed of the sender
5	Vehicle (a)	Sender's acceleration
6	Type(e)	Traffic message types, such as EEBL, PCN and RCN

in Figure 2 can be constructed. The model consists of three entities: A certificate authority (CA), a road side unit (RSU), and an on board unit (OBU) equipped with a vehicle. Among them, CA is responsible for distribution and revocation of certificates; RSU is responsible for publishing the normal and reliable traffic messages to vehicles within its communication scope; OBU is responsible for publishing, forwarding and receiving traffic messages [12].

However, in order to clearly illustrate the model, some connection parts are omitted here, such as the connection between the RSU and the CA. Where V_i is the node that publishes or forwards the traffic messages, and V_j is the node that receives the traffic messages.

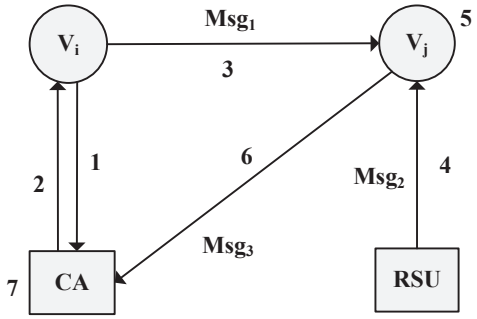


Figure 2: Forgery node detection model

Considering the non-repudiation of traffic messages, the effective feature of traffic messages, and the dynamics in the security requirements of the Internet of Vehicles, we design the communication message format as follows.

The communication messages between nodes are defined as follows:

$$Msg_1(Node\_Id_i,\ msgContent1_i).$$

Where $Node\_Id_i$ represents the node unique ID, and $msgContent1_i$ represents the traffic message sent by the node.

The communication messages sent by the RSU to the node is defined as follows:

$$Msg_2(Rsu\_Id_i,\ msgContent2_i)$$

Where Rsu_Id_i represents the RSU unique ID, and $msgContent2_i$ represents the traffic message sent by the RSU.

The feedback messages sent by the node to the CA is defined as follows:

$$Msg_3(Node\_Id_j,\ Node\_Id_i,\ \text{msgType}).$$

Where $Node\_Id_j$ represents the node unique ID of the receiving the traffic message, and $Node\_Id_i$ represents the node unique ID of the publishing or forwarding the traffic message. The $msgType$ is of the Boolean type, and the receiving node V_j informs the CA that the traffic message published or forwarded by the node V_i is normal (set to 0), or abnormal (set to 1).

Based on the above detection model, the seven steps are as follows:

Step 1. Node V_i requests a certificate from the CA. The node V_i needs to obtain communication and legal rights with other nodes in the network, and apply for a digital certificate to the CA according to its unique identity ID;

Step 2. The CA issues a certificate to V_i. The node V_i uses the digital certificate as an identifier that has communication authority in the network;

Steps 3, 4. Send the traffic message. The node V_i sends a traffic message $msgContent1_i$ to the node V_j, and the RSU sends a traffic message $msgContent2_i$ to the node V_j;

Step 5. Detect traffic messages. After receiving the traffic message of the node V_i, the node V_j starts detecting the traffic message according to the reliable traffic message sent by the RSU, and determines whether it is abnormal;

Step 6. Feedback to the CA. After the node V_j completes the detection of the traffic message published or forwarded by the node V_i locally, if the traffic message is normal, it is received; otherwise, it is discarded. At the same time, the node V_j sends a feedback message (Msg_3) to the CA;

Step 7. The CA updates the node reputation value (RV). The CA dynamically updates the reputation value of the node V_i according to the feedback message of the node V_j, and determines whether the node V_i is a forgery node.

4 Forgery Node Detection Process in the Internet of Vehicles

According to the above detection model, it can be seen that the forgery node detection process is as shown in Figure 3, and the specific detection steps are as follows.

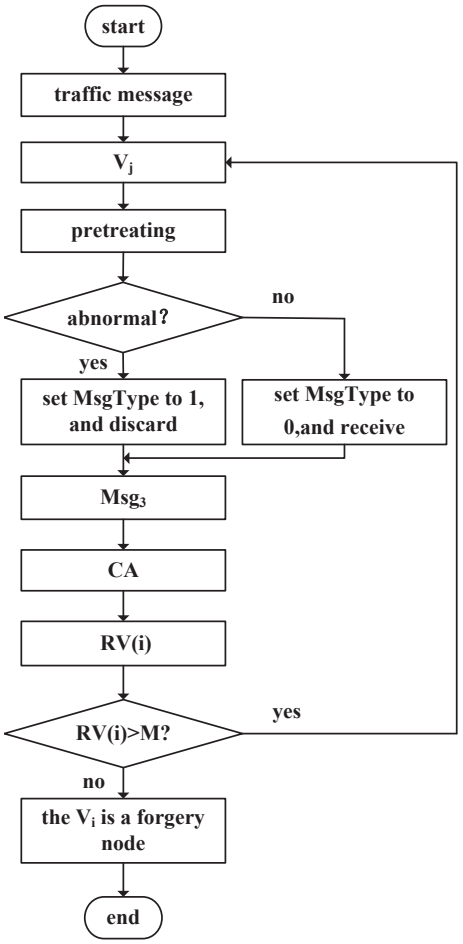


Figure 3: Forgery node detection process

Step 1. After receiving traffic message ($msgContent1_i$) and ($msgContent2_i$), the node V_j first preprocesses the traffic message according to the effective features of them;

Step 2. The node V_j performs a detection operation on the received traffic message. If it is a normal traffic message, it receives and sets the $msgType$ to 0. If it is an abnormal traffic message, it discards and sets the $msgType$ to 1. At the same time, the node V_j will send a feedback message (Msg_3) to the CA;

Step 3. After receiving feedback message, the CA updates the reputation value RV(i) of the node V_i, and compares the RV(i) with the threshold M, where M is the threshold of the node reputation, which can be set according to experience. If RV(i)>M, the CA continues to monitor its behavior; If RV(i)< M or RV(i)= M, then the node V_i is determined to be a forgery node, and the certificate issued to the node V_i is added to the revocation certificate list.

It can be seen that the detection process mainly includes three parts, namely, the traffic message is preprocessed, the traffic message is detected, and the node reputation value is dynamically updated.

4.1 Traffic Message Preprocessed

Preprocessing is to avoid the unnecessary computational overhead [22]. The traffic messages are preprocessed mainly from three aspects: the digital signature, the time validation, and the identity type verification. Firstly, the receiver verifies the integrity and non-repudiation of the traffic message by verifying the digital signature; then, using the batch authentication method to verify the timeliness, if the traffic message exceeds the time effective range, the traffic message is invalid, and the traffic message can be ignored. Finally, the traffic message sent by the RSU is used as a trained message, and the traffic message published or forwarded by the vehicle node is used as the detected message, as shown in Figure 4. The time validity of traffic messages is expressed as Equation (2):

$$t - t_0 < \Delta t. \tag{2}$$

Where t represents the time at which the node receives a traffic message, t_0 represents the time at which the traffic message was published or forwarded, and Δt represents the validity period of the traffic message.

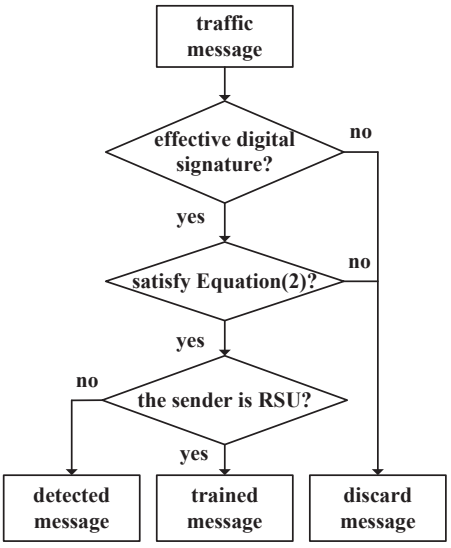


Figure 4: Traffic message preprocessed

4.2 Traffic Message Detected

At present, in the research field of intrusion detection algorithms, the main algorithms are the support vector machine (SVM), the clustering, naive bayes classifier(NBC),

the decision trees model (DTM), class association rules (CARS) and the deep learning [16]. The SVM is selected to realize the classification and detection of the traffic message in the Internet of Vehicles. The classification process is shown in Figure 5.

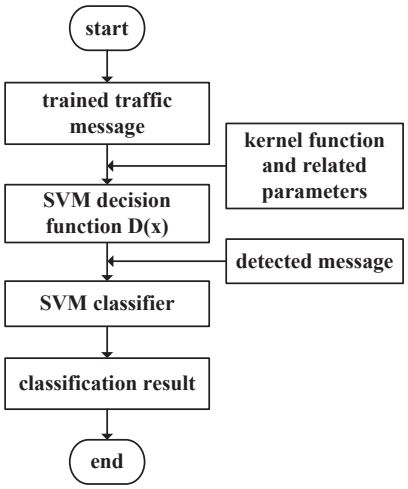


Figure 5: SVM classification process

The SVM algorithm classifies the traffic messages according to the effective feature vector EF of the multi-dimensional traffic message extracted by the Equation (1), and the classification result is a normal traffic message and an abnormal traffic message, where:
The SVM decision function is shown in Equation (3).

$$D(x) = sign[\sum_{i=1}^{n} \delta_i^* y_i K(x_i, x) + \theta^*]. \tag{3}$$

Where δ_i ($1 \leq i \leq$n) is the effective feature of the trained message x_i corresponds to the lagrangian factor, K($\bullet$) is the kernel function, and θ is the deviation.

The optimal classification hyperplane is shown in Equation (4).

$$\begin{cases} \phi(\omega, \varepsilon_i) = \frac{1}{2}||\omega||^2 + C\sum_{i=1}^{n} \varepsilon_i \\ y_i((\omega x_i) + b) \geq 1 - \varepsilon_i \qquad i = 1, 2, .., n \end{cases} \tag{4}$$

Where ε_i is the slack variable, C is the penalty factor, ω and b are the weight and threshold respectively, and n is the number of effective features of the traffic message.

4.3 Node Reputation Value Updated

After the traffic message is detected, the CA will automatically maintain a binary number based trust vector table to record the historical status of each node to publish or forward the traffic message based on the detection results of the node feedback. Currently, there are two methods for calculating the reputation value based on binary numbers: one is to calculate the reputation value according to the binary digital system, and the other is to calculate the reputation value by counting the number of the 0 or the 1 on the valid bit in the trust vector table [7].

Therefore, considering the features of the Internet of Vehicles, we design a new method for calculating the reputation value of the node by introducing the attenuation weights in combination with the existing two methods of calculating the reputation value.
The attenuation weight $g(k)$ represents the degree of attenuation of each bit in the binary valid bit in the trust vector table, and a valid bit represents a boolean judgment of the traffic message published or forwarded by node j to node i, the 1 and 0 respectively indicates an abnormal traffic message and a normal traffic message, and the conditions for satisfaction are as shown in Equation (5).

$$\sum_{k=1}^{m} g(k) = 1. \tag{5}$$

Where m is the number of the traffic messages communicated between nodes, and $0< g(k-1) < g(k) <1$.

Since the last calculated node reputation value should have different degrees of attenuation with traffic message detection time, the condition that the attenuation weight of the kth bit on the effective bit should satisfy is as shown in Equation (6).

$$g(k) = \frac{A}{t_t - t_k} \tag{6}$$

Where t_t is the current time, t_k is the time at which node j evaluates the kth the traffic messages sent by the node i, and A is the proportional coefficient.
Therefore, the calculation of the overall reputation value RV(i) of the node i can be expressed as shown in Equation (7).

$$RV(i) = 1 - \sum_{k=1}^{m} (\{1, 0\} * g(k)). \tag{7}$$

5 Forgery Node Detection Algorithm in the Internet of Vehicles

According to the above detection process, the designed the detection algorithm mainly includes:

Step 1. After receiving the traffic message, the node V_j preprocesses the message by using pre-processing function preTreat(), filters out the invalid traffic message, and verifies the identity of the sender. If the sender is an RSU, the traffic message to be sent is used as the trained traffic message. Otherwise, if the sender is a general vehicle node, the traffic message to be sent is used as the detected traffic message;

Step 2. The detected traffic message is sent as a parameter to the check() function, and traffic message is classified according to Equation (3) and Equation (4),

that is, when $D(x) = 0$, it is classified as a normal traffic message, and $msgType$ is set to 0. Otherwise, when $D(x) = 1$, it will be classified as an abnormal traffic message, and $msgType$ is set to 1;

Step 3. The vehicle node is determined by CA according to function isForgeryNode(), that is, if RV(i)> M, behavior is continuously monitored; otherwise, node V_i is determined to be the forgery node, and the certificate issued to node V_i is added to the revocation certifycate list. Among them, the main functions involved are:

1) PreTreat().
Preprocessing function.

```
Public void preTreat(String msg)
if (!imooc.jdkSign(msgNum)) then
  If the digital signature is incorrect, the
  node will discard the message discard();
else if (t − t0 > Δt) then
  /* If the Equation (2) is not met, the */
  /* message is discarded */
  discard();
else if (s==0) then
  /* If the sender is an RSU, the traffic */
  /* message is used as a training message */

  String trainMsg= msgContent2i ;
else
  /* If the sender is not an RSU, it is */
  /* used as a message to be detected. */
  String checkMsg = msgContent1i ;
end if
```

2) Check().
The traffic message detection function.

```
public static boolean check(String msg)
/* If it is a message sent by the RSU, the */
/* node trains it. */
print("start training.....");
String[] trainArgs={"msg2File"};
String[] modelFile=svmTrain.tain(trainArgs);

/* if it is not a message sent by the RSU, */
/* it is detected. */
print("start checkting.....");
String[] checkArgs={"msg1File"};
/* The node identifies and classifies the */
/* message according to Equation (3) */
/* and Equation (4).*/
Boolean result= modelFile.classify(checkArgs);
return result;
```

3) Update().
The reputation value update function.

```
public double update(String msg, boolean
msgType)
String nodeId=getNodeId(Msg1)
if (msgType == 0) then
  int[] Vtable= nodeIdVtable.vInsert(0);
else
  int[] Vtable= nodeIdVtable.vInsert(1)
  /* CA calculates the reputation value */
  /* of the node according to Equation (7).*/

  double RV(i)=sum(Vtable[i].g(k));
  return 1-RV(i)
end if
```

4) IsForgeryNode().
Forgery node decision function.

```
public boolean isForgeryNode (String rv)
if (rv > M) then
  return true;
else
  return false;
end if
```

6 Simulation and Analysis

Under the same conditions, the support vector machine, the decision tree model, class association rules and the naive bayes classifier are applied in the FNDA-IoV, and the performances of the detection efficiency and accurate detection rate of the four classification algorithms applied to the FNDA-IoV algorithm are compared.

6.1 Simulation Environment Configuration

A professional open source microscopic traffic simulation platform is the SUMO [3], two-way and six-lane highway environment is set, and experimental data is generated on the 6km road near the real vehicle driving position, speed, *etc.* Then a certain trace file is formed, and finally loads the network simulator NS2, the vehicle nodes is generated by reading the position, speed and other data of different vehicles at different times in the trace file [18]. Finally, the detection algorithm is simulated. The parameters are shown in Table 2.

6.2 Results and Analysis

Firstly, the simulation time is set to 50s, 100s, 150s, 200s, 250s, 300s to simulate the communication situation of each time period, and the node reputation value is calculated. Figure 6 shows the selected two nodes, namely the change in the reputation value of a legal node and a forgery node.

Figure 6 shows that with the passage of time, the reputation value of the legal node is rising. At 300s, the reputation value reaches 0.96, which is much higher than

Table 2: Simulation main parameters

type	name	value
Network scenes	Communication radius /m	300m
	MAC layer protocol	802.11p
	Simulation time	1000s
	Simulation area	1000m × 1000m
Traffic scenes	Number of lanes	6
	Number of nodes	200
	Number of forgery nodes	20
	Vehicle speed	20−60km/m
	Number of traffic messages	10 messages / vehicle
	Traffic message detection algorithm	SVM, NBC, DTM, CARS
	Reputation threshold	0.5

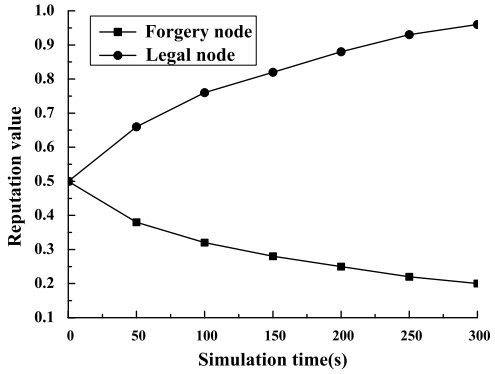


Figure 6: Node reputation value comparison

the threshold. The forgery node reputation value shows a downward trend. At 300s, the reputation value drops to about 0.2, which is far below the threshold. It can be seen that the FNDA-IoV algorithm can detect a forgery node whose reputation value is lower than the threshold.

Then, the detection overhead and the accuracy detection rate are analyzed.

1) The detection overhead. The detection overhead mainly measures the detection time required for detecting the forgery node, and the detection overhead is compared as shown in Figure 7. It can be seen that as the number of traffic messages published or forwarded by each vehicle is increasing, the detection time of the forgery nodes is gradually increasing.

   It is because the detection of the traffic messages published or forwarded by the vehicle takes time. The more traffic messages are detected, the longer the detection time required. However, The SVM has good recognition and generalization ability for non-linear and high dimensional data, and is suitable for traffic messages recognition and classification in the IoV. Therefore, the FNDA-IoV algorithm uses the SVM to detect traffic messages. Although the detection overhead of the algorithm is gradually increasing, it is slightly lower than a forgery node detection algorithm using the DTM, the CARS and the NBC.

2) The accuracy detection rate. The accurate detection rate refers to the probability that the detection algorithm can accurately detect the forgery node. The higher it is, the better the performance of the detection algorithm. The comparison results are shown in Figure 8.

As the number of the traffic messages published or forwarded by each vehicle is increasing, the accurate detection rate gradually is increasing. It is because the more messages published or forwarded by the node, the more accurate the judgment of the behavior of the node. At the same time, it can be seen that the FNDA-IoV algorithm designed by the SVM has a higher overall level than the NBC, the CARS and the DTM.

In summary, internal forgery nodes can be detected quickly and accurately by the FNDA-IoV algorithm in the Internet of Vehicles, and the security of the Internet of Vehicles is improved.

7 Conclusions

As a new type of wireless self-organizing network, Internet of Vehicles is well applied in the field of intelligent transportation. For the internal forgery node attack of the Internet of Vehicles, the detection efficiency and accuracy are improved by FNDA-IoV algorithm. However, the attack behavior of only a aspect of publishing or forwarding abnormal traffic messages is considered by the algorithm. The algorithm is considerd other aspects of the attack behavior, such as collusion between nodes, improved and optimized which will be the key tasks of the research work.

Acknowledgments

This research is supported by the National Natural Science Foundations of China under Grants No. 61862040,

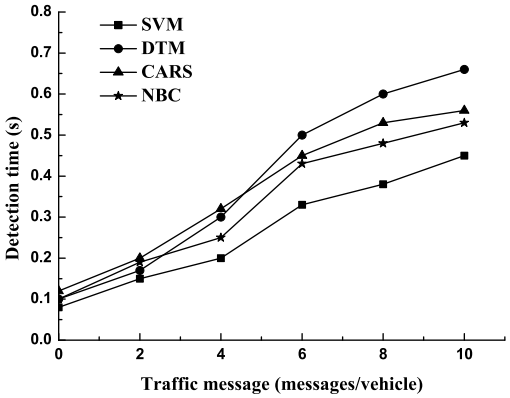


Figure 7: Detection overhead

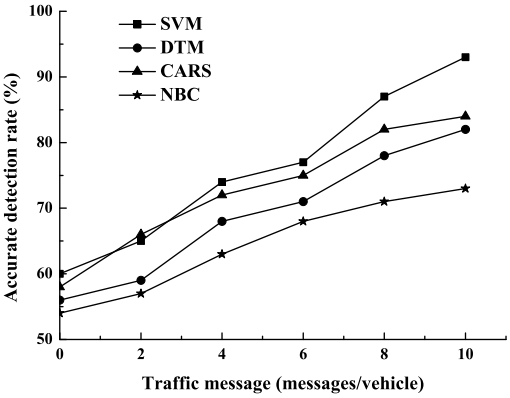


Figure 8: Accurate detection rate

No. 61762060 and No. 61762059. The authors gratefully acknowledge the anonymous reviewers for their helpful comments and suggestions.

References

[1] Z. Chen, L. Tian and C. Lin, "Trust evaluation model of cloud user based on behavior data," *International Journal of Distributed Sensor Networks*, vol. 14, no. 5, pp. 155–165, 2018.

[2] N. Fan and C. Q. Wu, "On trust models for communication security in vehicular ad-hoc networks," *Ad Hoc Networks*, vol. 90, pp. 101740, 2019.

[3] M. A. Hassan, U. Habiba and U. Ghani, "Asecure message passing framework for inter vehicular communication using blockchain," *International Journal of Distributed Sensor Networks*, vol. 15, no. 2, pp. 155–177, 2019.

[4] S. Ibrahim, M. Hamdy and E. Shaaban, "Towards an optimum authentication service allocation and availability in VANETs," *International Journal of Network Security*, vol. 19, no. 6, pp. 955–965, 2017.

[5] C. Khurana and P. Yadav, "Prevention of malicious nodes using genetic algorithm in vehicular ad hoc network," in *The Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC'18)*, pp. 700–705, 2018.

[6] L. Li and X. D. Li, "Detection mechanism of VANET abnormal nodes based on greenshield model," *Computer Engineering (in Chinese)*, vol. 44, no. 2, pp. 114–118, 2018.

[7] F. Li, Y. L. Si and Z. Chen, " Decision making method for opportunistic network security routing based on trust mechanism," *Journal of Software (in Chinese)*, vol. 29, no. 9, pp. 2829–2843, 2018.

[8] X. W. Liu and Y. L. Shi, "Detection of false traffic information in internet of vehicles based on weak classifier integration," *Journal of Communications (in Chinese)*, vol. 37, no. 8, pp. 58–66, 2016.

[9] Y. B. Liu , X. L. Song and Y. G. Xiao, "Car network authentication mechanism and trust model," *Journal of Beijing University of Posts and Telecommunications(in Chinese)*, vol. 40, no. 3, pp. 1–18, 2017.

[10] P. Ounsrimuang and S. Nootyaskool, "Classifying vehicle traffic messages from twitter to organize traffic services," in *The IEEE 6th International Conference on Industrial Engineering and Applications (ICIEA'19)*, pp. 705–708, 2019.

[11] S. Sharma and A. Kaul, "A survey on intrusion detection systems and honeypot based proactive security mechanisms in VANETs and VANET cloud," *Vehicular Communications*, vol. 12, pp. 138–164, 2018.

[12] Y. L. Shi and L. M. Wang, "A method for detecting anti-collusion sybil attack based on space-time analysis in VANETs," *Chinese Journal of Computers (in Chinese)*, vol. 41, no. 9, pp. 2148–2161, 2018.

[13] M. Sohail, L. Wang and S. Jiang, "Multihop interpersonal trust assessment in vehicular ad-hoc networks using three valued subjective logic," *IET Information Security*, vol. 13, no. 3, pp. 223–230, 2018.

[14] B. Subba, S. Biswas and S. Karmakar, "A game theory based multi layered intrusion detection framework for VANET," *Future Generation Computer Systems*, vol. 82, pp. 12–28, 2018.

[15] H. Tan, Z. Gui and I. Chung, "A secure and efficient certificateless authentication scheme with unsupervised anomaly detectionin VANETs," *IEEE Access*, vol. 6, pp. 74260–74276, 2018.

[16] S. Vhaduri and C. Poellabauer, "Multi-modal biometric-based implicit authentication of wearable device users," *IEEE Transactions on Information Forensics and Security*, vol. 18, pp. 955–965, 2019.

[17] Z. Wei and S. Yu, "Energy aware and trust based cluster head selection for ad-hoc sensor networks," *International Journal of Network Security*, vol. 20, no. 3, pp. 496–501, 2018.

[18] P. S. Xie, T. X. Fu and H. J. Fan, " An algorithm of the privacy security protection based on location service in the internet of vehicles," *International Journal of Network Security*, vol. 21, no. 4, pp. 556–565, 2019.

[19] Y. Xin and X. Feng, "A location dependent light weight sybil attack detection method in VANET,"

Journal on Communications (in Chinese), vol. 38, no. 4, pp. 110–119, 2017.

[20] W. Yang, M. R. Chen and G. Q. Zeng, "Cryptanalysis of two strongly unforgeable identity based signatures in the standard model," *International Journal of Network Security*, vol. 20, no. 6, pp. 1194–1199, 2018.

[21] W. Yu, Y. Li and Y. Xu, "Research on pseudo node detection algorithm in wireless sensor networks," *International Journal of Online Engineering*, vol. 13, no. 3, pp. 113–124, 2017.

[22] Q. Y. Zhang, W. J. Hu and S. B. Qiao, "Speech perceptual hashing authentication algorithm based on spectral subtraction and engery to entropy ratio," *International Journal of Network Security*, vol. 19, no. 5, pp. 752–760, 2017.

Biography

Peng-shou Xie was born in Jan. 1972. He is a professor and a supervisor of master student at Lanzhou University of Technology. His major research field is Security on Internet of Things. E-mail: xiepsh_lut@163.com.

Guo-qiang Ma was born in Jun. 1992. He is a master student at Lanzhou University of Technology. His major research field is network and information security. E-mail: magq1514@163.com.

Tao Feng was born in Dec. 1970. He is a professor and a supervisor of Doctoral student at Lanzhou University of Technology. His major research field is modern cryptography theory, network and information security technology. E-mail: fengt@lut.cn.

Yan Yan was born in Oct. 1980. She is a associate professor and a supervisor of master student at Lanzhou University of Technology. Her major research field is privacy protection, multimedia information security. E-mail: yanyan@lut.cn.

Xue-ming Han was born in Jan. 1990. He is a master student at Lanzhou University of Technology. His major research field is network and information security. E-mail: hxmhan@163.com.

Tripartite Authentication Protocol RFID/NFC Based on ECC

Yong-Shuang Wei and Jian-hua Chen
(Corresponding author: Yong-Shuang Wei)

Department of Mathematics and Statistics, Wuhan University
No. 299, Bayi Road, Wuchang 430072, Wuhan, China
(Email: yongshuang_wei@163.com)

(Received Mar. 9, 2019; Revised and Accepted Sept. 6, 2019; First Online Nov. 10, 2019)

Abstract

An RFID security three-party mutual authentication protocol based on elliptic curve cryptography (ECC) is designed in this paper. The proposed protocol not only satisfies most of basic characteristics of RFID system, such as mutual authentication, confidentiality, anonymity and others, but also resists tracking attack, denial of service attack, spoofing attack, *etc.* Being different from other RFID authentication protocols, our protocol is based on the assumption that the communication between reader and background is unsafe, so that tag, reader and background can mutually authenticate each other. In addition, the protocol provides a public secret co-negotiating key for the three participants to read and modify data in subsequent communication. According to the design of the protocol, it can apply to NFC system, which is evolved from the integration of RFID technology and interoperability technology. We further analyze the security of the protocol through the Burrows-Abadi-Needham logic (BAN-logic), which shows that the protocol can achieve mutual authentication and key agreement, as well as agree with RFID and NFC system.

Keywords: BAN Logic; ECC; Key Negotiation; RFID/NFC; Tripartite Authentication

1 Introduction

Radio Frequency Identification (RFID) is an emerging automatic identification technology developed in the 1980s. RFID technology uses a radio frequency signal to send and receive contactless information to authentication, through spatial coupling, that is alternating or electromagnetic field [18]. As the core supporting technology of the Internet of Things (IoT), RFID technology is widely used in logistics, transportation, medicine and industrial manufacturing, *etc.* And a complete RFID system consists of a reader, an electronic tag and a background sever.

Near Field Communication (NFC) technology, as a wireless peer-to-peer communication technology in the IoT, which is evolved from the integration of contactless RFID and interoperability technology, has made a good figure in the electronic payment and smart media. Compared with an RFID system, the slight difference is that an NFC device must be able to be a reader as well as a tag, and the connection between the background and the reader uses a wireless connection, thus we can treat an NFC device as a special RFID system [5, 8, 10].

With the rapid development and widespread application of RFID/NFC technology, the security and privacy issues of RFID/NFC systems have become increasingly prominent. It is currently the most effective method to protect the security and privacy of RFID systems by designing a high security authentication protocol with Public Key Cryptosystem (PKC). Under the premise of the same security in PKC, the elliptic curve cryptosystem (ECC) has become the preferred cryptosystem of RFID authentication protocol, due to its short key length, fast calculation speed and small occupied bandwidth [16, 22–26].

2 Related Work

In studies of RFID authentication protocol based on ECC, most of them focus on the security and efficiency. We briefly review these concerned works from two aspects: The basic security and the efficiency of security defense their protocol provide, and the goal of our proposed protocol.

2.1 Previous Research

As we all know, the RFID security authentication protocol based on ECC has become a hot spot. In 2007, Batina *et al.* [4] discussed the feasibility of an identification protocol based on ECC of the RFID tag, but the confidentiality of the tag's public key is not guaranteed, while the attacker can still obtain its public key, then the tag is tracked. In 2008, Lee *et al.* [12] proposed an ECC-based RFID authentication protocol, while the protocol

isn't resistant to spoofing and tracking attacks. In 2014, Moosavi *et al.* [15] gave a RFID authentication protocol relying on ECC and D-Quark lightweight ash, claiming that its solution is suitable for providing secure and resource-limited RFID implant system, but the required calculation time is not optimized when the tag still needs to calculate elliptic curve multiplication. In 2015, the ECC-based RFID protocol conceived by Ryu *et al.* [17] had a relatively good performance, however it couldn't provide the most basic mutual authentication. In 2016, Kang [9] analyzed and proposed an improved ECC-based RFID Grouping-proof authentication protocol to solve the problem existing grouping-proof protocols such as low grouping-proof efficiency, vulnerability to spoofing attack, tracking attack and other security threats. But it still can't prevent the illegal tag interference with reader authentication and the tag spoofing attack. In 2018, Zhang *et al.* [28] proposed an RFID mutual authentication protocol based on ECC, when thoroughgoing analysis shows that the interactive information C does not contain the information of the random point R_R of the reader, can't resist replay attack on the reader. In 2018, Chen *et al.* [6] proposed a multi-channels constructing method to build protocol model for formal analysis, then used it to verify RFID three-party authentication protocol based on NTRU cryptosystem, the result shows that an attack exists in this protocol. In 2019, Aghili *et al.* [1] shows that the protocol proposed by Fan *et al.* is vulnerable to secret disclosure and reader impersonation attacks. Moreover, they improved it to a protocol that is resistant to the attacks presented in the paper and the other known attacks in the context of RFID authentication.

2.2 Our Target

Through a large amount of literature analysis, we can know that almost all RFID authentication protocols are based on the assumption that the communication between the reader and the background is secure. We can only say that the back-end wired communication is more secure than the front-end over-the-air wireless transmission, but the system still faces the security problems that are common in traditional computer networks, which has a great impact on the security of the authentication protocol. Thus it's unreasonable to assume it's secure [16]. The protocol proposed in our paper negates this assumption. In addition, the difference between the NFC system and the RFID system is analyzed. This paper aims to design a security tripartite mutual authentication protocol based on ECC that is universal to RFID/NFC, which will be more practical.

3 The RFID/NFC Protocol Based on ECC

We introduce our tripartite authentication protocol RFID/NFC based on ECC in this section firstly. Then the two phases' details of our protocol: initialization and authentication, are described as follow.

3.1 Protocol Description

We propose a security RFID/NFC tripartite authentication protocol based on ECC, with good security and anti-attack capabilities. Our paper doesn't support the assumption that the connection of the background and the reader is a wired connection so that the channel between them is safe. Therefore, the protocol proposed in the paper can also be applicable to NFC, which the communication between reader and background is insecure, and has better practicability.

3.2 Initialization Phase

Definition of the relevant symbols in the protocol are explained in Table 1.

Table 1: Summary of symbols in our protocol

Symbol	Symbol's Description
P	Base point on the elliptic curve
R_S, R_P	Private and public key of Reader
T_S, T_P	Private and public key of Tag
r_R, r_T	Random number
AR, AT	Authentication information
VR, VT, VB	Verification information
K_T, K_R, K_B	Co-negotiating secret key

In the initialization phase, the reader randomly selects a number R_S as its private key and calculates its public key $R_P = [R_S]P$ accordingly. While the tag does the same thing, randomly selects a random number T_S as its private key, and calculates its corresponding public key $T_P = [T_S]P$. We specify the public key of the tag as its unique identifier in our protocol. The background server stores the public and private keys of both the reader and the tag. Each tag stores its own public and private key information and the public key of the reader, while the reader only stores its own public and private key information for RFID system, plus the public key of tag for NFC system.

3.3 Authentication Phase

As shown in Figure 1, the specific mutual authentication process among the tag, the reader, and the background is as follows:

Step 1: The reader chooses a random integer number r_R that belongs to Z_q, and calculates the point $R_R = [r_R]P$ on the elliptic curve, then sends an query and R_R to the tag.

Step 2: The tag also selects a random integer r_T in Z_q, calculating the corresponding point $R_T = [r_T]P$ on

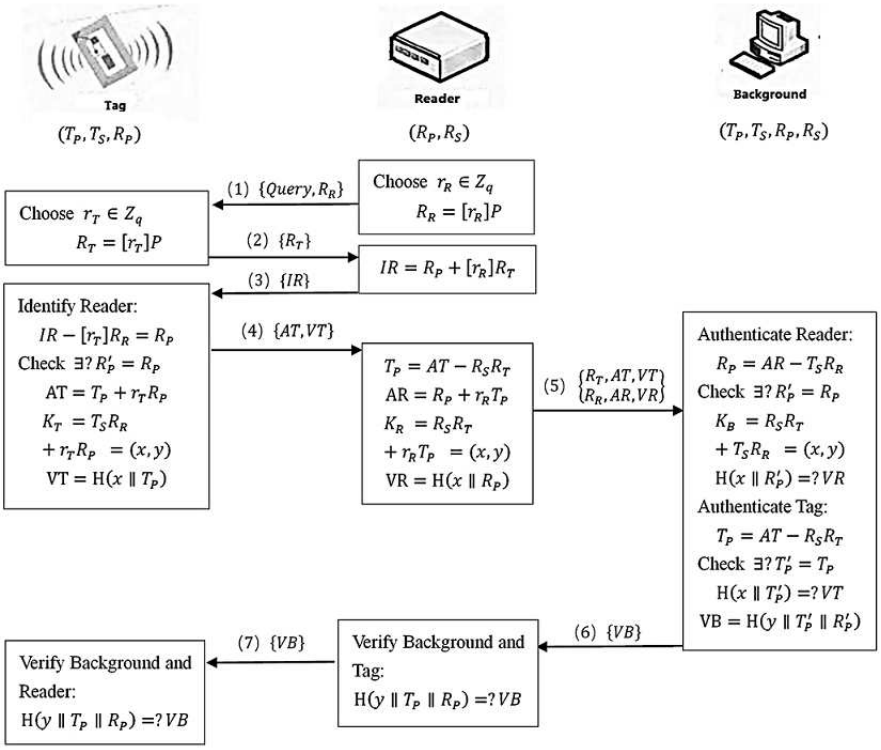


Figure 1: ECC-based RFID/NFC authentication protocol

the elliptic curve. Then it sends R_T to respond the reader.

Step 3: The reader calculates the $IR = R_P + [r_R]R_T$ and sends it, so that the tag identify the reader.

Step 4: The tag identify the reader by counting $R_P = IR - [r_R]R_T$, and it searches whether there is a R'_P equal to R_P. If established, the tag thinks the reader is legal and calculates the Authentication variable $AT = T_P + r_T R_P$ of the tag. What's more, it is possible to calculate the co-negotiating secret key $K_T = T_S R_R + r_T R_P = (x, y)$, and finally calculates the tag verification amount $VT = H(x \| T_P)$for background verification. Then sends AT, VT to the reader.

Step 5: The reader calculates $T_P = AT - R_S R_T$ from the received AT, and counts its authentication variable $AR = R_P + r_R T_P$, co-negotiating secret key $K_R = R_S R_T + r_R T_P = (x, y)$, and the verification variable $VR = H(x \| R_P)$ for the background authenticate the reader. Then the associated amount of the tag and the reader R_T, AT, VT, R_R, AR, VR are passed to the background server.

Step 6: After receiving the message transmitted by the reader, the background server first verifies the legitimacy of the reader by calculating the $R_P = AR - T_S R_R$, and then retrieves whether it's the same as the stored R'_P or not. If it exists, the background calculates the co-negotiating secret key $K_B = R_S R_T + T_S R_R = (x, y)$, and $H(x \| R'_P)$ to check whether it is equal to VR or not. After the reader passing the verification, the background verifies whether the information of tag collected by the reader is legal, calculating the $T_P = AT - R_S R_T$, and retrieves whether there is corresponding T'_P in the repository. Based on it, the background calculates $H(x \| T'_P)$ and judges whether it is equal to VT. Finally, the background calculates its authentication variable $VB = H(y \| T'_P \| R'_P)$ for the reader and the tag to authenticate the legitimacy of the background server, and sends it to the reader.

Step 7: The reader verifies the background server and the tag, when calculated $H(y \| T_P \| R_P)$ and judged whether it is equal to the received VB. And it sends VB to the tag.

Step 8: Lastly, based on the information that the tag receives, the tag verifies the legitimacy of the reader and the background by calculating whether $H(y \| T_P \| R_P)$ is equal to VB. The mutual authentication among the tag, the reader and the background is accomplished, and K_T, K_R, K_B are the same, as a session key is used for subsequent communication.

4 Security Analysis

A safe RFID/NFC system should be able to provide mutual authentication, confidentiality, anonymity, forward security, scalability. As well as resist tracking attack, denial of service attack, spoofing attack, and replay

attack, *etc.* [11,20,27,30] Besides, considering the practicality of RFID/NFC, it should also be able to provide a co-negotiating secret key for subsequent communication. The security performance and scalability of the protocol in our paper has been greatly improved, satisfying the basic security requirements as mentioned above.

4.1 Qualitative Analysis

We give qualitative analysis of our protocol from nine aspects:

- **Mutual authentication:** The tag verifies the reader by judging $R_P^{'} = R_P$ to preliminarily certificate; then it uses K_T to authenticate the background and the reader further with VB.

  To the reader, it also authenticates the background with VB, K_R. Furthermore, VB from the background contains the legal public key T_P of the tag, and the reader can confirm the validity of the tag through the background ulteriorly.

  When received the information, the background retrieves whether there is $R_P^{'} = R_P$, to initially determine that the reader is legal; then uses K_B, VB to authenticate the reader. In the same way, the certification of the tag in the background can be obtained. As a result, the protocol completes the mutual authentication of the three parties.

- **Confidentiality:** In the process of authentication, the public key T_P of the tag is used as its unique identifier, which is calculated by $T_P = AT - R_S R_T$. The public key of the reader is also calculated by $R_P = IR - [r_R]R_T$. Both of them don't transmitted on the channel. Even if the attacker intercepts the interactive information $R_R, R_T, AR, AT, VR, VT, VB$ on the wireless channel, due to the discrete logarithm problem of elliptic curve and the randomness and unipolarity of the Hash function which are unable to be solved based on today's computer calculation, either T_P or R_P is not derived. This ensures the confidentiality of the tag identity and the reader's public key.

- **Anonymity:** As we know, neither the public key of the tag T_P nor the reader R_P is transmitted over the channel directly. Due to the security of the elliptic curve cryptosystem, the attacker can't calculate the corresponding private key and the identity of the parties from the interaction information. So the protocol can provide anonymity of the tag, the reader and the background.

- **Forward security:** Assuming that the attacker can attack on the maximum degree, which gets the public key T_P, R_P and all the interaction information $\{R_R, R_T, AR, AT, VR, VT, VB\}$, the attacker still can't calculate the random number r_R, r_T through these, let alone T_S, R_S. Thus it is impossible to bind the obtained interactive information with the specific tag or the reader, and the protocol has good forward security.

- **Scalability:** For RFID system, the reader doesn't need to store the unique identifier T_P of the tag by calculating it, therefore, a large memory reduction can be achieved for a large number of tags. Similarly, if the memory requires high memory, the tag do the same thing to calculate R_P instead of storing it [19]. But it does in NFC system. In addition, since the public key are used as the unique identifier, the identity validity period can be added to the identity identifier, when the identity is invalid, it can no longer participate in encryption and decryption and authentication, which makes the protocol more practical.

- **Resist tracking attack:** According to the confidentiality, the attacker can't get T_P, R_P. The reader and the tag will generate new random numbers in each new session, hence the interaction information is also fresh at each time. Unpredictable changes in the session make it impossible for the attacker to track the tag or the reader.

- **Resist denial of service attack (DoS attack):** The guarantee of anonymity enables T_P and R_P to be effectively protected, and the private keys of them don't need to be updated, as a result, the shared secret information doesn't need to be updated synchronously among the tag, the reader and the background. In consequence, the protocol can resist denial of service attack.

- **Resist spoofing attack:** If the attacker wants to impersonate a legitimate tag to deceive the reader, it needs to forge a legitimate authentication message R_T, AT, VT, since there is no legal tag identity T_P and T_S, and K_T. The attacker can't generate valid authentication messages AT, VT, so that it can't deceive the background. In case that the attacker wants to impersonate a legitimate reader to spoof the tag, it is necessary to forge a legitimate authentication message AR, VR, but it can't calculate T_P, and K_R. At this point, the attacker is even more impossible to spoof the background.

  Assume that the attacker wants to impersonate the background spoofing the tag and the reader, it's necessary to forge a legitimate authentication message VB. Because there is no legal T_P, R_P, and K_B, the attacker can't generate a valid VB. The protocol is resistant to spoofing attack.

- **Resist replay attack:** Suppose the attacker replays the tag by intercepting the interactive information R_R and AR, VR. While the tag generates a new random number r_T in each session, and it can pass the verification $H(y \| T_P \| R_P) = ?VB$ to determine

whether it is attacked. It's Similar to the reader. For this reason, the protocol resists replay attack.

4.2 Formal Analysis

The formal analysis method is a standardized method, judging whether the authentication protocol itself meets the security objectives, and whether there are security vulnerabilities. It is divided into a structural method based on reasoning, an attack-based structural method and a theorem-based proof method. Burrows-Abadi-Needham (BAN) logic [3, 21] is an industry-recognized milestone in the formal analysis for security authentication protocols. BAN-based logic is widely used in the field of authentication protocol analysis. In this subsection, we adopt the widely-accepted BAN logic to demonstrate that the proposed authentication protocol guarantees mutual authentication and secure session key establishment between the communicating parties.

4.2.1 Basic Terms of BAN Logic

We explain the important notations of BAN logic in Table 2.

Table 2: Notations of BAN logic

Notation	Notation's Description
$P \models X$	P trusts the statement X
$P \lhd X$	P sees X
$P \mid\sim X$	P once said X
$P \Rightarrow X$	P can rule X
$P \overset{SK}{\leftrightarrow} Q$	P and Q share the secret key SK to communicate between each other
$\overset{K}{\longrightarrow} P$	K is the public key of P
$P \overset{X}{\rightleftharpoons} Q$	X is secret information between P and Q
$\#(X)$	X is fresh
$\{X\}_K$	Ciphertext obtained by encrypting X with key K
(X, Y)	X or Y is one part of (X,Y)

Next, the inference rules of BAN logic are shown.

Rule 1: Message-meaning Rule:

$$\frac{P \models P \overset{K}{\longleftrightarrow} Q, P \lhd \{X\}_K}{P \models Q \mid\sim X}$$

Rule 2: Jurisdiction Rule:

$$\frac{P \models Q \Rightarrow X, P \models Q \models X}{P \models X}$$

Rule 3: Nonce-verification Rule:

$$\frac{P \models \#(X), P \models Q \mid\sim X}{P \models Q \models X}$$

Rule 4: Belief Rule:

$$\frac{P \models X, P \models Y}{P \models (X, Y)}$$

Rule 5: Freshness Rule:

$$\frac{P \models \#(X)}{P \models \#(X, Y)}$$

Rule 6: Message-sink Rule:

$$\frac{P \models P \overset{K}{\longleftrightarrow} Q, P \lhd \{X\}_K}{P \lhd X} \qquad \frac{P \lhd (X, Y)}{P \lhd X}$$

Rule 7: Hash Rule:

$$\frac{P \models Q \mid\sim H(X), P \lhd X}{P \models Q \mid\sim X}$$

4.2.2 BAN Logic Analysis of Protocol

There are three participants in our protocol, include T(tag), (R)reader, and B(background).In the protocol for RFID system, T stores the public key of R; for NFC system, R, which is also as tag now, stores the public key of T, which is as the reader at the same time. The background holds the public and private keys of both T and R, while there is no assumption that communication between R and B is safe. We use BAN logic to formally analyze the protocol in following part, which is mainly divided into message idealization, initialization hypothesis, security goal and certification process.

- **Message Idealization:**

$$T \lhd \{R_R, \{R_P\}_{R_P}, H(y \| T_P' \| R_P')\} \quad (1)$$

$$R \lhd \{T_R, \{T_P, r_T\}_{T_P}, H(x \| T_P), H(y \| T_P' \| R_P')\} \quad (2)$$

$$B \lhd \{R_T, \{T_P, r_T\}_{T_P}, H(x \| T_P), R_R, \{R_P, r_R\}_{R_P}, H(x \| R_P)\}$$

- **Initialization Hypothesis:**
  H_1: **Validity of the keys**

$$T \models T \overset{R_P}{\longleftrightarrow} R \quad (3)$$

$$R \models R \overset{T_P}{\longleftrightarrow} T \quad (4)$$

$$T \models T \overset{T_P'}{\longleftrightarrow} B$$

$$B \models B \overset{T_P}{\longleftrightarrow} T$$

$$R \models R \overset{R_P'}{\longleftrightarrow} B$$

$$B \models B \overset{R_P}{\longleftrightarrow} R$$

H_2: **Authority of the subjects**

$$T \models R|\Rightarrow R_P \quad (5)$$
$$R \models T|\Rightarrow T_P \quad (6)$$
$$T \models B|\Rightarrow T_P^{'}$$
$$B \models T|\Rightarrow T_P$$
$$R \models B|\Rightarrow R_P^{'}$$
$$B \models R|\Rightarrow R_P$$

H_3: **Freshness of random numbers**

$$T \models \#(R_R) \quad (7)$$
$$R \models \#(R_T) \quad (8)$$

- **Security Goal:**
There are two main goal to achieve, where we first to authenticate the tag, reader and background to each other, then show that they agree on a session key.

G_1: **Primary goal**

$$T \models T \overset{K_T}{\longleftrightarrow} R$$
$$R \models R \overset{K_R}{\longleftrightarrow} T$$
$$R \models R \overset{K_R}{\longleftrightarrow} B$$
$$B \models B \overset{K_B}{\longleftrightarrow} R$$

where $K_T = K_R = K_B$.

G_2: **Secondary goal**

$$T \models R|\equiv T \overset{K_T}{\longleftrightarrow} R$$
$$R \models T|\equiv R \overset{K_R}{\longleftrightarrow} T$$
$$R \models B|\equiv R \overset{K_R}{\longleftrightarrow} B$$
$$B \models R|\equiv B \overset{K_B}{\longleftrightarrow} R$$

- **Certification Process:**
Proof of G_1:
From Equations (1), (3), and Rule 1: $\frac{P|\equiv P \overset{K}{\longleftrightarrow} Q, P \lhd \{X\}_K}{P|\equiv Q|\sim X}$. We have

$$T|\equiv R|\sim (R_R, \{R_P\}_{R_P}, H(y\|T_P^{'}\|R_P^{'})) \quad (9)$$

From Equation (7) and Rule 5: $\frac{P|\equiv \#(X)}{P|\equiv \#(X,Y)}$. Derive

$$T|\equiv \#(R_R, \{R_P\}_{R_P}, H(y\|T_P^{'}\|R_P^{'})) \quad (10)$$

From Equations (9), (10), and Rule 3: $\frac{P|\equiv \#(X), P|\equiv Q|\sim X}{P|\equiv Q|\equiv X}$. With

$$T|\equiv R|\equiv (R_R, \{R_P\}_{R_P}, H(y\|T_P^{'}\|R_P^{'})) \quad (11)$$

From Equation (11) and Rules 3, 5. We can get

$$T|\equiv R|\equiv (R_P, H(y\|T_P^{'}\|R_P^{'}))$$
$$T|\equiv R|\equiv \#(R_P, H(y\|T_P^{'}\|R_P^{'}))$$

From Equations (5) and (6), there is

$$T|\equiv T \overset{K_T}{\longleftrightarrow} R,\ T|\equiv \#(K_T) \quad (12)$$

From Equations (2), (4), and Rule 1, have

$$R|\equiv T|\sim \{R_T, \{T_P, r_T\}_{T_P}, H(x\|T_P), H(y\|T_P^{'}\|R_P^{'})\} \quad (13)$$

From Equations (13) and (8), get

$$R|\equiv T|\equiv R \overset{K_R}{\longleftrightarrow} T \quad (14)$$

From Equations (14) and (6), we know

$$R|\equiv R \overset{K_R}{\longleftrightarrow} T$$

One part of G_1 is certified, proof of the rest can be obtained by analogy.

Proof of G_2
From Equation (1) and Rule 6: $\frac{P \lhd (X,Y)}{P \lhd X}$. we have

$$T \lhd H(y\|T_P^{'}\|R_P^{'}) \quad (15)$$

From Equations (12), (15), and Rule 1, with

$$T|\equiv R|\sim K_T \quad (16)$$

From Equations (12) and (16) again, we final get

$$T \models R|\equiv T \overset{K_T}{\longleftrightarrow} R$$
$$R \models T|\equiv R \overset{K_R}{\longleftrightarrow} T$$

The similar to the other part of G_2.

4.2.3 Conclusion of Formal Analysis

In summary, the BAN logic formal analysis method proves that the proposed protocol can achieve the expected goal, and also shows that the protocol is safe and reliable in theory.

5 Performance Analysis

In this section, we analyze the performance of the improved protocol from two aspects: The practical advantage and the security comparison with other protocol.

5.1 Practical Advantage

The existing ECC-based security authentication protocol of RFID face the threat brought by the traditional computer network even the traditional computer network communication. So the assumption that the communication between reader and background is secure, which is obviously unreasonable. The protocol in our paper abandons this hypothesis and makes it more scientific.

Beyond that the protocol negotiates the secret key $K_T = K_R = K_B$ in the authentication process, which facilitates subsequent communication between each other. Based on this, the tag can support the reading and writing function, so that information of the target can be updated at any time in real life. This is also not available in many current RFID protocols. All of these improvements makes the tag can be used as a reader as well, so our protocol can be applied to NFC systems, which is more practical.

5.2 Security Comparison

We have shown the security analysis of our protocol above, now we compare our protocol to the latest related protocols in terms of security [2].

Table 2 shows the security comparison of our protocol to Zhang *et al.*'s protocol [29], Liao *et al.*'s protocol [13], Liu *et al.*'s protocol [14] from the necessary security of RFID system, where "$\surd$" means satisfy, "$\times$" means not satisfy.

Table 3: Security properties comparison

Requirements	[29]	[13]	[14]	Our
Mutual authentication	$\times$	$\surd$	$\surd$	$\surd$
Confidentiality	$\surd$	$\surd$	$\times$	$\surd$
Anonymity	$\surd$	$\surd$	$\surd$	$\surd$
Forward security	$\surd$	$\surd$	$\surd$	$\surd$
Scalability	$\times$	$\times$	$\times$	$\surd$
Tracking attack	$\surd$	$\surd$	$\surd$	$\surd$
DoS attack	$\times$	$\surd$	$\times$	$\surd$
Spoofing attack	$\surd$	$\surd$	$\surd$	$\surd$
Replay attack	$\surd$	$\surd$	$\surd$	$\surd$
Mobile environment	$\times$	$\times$	$\times$	$\surd$

As illustrated in Table 3, Zhang *et al.*'s protocol only satisfies one-way authentication from the reader to the tag. Our protocol not only satisfies the two-way authentication of the tag and the reader, but also contents the mutual authentication among the tag, the reader and the background. Liu *et al.*'s protocol doesn't meet the basic confidentiality, and our protocol can solve this problem well. Both Zhang *et al.*'s protocol and Liu *et al.*'s protocol are not resistant to denial of service attack. Instead, our protocol is resistant to multiple attacks including denial of service attack. In addition, all of them have no scalability, which is necessary to the large-scale application of RFID in the IoT, while the tag that our protocol can satisfy and has good scalability. Relatively speaking, our protocol is also applicable to NFC, so only it can be applied to smart device environments such as mobile phones.

6 Conclution

A security tripartite authentication protocol based on ECC of RFID/NFC system is designed in our paper. Since we assume that the communication between reader and background is insecure, tag, reader and background can achieve mutual authentication, which is more scientific and reasonable. Apart from this, they negotiate a secret co-negotiating key for subsequent communication. We through qualitative analysis the basic security of the protocol, and the result show that our protocol can provide mutual authentication, confidentiality, anonymity,etc. As well as resist tracking attack, denial of service attack, spoofing attack, *etc.* Then we further formal analyze the security and aim of our protocol by the BAN logic, while result of the analysis indicated that our protocol achieves the goals that tripartite authentication and key agreement. Compare our protocol to the latest related protocols in security, we can know that our protocol has greater security and better availability. Besides, the protocol, where uses the public key as the identity of tag or reader can provide and add more information to it. To sum up, it can not only solve the problem effectively, which current and potential security issues faced by current RFID systems, but also be applied to NFC systems.

References

[1] S. F. Aghili and H. Mala, "Security analysis of an ultra-lightweight RFID authentication protocol for m-commerce," *International Journal of Communication Systems*, vol. 32, no. 3, pp. e3837, 2019.

[2] P. Alexander, R. Baashirah, and A. Abuzneid, "Comparison and feasibility of various RFID authentication methods using ECC," *Sensors*, vol. 18, no. 9, pp. 2902, 2018.

[3] A. Arfaoui, A. Kribeche, and S. M. Senouci, "Context-aware anonymous authentication protocols in the internet of things dedicated to e-health applications," *Computer Networks*, vol. 159, pp. 23–36, 2019.

[4] L. Batina, J. Guajardo, T. Kerins, N. Mentens, P. Tuyls, and I. Verbauwhede, "Public-key cryptography for RFID-tags," in *Fifth Annual IEEE International Conference on Pervasive Computing and Communications Workshops (PerComW'07)*, pp. 217–222, 2007.

[5] S. Bojjagani and V. N. Sastry, "A secure end-to-end proximity NFC-based mobile payment protocol," *Computer Standards & Interfaces*, pp. 103348, 2019.

[6] J. Chen, M. Xiao, K. Yang, W. Li, and X. Zhong, "Formal analysis and verification for three-party authentication protocol of RFID," in *National Conference of Theoretical Computer Science*, pp. 46–60, 2018.

[7] Y. L. Chi, C. H. Chen, I. C. Lin, M. S. Hwang, "The secure transaction protocol in NFC card emulation mode," *International Journal of Network Security*, vol. 17, no. 4, pp. 431–438, 2015.

[8] 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.

[9] K. Hong-yan, "Analysis and improvement of ECC-based grouping-proof protocol for RFID," *International Journal of Control and Automation*, vol. 9, no. 7, pp. 343–352, 2016.

[10] W. Huo, Q. Dong, and Y. Chen, "ECC-based RFID/NFC mutual authentication protocol," in *The 2nd International Workshop on Materials Engineering and Computer Sciences*, 2015. DOI: 10.2991/iwmecs-15.2015.31.

[11] M. Khalid, U. Mujahid, and N. ul I. Muhammad, "Ultralightweight RFID authentication protocols for low-cost passive RFID tags," *Security and Communication Networks*, vol. 2019, pp. 25, 2019.

[12] Y. K. Lee, L. Batina, and I. Verbauwhede, "EC-RAC (ECDLP based randomized access control): Provably secure RFID authentication protocol," in *IEEE International Conference on RFID*, pp. 97–104, 2008.

[13] Y. P. Liao and C. M. Hsiao, "A secure ECC-based RFID authentication scheme integrated with ID-verifier transfer protocol," *Ad Hoc Networks*, vol. 18, pp. 133–146, 2014.

[14] G. Liu, H. Zhang, F. Kong, and L. Zhang, "A novel authentication management RFID protocol based on elliptic curve cryptography," *Wireless Personal Communications*, vol. 101, no. 3, pp. 1445–1455, 2018.

[15] S. R. Moosavi, E. Nigussie, S. Virtanen, and J. Isoaho, "An elliptic curve-based mutual authentication scheme for RFID implant systems," *Procedia Computer Science*, vol. 32, pp. 198–206, 2014.

[16] Y. Pan, Z. Shan, Q. Dai, and F. Yue, "CPK-ECC based mutual authentication protocol for large-scale RFID system," *Journal on Communications*, vol. 38, no. 8, pp. 165–171, 2017.

[17] E. K. Ryu, D. S. Kim, and K. Y. Yoo, "On elliptic curve based untraceable RFID authentication protocols," in *Proceedings of the 3rd ACM Workshop on Information Hiding and Multimedia Security*, pp. 147–153, 2015.

[18] H. Shen, J. Shen, M. K. Khan, and J. H. Lee, "Efficient RFID authentication using elliptic curve cryptography for the internet of things," *Wireless Personal Communications*, vol. 96, no. 4, pp. 5253–5266, 2017.

[19] X. Tan, M. Dong, C. Wu, K. Ota, J. Wang, and D. W. Engels, "An energy-efficient ECC processor of UHF RFID tag for banknote anti-counterfeiting," *IEEE Access*, vol. 5, pp. 3044–3054, 2016.

[20] A. Tewari and B. B. Gupta, "Cryptanalysis of a novel ultra-lightweight mutual authentication protocol for IoT devices using RFID tags," *The Journal of Supercomputing*, vol. 73, no. 3, pp. 1085–1102, 2017.

[21] L. Tingyuan, L. Xiaodong, Q. Zhiguang, and Z. Xuanfang, "An improved security protocol formal analysis with ban logic," in *International Conference on Electronic Commerce and Business Intelligence*, pp. 102–105, 2009.

[22] C. H. Wei, M. S. Hwang, A. Y. H. Chin, "A mutual authentication protocol for RFID", *IEEE IT Professional*, vol. 13, no. 2, pp. 20–24, Mar. 2011.

[23] 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.

[24] 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.

[25] 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.

[26] 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.

[27] G. Xu, Y. Ren, Y. Han, X. Li, and Z. Feng, "Privacy protection method based on two-factor authentication protocol in frid systems," *IEICE Transactions on Information and Systems*, vol. 99, no. 8, pp. 2019–2026, 2016.

[28] X. Zhang and Y. Guo, "Research on RFID system security authentication protocol based on elliptic curve cryptography," *Net Information Security*, vol. 18, no. 10, pp. 51–61, 2018.

[29] X. Zhang, L. Li, Y. Wu, and Q. Zhang, "An ECDLP-based randomized key RFID authentication protocol," in *International Conference on Network Computing and Information Security*, vol. 2, pp. 146–149, 2011.

[30] L. Zheng, Y. Xue, L. Zhang, and R. Zhang, "Mutual authentication protocol for RFID based on ECC," in *IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC)*, vol. 2, pp. 320–323, 2017.

Biography

Yong-Shuang Wei, is currently a Master student in the Department of Mathematics and Statistics at Wuhan University, China. She holds a Bachelor of Science degree in Applied Mathematics at Chongqing University, China. Her main research interests are cryptography and information security, especially elliptic curve cryptography.

Jian-Hua Chen, is a Professor in the Department of Applied Mathematics at Wuhan University, Hubei, China. He is the Director of the Information Security Research Center, and the main designer of the SM2 Signature Algorithm for ISO/IEC International Standards. His research interests include number theory and information security, elliptic curve cryptography.

Secure Sharing of Data for Dynamic Group in Public Cloud

Cungang Yang and Celia Li
(Corresponding author: Cungang Yang)
Department of Electrical and Computer Engineering, Ryerson University
350 Victoria Street, Toronto, Canada
(Email: cungang@gmail.com)
(Received June 2, 2019; Revised and Accepted Dec. 28, 2019; First Online Feb. 26, 2020)

Abstract

Cloud Computing has been envisioned as the next-wave system architectures of IT. Cloud storage service brings benefit of not only low cost and scalability, but also great appropriateness for group sharing, such as e-learning resource storage. In this paper, a scalable and efficient group sharing method for public cloud is proposed. It is based on the key tree approach and the well-known Diffie-Hellman cryptographic protocol. Security and performance analysis shows that our proposed scheme is secure and highly efficient for public cloud based secure storage.

Keywords: Cloud Storage; Group Key Management; Group Sharing; Proxy Re-Encryption

1 Introduction

With the development of cloud services and social networks, a group can be easily organized between some people over Internet due to the same interests, so that group applications with the aid of cloud servers become possible and attract more and more attentions [4, 8, 11, 18].

Cloud storage is an extensive new service in the field of cloud computing, which is a distributed file system that using software integrate various types storage to supply users with data storage and access service. Users can easily share their data and reduce the overhead of building and maintaining their data center.

Despite of the advantages of cloud storage, there are various challenges on the privacy and security of users' data. For example,

1) The cloud is usually maintained and managed by a semi-trusted third party (cloud provider).

2) While it is desirable for the data owner to share his/her private data with intended recipients, it presents an even more challenging problem since we have to make sure that except the intended recipients, nobody, including the cloud providers, can obtain any useful information from the encrypted data.

In order to ensure that only authorized users can access the shared data, the data are encrypted using a cryptographic key known as the *group key*. Data are encrypted with the group key and stored in the cloud. The group key should only be known by authenticated and authorized members of a group. Authorized users can download the encrypted file and decrypt them with the group key. But how to distribute and update group keys is an important problems. We can use digital envelope to address this problem. A digital envelope is a secure electronic data container that is used to protect a message through encryption and data authentication. A digital envelope allows users to encrypt data with the speed of secret key encryption and the convenience and security of public key encryption. For example, when Bob wants to send a confidential message M to Alice, she can generate a digital envelop for M and send the envelop to Alice. On the sender's side the procedure is as follows:

1) Bob randomly generates a secret key K.

2) Bob encrypts M with K, $E(K, M)$.

3) Bob encrypts K with Alice's public key $E(Pub_A, K)$.

4) Bob concatenates $E(K, M)$ with $E(Pub_A, K)$ and sends the result to Alice as digital envelop.

Upon receipt of $E(K, M)$ and $E(Pub_A, K)$, Alice uses her private key PR_A to decrypt the message. The procedure is as follows:

1) Alice decrypts $E(Pub_A, K), K)$ with PR_A.

2) Alice decrypts $E(K, M)$ with K. The result is M.

Before uploading a file to cloud servers, the data owner symmetrically encrypts the file with a randomly chosen session key. The data owner also uploads a digital envelope. Group members should timely get the updated key from cloud servers to get the group private key. When a

group member requests to download a file, he/she sends a request to cloud servers. Cloud servers respond with a random number. The group member uses the current agreed group private key to sign the number. Cloud servers use the agreed group public key to verify the signature. If passed, they send the encrypted file and the specific digital envelope to the member. Mi open the envelop and get the requested file [15].

The architecture of secure shared cloud storage is shown in Figure 1 [15] where the group leader opens up a sharing area in the cloud to form a group application. Then, he/she grants the group members the right to implement data management. All the data in this group are available to all the group members, while they remain private towards the outsiders of the group including the cloud provider. To prevent the cloud provider to access the original date, proxy re-encryption is used to provide ciphertext updating in cloud environment. By this way, most computational intensive operations of ciphertext updating can be transferred to cloud servers, without reveal any content of ciphertext to the cloud provider.

Proxy re-encryption is a technique which allows a semi-trusted server transform the ciphertexts encrypted by one party's public keys into proxy ciphertexts encrypted by another party's public keys [2,3,5,10]. Under this circumstance, the two parties can conduct secure data sharing over the same plaintext without exchanging their private keys or worrying about sensitive data leakage in the semi-trusted server. Proxy re-encryption is an cryptographic primitive in which one person (Take the user A for example) allows a semi-trusted cloud provider to reencrypt his/her message that will be sent to another designated person (Take the user B for example). A should generate a proxy re-encryption key by combining his/her secret key with B's public key. This re-encryption key is used by the proxy as input of the re-encryption function, which is executed to convert a ciphertext encrypted under A's public key into another ciphertext that can be decrypted by B's private key. Except for converting, the proxy (cloud provider) cannot see the underlying data contents.

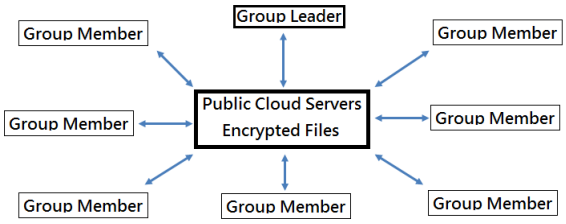


Figure 1: Network model

There are two kinds of users:

1) Group Leader: Only one group leader in a group, who is the group creator. The group leader buys or obtains storage and computing resource from the cloud provider.
2) Group member: Each group member can implement file download and upload operations in the group.

Each group member can get some related public information from cloud servers and compute the group key pair. The group membership can change overtime: each group member except the group leader can leave or apply to join the group at his/her will. All group members can negotiate a group key pair (the group public key and the group private key) with the help of cloud servers. This group key pair is used to protect the data shared in the group. Group members' leaving and joining can launch key updating process. Every time a membership change occurs, the group key must be changed to ensure backward and forward secrecy. Backwards secrecy guarantees that a new member joining the group does not have access to any old group keys. This ensures that a member cannot decrypt messages sent before it joins the group. Forward secrecy requires that a member leaving the group does not have access to any future group keys. This ensures that a member cannot decrypt future messages after it leaves the group.

Some researchers have emphasized the integrity and availability of outsourced data. Wang proposed a homomorpic distributed verification protocol using pseudorandom data to ensure cloud storage security [13]. The protocol focuses on the storage correctness as well as verifies misbehaving servers. However, Pseudorandom data does not cover the entire data while identifies the cloud servers, some data corruptions maybe missing such that the protocol do not provide full protection for cloud storage. Kamara proposed a framework of a cryptographic storage service which considers the issue of building a secure cloud storage service on cloud infrastructure where the service provider is not fully trusted by the user [6]. It is made up of three basic components and realizes encryption storage and integrity validation by a group of protocols. However, this method is hard to build since it considers at a high level, needs to modify lots of the source code of cloud storage platform. In addition, users have to query data owner to access the shared data, which will make a communication bottleneck as the number of users increases rapidly. Yeh proposed a secure group communication and data sharing scheme using public key cryptography [17]. However, using such asymmetric cryptography in group needs a PKI infrastructure and the trusted Certificate Authority in the system and each entity need to query the public key of other entity, which will be a overhead as large amount of group members.

Ateniese [1] proposed a proxy re-encryption scheme to manage distributed file systems that attempt to achieve secure data storage in the semi-trusted party. Based on bilinear maps, the scheme offers improved security guarantees. An example of group data sharing in cloud computing was proposed by Liu [9]. In [9], a secure scheme was proposed to support anonymous data sharing in cloud computing. In this paper, we proposed a scheme named CTDH, Our scheme supports the updating of the group

key pair whenever group members joining or leaving. Our approach transfers most of the computational complexity and communication overhead of group key updates to cloud servers without leaking their privacy.

The remainder of this paper is organized as follows. In Section 2, we proposed the group sharing method. The security proof of the scheme is also provided. We analyze the performance of the proposed scheme in Section 3. Section 4 summarizes the paper.

2 Our Proposed Group Sharing Method

The group is composed of the entity with similar interest or common purpose such as society, work-group, or e-learning team. Initially, the group manager as group leader takes charge of key management activities including key generation and key updating. We assume that the group key creation and updates are proceeded in secure channel.

Our scheme uses the key tree structure for scalability and minimal key computation and distribution. It employs the Diffie-Hellman (DH) cryptographic algorithm to avoid the need of setting up a shared key between a member and its group leader. Although each member contributes to the group key, the group leader is responsible for generating and distributing partial keys.

Our scheme is similar to TGDH [7, 12] with respect to the use of the key tree structure and the DH algorithm. They differ in several aspects, as follows:

- In our scheme, the group leader is responsible for computing and distributing the partial keys to the group when a new member joins or an existing member leaves. In TGDH, when a new member joins the group, the current right-most member in the logical key tree is responsible for computing and distributing the partial keys to the other members of the group.
- In TGDH, each member in the group needs to store the entire key tree. In our scheme, only the group leader needs to store the key tree. Each member needs to store only the partial keys of the nodes on the path from itself to the root of the key tree.
- Because the group key updates are now done by the group leader in our scheme instead by the right-most member, each intermediate node in the logical key tree requires two secret keys. These secret keys are changed whenever the partial key associated with the intermediate node needs to be updated upon a join or leave event. The use of these secret keys at intermediate nodes is described in Section 2.1.2.
- It has been shown that the key tree approach gives best performance with 4-ary trees [14, 16]. However, TGDH must work only on binary trees in order to maintain the optimal number of rekeying computations and messages. Our scheme, on the other hand, can support trees of any degree without affecting the complexity of key computation and communication.

In the following sub-sections, we discuss how the group key is computed and updated upon join and leave operations.

2.1 Group Key Generation

2.1.1 Key Tree

The key tree is a logical data structure.

The nodes are numbered as follows. $K_{i,j}$ denotes the content of node $< i, j >$.

Each leaf node $< i, j >$ is associated with a group member and contains the member's secret key (which is generated by the member himself) denoted by $s_{i,j}$.

$$K_{i,j} = s_{i,j}. \tag{1}$$

In the logical key tree displayed by Figure 2, the leaf nodes contain the secret keys $s_{3,0}, ...s_{3,3}, s_{2,2}$ and $s_{2,3}$ of the six group members.

The content of each non-leaf node $< i, j >$ is called a proper key and computed from the contents of its two children $< i+1, 2j >$ and $< i+1, 2j+1 >$ in a recursive manner:

$$K_{i,j} = f(K_{i+1,2j}, K_{i+1,2j+1}). \tag{2}$$

For example, $K_{1,1} = f(s_{2,2}, s_{2,3})$.

By this recursive definition, the proper key $K_{i,j}$ of a non-leaf node $< i, j >$ is made up of the secret keys of the group members in the subtree rooted at node $< i, j >$. For instance, $K_{1,0}$ is the result of $s_{3,0}$, ..., $s_{3,3}$.

The content of the root node $< 0, 0 >$ is the *group key* used by the data source to encode a message and by the group members to decode the message. The group key $K_{0,0}$ is made up of the secret keys of all the group members (i.e., the leaf nodes).

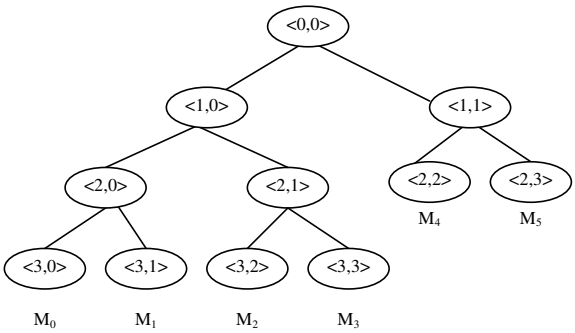


Figure 2: A logical key tree

In the logic key tree, we denote the group key $K(0, 0)$ and $g_K(0, 0)$ as the group private key PrKG and group public key PuKg. The established group key pair is shared by all group members: PrKG = $K_{0,0}$ and PuKG = $g_K(0, 0)$.

2.1.2 Group Key Generation

Each member contributes a share to the group key, but the group leader computed required partial keys and distribute to the whole group. Each member extracts the necessary partial keys from the leader's broadcast message, then combines with his own secret key to compute the group key as follows.

The required variable for each leaf node is following:

- Each member $< i, j >$ generates its own secret key $s_{i,j}$ to contribute towards the group key. It sends $g^{s_{i,j}}$ to the leader only.

Each non-leaf node $< i, j >$ requires the following variables:

- a proper key $K_{i,j}$ computed by the leader. The computation of $K_{i,j}$ is different from that in TGDH though, as will be shown shortly.
- a secret key $S_{i,j}$ generated by the leader for node $< i, j >$. The leader generates a new key $S_{i,j}$ when $K_{i,j}$ needs to be updated (i.e., node $< i, j >$ is on the path from the root to the joining/leaving member in the key tree).
- a hidden key $R_{i,j}$, also generated by the leader for node $< i, j >$. The hidden key $R_{i,j}$ is updated along with the secret key $S_{i,j}$.

Each member $< i, j >$ also has a proper key $K_{i,j}$, where $K_{i,j} = s_{i,j}$, the secret key created by the member himself.

If node $< i, j >$ is a non-leaf node, its proper key $K_{i,j}$ is computed from $S_{i,j}$, $R_{i,j}$ and the two proper keys of the two children of nodes $< i, j >$ as follows:

$$\begin{aligned} K_{i,j} &= (g^{R_{i,j}} \times g^{K_{i+1,2j}})^{S_{i,j}} \times (g^{S_{i,j}})^{K_{i+1,2j+1}} \\ &= (g^{R_{i,j}} \times g^{K_{i+1,2j+1}})^{S_{i,j}} \times (g^{S_{i,j}})^{K_{i+1,2j}}. \end{aligned}$$

The group key $K_{0,0}$ is computed recursively using Equations (1) and (2).

Note that the proper key $K_{i,j}$ of node $< i, j >$ contains the proper keys of its two children $K_{i+1,2j}$ and $K_{i+1,2j+1}$. In addition, it contains the secret key $S_{i,j}$ and hidden key $R_{i,j}$ generated by the leader.

If we would like to use the concept of blinded keys as in TGDH, we can re-write $K_{i,j}$ as follows:

$$\begin{aligned} K_{i,j} &= B_{i+1,2j} \times (g^{S_{i,j}})^{K_{i+1,2j+1}} \\ &= B_{i+1,2j+1} \times (g^{S_{i,j}})^{K_{i+1,2j}}, \end{aligned}$$

where

$$B_{i+1,2j} = (g^{R_{i,j}} \times g^{K_{i+1,2j}})^{S_{i,j}},$$

and

$$B_{i+1,2j+1} = (g^{R_{i,j}} \times g^{K_{i+1,2j+1}})^{S_{i,j}}$$

The secret key $S_{i,j}$ and the hidden key $R_{i,j}$ generated by the leader are used to assure that as long as the leader's contribution is chosen at random, even a coalition of all other parties will not be able to have any means of controlling the final value of the group key. Therefore, the protocol are fairer and more secure in order to prevent some parties having any kind of advantage over the others.

Group key pair will then be calculated as the group public key PrKG = $K_{0,0}$ and the group private key PuKg = $g_K(0, 0)$. The group leader generates a shared secret key K, encrypts shared data with K. The group leader then encrypt K with the group public key PrKG and upload the digital envelop and the encrypted data to the cloud.

2.2 CTDH Implementation

2.2.1 Group Key Refresh/Reinforce

The group key may need to change periodically and may not be related to any change of group membership. The purpose of refreshing the group key periodically is to prevent an adversary from having a sufficient time or resources to break the key. This process is initiated and performed by the group leader which generates a new secret key and a new hidden key, computes and multicasts the blinded value to the whole group. We show the group key refresh/reinforce operation below.

- Randomly choosing a non-leaf node $< i, j >$ and generating a new secret key $S'_{i,j}$ and a new hidden key $R'_{i,j}$ to replace $S_{i,j}$ and $R_{i,j}$.
- Updating the proper keys and partial keys as a result of the above secret key and hidden key changes.
- Broadcasting the updated partial keys and $g^{S'_{i,j}}$ to all members in the group.
- Each member re-computes the group key using the algorithm described in Section 2.1.2.

2.2.2 Join

When a new member joins the group, the member will sends a join request to the group leader. The join request triggers the rekeying procedure. The leader determines the insertion point in the key tree. It then adds a new member node and a new internal node. Once the key tree is updated, the leader generates a secret key $S_{i,j}$ and a hidden key $R_{i,j}$ for the new internal node $< i, j >$. Next, the leader computes all blinded keys and broadcast them with $g^{S_{i,j}}$ ($S_{i,j}$ is the secret key of new internal node $< i, j >$). All members can then compute the new group key.

Group members do not need to maintain a copy of the key tree like in TGDH as they only need to know the blinded keys of its sibling along the path to the root node. Also the future rekeying does not require member having addition information about the key tree. Thus, CTDH only requires member to know the least information to compute the group key.

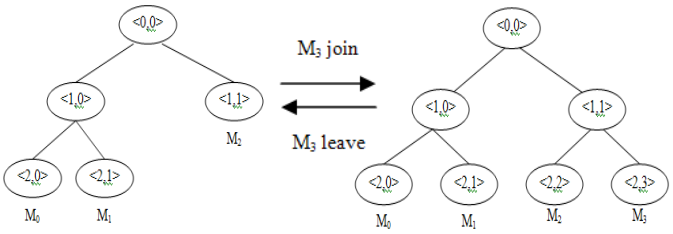


Figure 3: An example of CTDH implementation

Following is an example that illustrates the rekeying operations.

In Figure 3, when member M_3 wishes to join the group by sending its own blinded key to the group leader. The group leader performs the following operations to update the group key:

- Creating two new nodes as children of node <1, 1>.
- Moving M_2 and its proper key from node <1, 1> to node <2, 2>.
- Assigning node <2, 3> to the new member M_3. Member M_3 generates its own secret key $s_{2,3}$ and sends $g^{s_{2,3}}$ to the group leader.
- In general, the proper keys of the nodes on the path from the new member to the root are updated. Since node $<1,1>$ becomes a non-leaf node, the group leader needs to assign it with a secret key $S_{1,1}$ and a hidden key $R_{1,1}$.
- Updating the following blinded keys as a result of the above secret key and hidden key changes and tree restructuring: $B_{2,2}$, $B_{2,3}$ (two newly created partial keys) and $B_{1,1}$. Note that the blinded keys $B_{2,0}$, $B_{2,1}$, $B_{1,0}$ are not changed. The new blinded keys $B_{2,2}$, $B_{2,3}$ and $B_{1,1}$ are computed as follows:

$$\begin{aligned} B_{2,2} &= (g^{R_{1,1}} \times g^{s_{2,2}})^{S_{1,1}} \\ B_{2,3} &= (g^{R_{1,1}} \times g^{s_{2,3}})^{S_{1,1}} \\ B_{1,1} &= (g^{R_{0,0}} \times g^{K_{1,1}})^{S_{0,0}}. \end{aligned}$$

- Broadcasting the updated blinded keys $B_{2,2}$, $B_{2,3}$,$B_{1,0}$, $B_{1,1}$ and $g^{S_{1,1}}$ to all members in the group. Each member re-computes the group key using the algorithm described in Section 2.1.2.

  For example, member M_0 and M_1 can compute the new group key as follows:

$$K_{0,0} = B_{1,1} \times (g^{S_{0,0}})^{K_{1,0}}$$

  Member M_2 and M_3 can compute $K_{1,1}$, $K_{0,0}$ as follows:

$$\begin{aligned} K_{1,1} &= B_{2,3} \times (g^{S_{1,1}})^{s_{2,2}} \\ &= B_{2,2} \times (g^{S_{1,1}})^{s_{2,3}} \\ K_{0,0} &= B_{1,0} \times (g^{S_{0,0}})^{K_{1,1}} \end{aligned}$$

The group key pair shared by all group members will then be updated as the group private key PrKG = $K_{0,0}$ and the group public key PuKg = $g_K(0,0)$.

2.2.3 Leave

When an existing member leaves the group, the group key needs to be changed as well. We start with n member and assume member M_l leaves the group. The leader in this case first update the key tree by deleting the leaf node corresponding to M_l. The former sibling of M_l is promoted to replace M_l's parent node. That is to say that an internal node, the parent of M_l, is also deleted. Once the key tree is updated, the leader generates a new secret key $S_{i,j}$ and a new hidden key $R_{i,j}$ for the new parent node $< i, j >$ of M_l's former sibling. Next, the leader computes all blinded keys and broadcast them with $g_{S_{i,j}}$. This allows all members to compute the new group key.

Looking at the setting that Figure 3 shows, if member M_3 leaves the group, the leader performs the following rekeying procedures.

First, node $< 1, 1 >$, $< 2, 3 >$ are deleted. After updating the key tree, the leader picks a new secret key $S_{0,0}$ and a new hidden key $R_{0,0}$ for node $< 0, 0 >$.

- Removing node $< 2, 3 >$.
- Replace internal node $< 1, 1 >$ with member M_2's node.
- Generate a new secret key $S'_{0,0}$ and a new hidden key $R'_{0,0}$ for node $< 0, 0 >$. As node $< 1, 1 >$ becomes a leaf node, no secret key generated by the leader are needed for this node. Thus, previously generated secret key $R_{<1,1>}$ and $S_{<1,1>}$ by the group leader are discarded.
- Updating the proper keys $K_{<0,0>}$ (the group key) as a result of the above tree restructuring. In general, the proper keys of the nodes on the path from the new member to the root are updated.
- Updating the partial keys $B_{<1,0>}$, $B_{<1,1>}$ as a result of the above secret key changes and tree restructuring.

$$B_{1,0} = (g^{R'_{0,0}} \times g^{K_{1,0}})^{S'_{0,0}}$$

$$B_{1,1} = (g^{R'_{0,0}} \times g^{s_{1,1}})^{S'_{0,0}}$$

- Broadcasting the updated partial keys $B_{1,0}$, $B_{1,1}$, and $g^{S'_{0,0}}$ to all members in the group. Each member re-computes the group key using the algorithm described in Section 2.1.2.

  For example, member M_0 and M_1 can compute the new group key as follows:

$$K_{<0,0>} = B_{<1,1>}(g^{S'_{<0,0>}})^{K_{<1,0>}}$$

  Member M_2 can compute the new group key as follows:

$$K_{<0,0>} = B_{<1,0>}(g^{S'_{<0,0>}})^{s_{<1,1>}}$$

The group key pair shared by all group members will then be updated as the group public key $\text{PuKg} = g_K(0,0)$ and the group private key $\text{PrKG} = K_{0,0}$.

2.2.4 Updates on the Cloud Server

When a group member joins or leaves, all digital envelopes related to the sharing data in this group should be also updated and encrypted by the new group public key. After the group public key and the group private key are updated, the group leader will re-compute a proxy re-encryption key from the version of group public key (PuKG) used in the existing digital envelopes to the new updated version (PuKG1). The leader then uploads the updated information into the cloud. With this proxy re-encryption key, cloud servers can update all existing digital envelopes to be encrypted under the new updated group public key PuKG1. This method can delegate most of the computation intensive operations to cloud servers without disclosing the encrypted data contents and keys in all digital envelopes.

2.2.5 Upload and Download Data

Before uploading a file to cloud servers, the data owner encrypts the file with a randomly chosen session key K. Together with uploading the encrypted sharing file the data owner also uploads a digital envelope (asymmetrically encrypt the session key K with the group public key PuKG), which is currently used.

When a group member requests to download a file he/she sends a request to cloud servers. Cloud servers send the encrypted file and the specific digital envelope to the group member. The group member decrypts the digital envelop to get key K and then decrypts the request file using key K.

2.3 Security Proof

Assuming the DH algorithm is secure, we show that CTDH is secure using a binary tree as an example. In this proof, all operations are assumed to be performed mod p.

Theorem 1. *Let T be a key tree of height m. l be the level of T. Let $s_0, \ldots, s_n$ be secret keys of the members of the group. The shared group key K derived by any member in the application of CTDH is secure.*

Proof. The proof is based on induction over level l ($1 \leq l \leq m$).

Base Case: $l = 1$. Let member M_0 and M_1 be rooted at node $< m-1, 0 >$ as shown in figure 4. Denote the secret keys of M_0 and M_1 as $s_{m,0} = s_0$ and $s_{m,1} = s_1$ respectively. Let $S_{m-1,0}$ and $R_{m-1,0}$ be the secret key and hidden key of the node $< m-1, 0 >$ generated by the group leader. We show that the shared key K derived by any member in the group is secure.

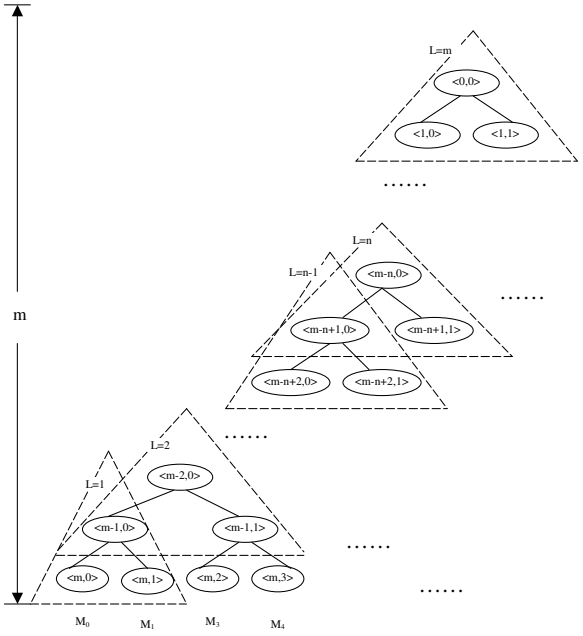


Figure 4: An example of group key calculation

The group leader computes the blinded key for member M_0 and M_1 as follows.

$$B_{m,0} = (g^{R_{m-1,0}} \times g^{s_{m,0}})^{S_{m-1,0}},$$

and

$$B_{m,1} = (g^{R_{m-1,0}} \times g^{s_{m,1}})^{S_{m-1,0}}.$$

After receiving $B_{m,0}$, $B_{m,1}$ and $g^{S_{m-1,0}}$ broadcast by the group leader, member M_0 can derive the shared key as $K = K_{m-1,0} = B_{m,1} \times (g^{S_{m-1,0}})^{s_{m,0}}$.

Since the shared key $K_{m-1,0}$ requires member $M_{m,0}$'s secret key, any passive adversary who gets the broadcast message cannot derive the group key without knowing that secret key.

In case the broadcast message is captured, the adversary tries to use $B_{m,0}$, $B_{m,1}$ and $g^{S_{m-1,0}}$ to derive the shared key $K_{m-1,0}$. Since $(g^{S_{m-1,0}})^{s_{m,0}}$ is the result of the two party Diffie-Hellman key exchange, which is assumed to be indistinguishable in polynomial time from a random number.

Moreover, from the public known information of $B_{m,0}$, $B_{m,1}$ and $g^{S_{m-1,0}}$, the adversary also cannot derived the shared key $K_{m-1,0}$ because the hidden key $R_{m-1,0}$ has never been released to anyone, even the group member. Only the group leader knows what the hidden key it factors into the shared key. Thus, from the broadcast information itself, the adversary cannot derive the shared key. Therefore, the shared key $K_{m-1,0}$ is secure.

Induction hypothesis: The shared key K derived by any member in the application of CTDH, at level $l = n-1$ ($n \geq 2$)is secure.

Induction step: At level $l = n$, $K_{m-n+1,0}$ and $K_{m-n+1,1}$ are the shared values of two subtrees rooted at node $< m-n+1, 0 >$ and $< m-n+1, 1 >$, respectively.

The group leader computes the blinded key for node $< m-n+1, 0 >$ and $< m-n+1, 1 >$ as follows

$$B_{m-n+1,0} = (g^{R_{m-n,0}} \times g^{s_{m-n+1,0}})^{S_{m-n,0}},$$

and

$$B_{m-n+1,1} = (g^{R_{m-n,0}} \times g^{s_{m-n+1,1}})^{S_{m-n,0}}$$

respectively.

After receiving $B_{m-n+1,0}$, $B_{m-n+1,1}$ and $g^{S_{m-n,0}}$ broadcast by the group leader, any member (such as M_0) who belongs to the subtree rooted at node $< m-n+1, 0 >$ can derive the shared key at level n as follows:

$$K_{<m-n,0>} = B_{m-n+1,1} \times (g^{S_{m-n,0}})^{K_{m-n+1,0}}.$$

From the induction hypothesis above, we know that $K_{m-n+1,0}$ is secure and only known by the member who are the children of this subtrees. Thus, the shared group key $K_{<m-n,0>}$ at level n is secure. We can conclude that the shared key of any level l $(1 \leq l \geq m)$ is secure. Therefore, the shared group key $K = K_{0,0}$ of level m derived by any member in the application of CTDH is secure. Furthermore, if extending the inductive method to n-ary key tree, we can apply the same derivation method and show that the derived group key of n-ary key tree is secure as well. □

Theorem 2. *When a new member joins or an exiting member leaves the group, the application of CTDH is still secure. We use Figure 3 as an an example to shown that CTDH provide backward secrecy and forward secrecy.*

First, CTDH provides backward secrecy. Backward secrecy states that a new member who knows the current group key cannot derive any previous group keys. In Figure 3, once a new member M_3 joins the key tree, the group leader for this join event generates a new secret key and a new hidden key and must involve the new member's secret key into the new group key, consequently, previous group key is changed. Therefore, the information learned by the new member with respect to the prior key tree is exactly same as the information of an outsider. Hence, the new member does not gain any advantage compared to a passive adversary.

Second, CTDH provides forward secrecy. Forward secrecy requires that a member who knows a contiguous subset of old group keys cannot discover subsequent group keys once it leaves the group. In Figure 3, once the current member M_3 leaves the group, the group leader refreshes the secret key and hidden key, therefore, all keys known to leaving members will be changed accordingly. Moreover, the new group key will subtract the member M_3's secret. Thus, the information learned by M_3 with respect to the new key tree is exactly same as the information of a passive adversary. This proves that CTDH provides both backward and forward secrecy.

3 Evaluation

In this section, we evaluate our proposed CTDH protocol with Tree-based Group Diffie-Hellman (TGDH) protocol in terms of computation and communication costs.

3.1 Computation Analysis

Table 1 shows computation comparison between TGDH and CTDH when a new member joins the group or an existing member leaves the group. CTDH needs to compute $3\log_2 n + 3$ exponentials while $4\log_2 n$ exponentials are required by TGDH upon joining. Upon leaving, $2\log_2 n + 1$ exponentials are computed using CTDH, where $3\log_2 n - 3$ exponentials are needed by TGDH.

Table 1: Computation comparison

Description		TGDH	CTDH
Join	Group Leader	$2\log_2 n$	$\log_2 n + 3$
	Existing member	at most $\log_2 n - 1$	at most $\log_2 n - 1$
	New member	$\log_2 n + 1$	$\log_2 n + 1$
	Total exponentials	$4\log_2 n$	$3\log_2 n + 3$
Leave	Group Leader	$2\log_2 n - 2$	$\log_2 n + 2$
	Existing member	at most $\log_2 n - 1$	at most $\log_2 n - 1$
	Total exponentials	$3\log_2 n - 3$	$2\log_2 n + 1$

Figure 5 shows that if there are 8 members before a new member joins the group, CTDH has same exponential computation as TGDH. Only if there is less than 8 members in the group before new member joins, TGDH requires 1 or 2 exponential computations less than CTDH. Since these 1 or 2 more exponentials are actually computed by the powerful node in CTDH, they should not cause any big delay comparing with TGDH. If there is more than 8 member in the group before new member joins, CTDH outperforms TGDH.

Figure 6 shows that if there are 16 members before a member leaves the group, CTDH has same exponential computation as TGDH. Only if there are less than 4 members in the group before anyone leaves, TGDH requires 1 or 2 exponential computations less than CTDH. Since group leader is much more powerful than a mobile device in terms of computation, memory, power supply, these 1 or 2 exponentials should not cause any big delay comparing with TGDH. If the gorup has more than 16 members before a member leaves, CTDH outperforms TGDH.

3.2 Communication Analysis

Upon member joins the group, TGDH needs 2 broadcast messages to distribute the partial keys, while CTDH needs 1 unicast and 1 broadcast to update the partial keys. Although CTDH and TGDH require same number

of messages, one unicast has less communication cost than a broadcast when sending same size message.

A more relevant measure for a group key management is the latency that a user experiences from the moment the group change was detected, until the new secure group is established. This time is greater than only analytical cryptographic cost, since it includes network latency. Our simulation results in Figures 7- 8 shows that CTDH provides best communication efficiency. Thus, CTDH perform much better than TGDH in terms of communication delay upon joining or leaving.

4 Conclusions

In this paper, a scalable and efficient group sharing method, CTDH, is proposed for public cloud. CTDH is scalable and efficient thanks to the key tree structure. It uses the contributory approach based on the Diffie-Hellman cryptographic algorithm to avoid the need of setting up pairwise secure channels among members We briefly analyze the security and performance of CTDH. We conclude that CTDH not only provides same level of security as TGDH, but also outperforms TGDH in terms of computation and communication costs. In the future, we would like to present a secure and fault-tolerant key agreement for group data sharing in a cloud storage scheme.

References

[1] G. Ateniese, K. Fu, M. Green and S. Hohenberger, "Improved proxy re-encryption schemes with applications to secure distributed storage," *ACM Transactions on Information and System Security*, vol. 9, no. 1, pp. 1–30, 2006.

[2] M. G. G. Ateniese, K. Fu and S. Hohenberger, "Improved proxy reencryption schemes with applications to secure distributed storage," *ACM Transactions on Information and System Security*, vol. 9, no. 1, pp. 1-30, 2006.

[3] C. I. Fan, "Provably secure timed-release proxy conditional reencryption," *IEEE Systems Journal*, vol. 11, no. 4, 2017.

[4] L. Garton, C. Haythornthwaite and B. Wellman, "Studying online social networks," *Journal of Computer-Mediated Communication*, vol. 3, no. 1, pp. 75-106, 2006.

[5] L. Greenwald, K. Rohloff and D. Stott, "Secure proxy-reencryption-based inter-network key exchange," in *IEEE Military Communications Conference (MILCOM)*, pp. 780-785, 2018.

[6] S. Kamara and K. Lauter, "Cryptographic cloud storage," in *International Conference on Financial Cryptography and Data Security*, pp. 136-149, 2010.

[7] Y. Kim, A. Perrig and G. Tsudik, "Tree-based group key agreement," *ACM Transactions on Information and System Security*, vol. 7, no. 1, pp. 60-96, 2004.

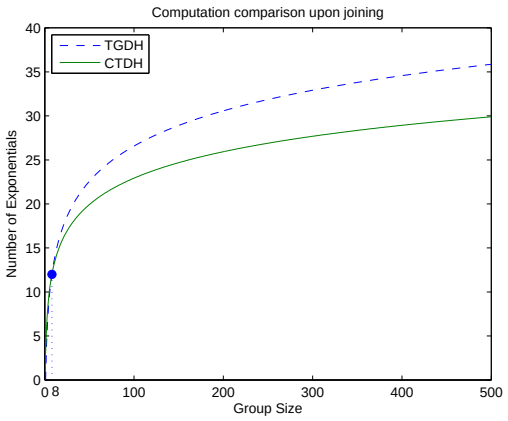


Figure 5: Computation comparison upon joining

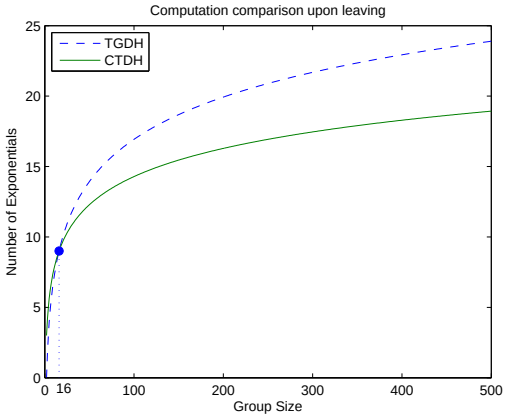


Figure 6: Computation comparison upon leaveing

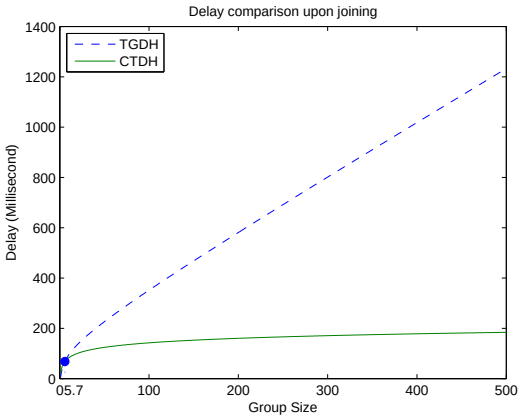


Figure 7: Delay comparison upon member joining

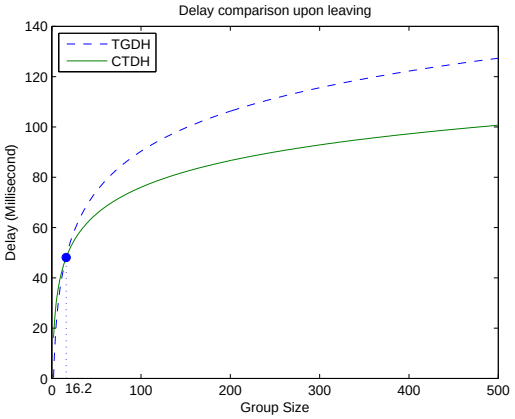


Figure 8: Delay comparison upon member leaving

[8] L. S. Lai and E. Turban, "Groups formation and operations in the web2.0 environment and social networks," *Group Decision Negotiation*, vol. 17, no. 5, pp. 387-402, 2008.

[9] X. Liu, Y. Zhang, B. Wang and J. Yan, "Mona: Secure multi-owner data sharing for dynamic groups in the cloud," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 6, pp. 1182–1191, June 2013.

[10] Z. C. J. Shao, P. Liu and G. Wei, "Multi-use unidirectional proxy reencryption," in *IEEE International Conference on Communications*, pp. 1-5, 2011.

[11] W. Song, "A practical group key management algorithm for cloud data sharing with dynamic group," *China Communications*, vol. 16, no. 6, 2016.

[12] V. R. Thakare and K. J. Singh, "Ternary tree based TGDH protocol for dynamic secure group data sharing in healthcare cloud," in *International Conference on Inventive Computation Technologies (ICICT'16)*, 2016. DOI: 10.1109/INVENTIVE.2016.7823294.

[13] C. Wang, Q. Wang, K. Ren, W. Lou, "Ensuring data storage security in Cloud Computing," in *The 17th International Workshop on Quality of Service*, pp. 1-9, 2009.

[14] C. Wong, M. Gouda and S. Lam, "Secure group communications using key graphs" *IEEE/ACM Transactions Networking*, vol. 8, pp. 16-30, 2000.

[15] K. Xue, P. Hong, "A dynamic secure group sharing framework in public cloud computing," *IEEE Transactions on Cloud Computing*, vol. 2, no. 4, 2014

[16] O. Yağan, "Zero-one laws for connectivity in inhomogeneous random key graphs," *IEEE Transactions on Information Theory*, vol. 62, no. 8, 2016.

[17] C. H. Yeh, Y. M. Huang, T. I. Wang and H. H. Chen, "DESCV—A secure wireless communication scheme for vehicle ad hoc networking," *Mobile Networks and Applications*, vol. 14, no. 5, pp. 611-624, 2009.

[18] Y. Zhang, "Authorized identity-based public cloud storage auditing scheme with hierarchical structure for large-scale user groups," *China Communications*, vol. 11, no. 15, 2018.

Biography

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

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

A Note on One Popular Non-Interactive Zero-Knowledge Proof System

Zhengjun Cao[1,2], Xiqi Wang[1], and Lihua Liu[3]
(Corresponding author: Lihua Liu)
Department of Mathematics, Shanghai University, No.99, Shangda Road, Shanghai, China[1]
State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications[2]
No.10, Xitucheng Road, Beijing, China
Department of Mathematics, Shanghai Maritime University, No.1550, Haigang Ave, Shanghai, China[3]
(Email: liulh@shmtu.edu.cn)
(Received May 4, 2019; Revised and Accepted Sept. 4, 2019; First Online Sept. 21, 2019)

Abstract

At Eurocrypt'06, Groth *et al.* have proposed one non-interactive zero-knowledge (NIZK) proof system for plaintext being 0 or 1 [its revision published by J. ACM, 59(3), 1-35, 2012]. Based on the system, they presented the first perfect NIZK argument system for any NP language and the first secure NIZK argument with universal composability for any NP language in the presence of a dynamic/adaptive adversary. In this note, we remark that in the scheme the prover is not compelled to invoke any trapdoor key to generate witnesses. The mechanism is dramatically different from the previous works, such as Blum-Feldman-Micali proof system and Blum-Santis-Micali-Persiano proof system. We find if the trapdoor key is available to the prover then he can cheat the verifier to accept a false claim. The characteristic is essentially incompatible with the general primitive of zero-knowledge proof, which does not require any extra trust.

Keywords: Bilinear Groups with Composite Order; Extended Euclid Algorithm; Non-interactive Zero-knowledge Proof; Subgroup Decision Problem

1 Introduction

Non-interactive zero-knowledge (NIZK) proof in the common random string model, introduced by Blum *et al.* [4], plays a key role in many constructions, including digital signatures [11, 25], E-voting [14], Shuffle [2, 27], polynomial evaluation [3], arithmetic circuits [7, 8] and multiple-party computation [1, 9, 20, 26]. In 1988, Blum *et al.* [4] constructed some computational NIZK proof systems for proving a single statement about any NP language. In 1991, they [5] presented the first computational NIZK proof system for multiple theorems. These systems are based on the hardness of deciding quadratic residues modulo a composite number. In 1998, Kilian and Petrank [21] designed an efficient noninteractive zero-knowledge proof system for NP with general assumptions.

In 1999, Feige *et al.* [10] developed a method to construct computational NIZK proof systems based on any trapdoor permutation. Goldreich *et al.* [13] discussed the possibility of converting a statistical zero knowledge (SZK) proof into a NIZK proof. In 2001, Santis *et al.* [23, 24] investigated the robustness and randomness-optimal characterization of some NIZK proof systems. In 2003, Sahai and Vadhan [22] presented an interesting survey on SZK. Groth *et al.* [15, 16, 19] designed some linear algebra with sub-linear zero-knowledge arguments and short pairing-based NIZK arguments. In 2015, Gentry *et al.* [12] discussed the problem of using fully homomorphic hybrid encryption to minimize NIZK proofs.

At Eurocrypt'06, Groth, Ostrovsky and Sahai [17] designed a popular NIZK proof system for plaintext being 0 or 1 using bilinear groups with composite order. The refined version [18] was published by Journal of ACM in 2012. The behind intractability of this work is the subgroup decision problem introduced by Boneh *et al.* [6]. Based on the basic NIZK proof system, they presented one NIZK proof for circuit satisfiability. Furthermore, they constructed the first perfect NIZK argument system for any NP language and the first secure NIZK argument with universal composability for any NP language in the presence of a dynamic/adaptive adversary. They claimed it has resolved a central open problem concerning NIZK protocols.

In this note, we show that in Groth-Ostrovsky-Sahai proof system the prover is not compelled to invoke any trapdoor key to generate witnesses. The mechanism was dramatically different from the previous works, such as Blum-Feldman-Micali proof system [4] and Blum-Santis-Micali-Persiano proof system [5]. They did adopt a different security model although it was not specified explicitly. We also find that if the trapdoor key is available to the prover then he can cheat the verifier to accept a false

claim. The characteristic is radically incompatible with the general primitive of zero-knowledge proof. That is, the popular NIZK proof system requires extra trust to set its parameters. This shortcoming renders itself vulnerable to inner attacks.

2 Review of Groth-Ostrovsky-Sahai NIZK Proof System

The system [17] can be described as follows.

- Common reference string. Let $\mathbb{G}, \mathbb{G}_1$ be two cyclic groups of order n, where $n = pq$ and p, q are primes such that it is difficult to factor n. $\hat{e} : \mathbb{G} \times \mathbb{G} \to \mathbb{G}_1$ is a bilinear map. We require that $\hat{e}(g, g)$ is a generator of $\mathbb{G}_1$ if g is a generator of $\mathbb{G}$. Pick a generator $h \in \mathbb{G}_q$, where $\mathbb{G}_q \subset \mathbb{G}$ is of order q. The common reference string is $\sigma = (n, \mathbb{G}, \mathbb{G}_1, \hat{e}, g, h)$.

- Statement. The statement is an element $c \in \mathbb{G}$. The claim is that there exists a pair $(m, w) \in \mathbb{Z}^2$ so $m \in \{0, 1\}$ and $c = g^m h^w$.

  Proof. Given (σ, c, m, w), check $m \in \{0, 1\}$ and $c = g^m h^w$. Return failure if check fails. It proceeds as follows. Pick $r \in \mathbb{Z}_n^*$, compute

$$\begin{aligned} \pi_1 &= h^r \\ \pi_2 &= (g^{2m-1} h^w)^{wr^{-1}} \\ \pi_3 &= g^r. \end{aligned}$$

  Return $\pi = (\pi_1, \pi_2, \pi_3)$. □

- Verification. Given the parameter σ and c, π, check $c \in \mathbb{G}$, $\pi \in \mathbb{G}^3$, and verify that

$$\begin{aligned} \hat{e}(c, cg^{-1}) &= \hat{e}(\pi_1, \pi_2) \\ \hat{e}(\pi_1, g) &= \hat{e}(h, \pi_3). \end{aligned}$$

Correctness . It is easy to check that

$$\begin{aligned} \hat{e}(c, cg^{-1}) &= \hat{e}(g^m h^w, g^{m-1} h^w) \\ &= \hat{e}(g, g)^{m(m-1)} \hat{e}(g, h)^{(2m-1)w} \hat{e}(h, h)^{w^2} \\ \hat{e}(\pi_1, \pi_2) &= \hat{e}(h^r, (g^{2m-1} h^w)^{wr^{-1}}) \\ &= \hat{e}(g, h)^{(2m-1)w} \hat{e}(h, h)^{w^2} \end{aligned}$$

If $m \in \{0, 1\}$, then $\hat{e}(c, cg^{-1}) = \hat{e}(\pi_1, \pi_2)$.

If $m \notin \{0, 1\}$, it seems that Alice has to solve the discrete logarithms among $\hat{e}(g, h), \hat{e}(g, g), \hat{e}(h, h)$, which are reduced to the discrete logarithm of h to g. This possibility can be eradicated in advance by asking Alice and Bob agree to a random seed to a pseudorandom generator for generating g. Based on the observations, Groth *et al.* concluded that the scheme was secure against either the prover's attack or the verifier's attack.

3 Analysis of Groth-Ostrovsky-Sahai NIZK Proof System

For convenience, we will call the prover, Alice, and the verifier, Bob. We now consider the following problems.

3.1 What is the True Statement

Give $c \in \mathbb{G}$, Alice claims that c is of the form $g^m h^w$ for some $(m, w) \in \{0, 1\} \times \mathbb{Z}_n$. This is equivalent to checking whether c or c/g is in the subgroup $\mathbb{G}_q$.

If the trapdoor key q is available, then it suffices to check that

$$c^q = 1, \quad \text{or} \quad (c/g)^q = 1.$$

However, the trapdoor key cannot be directly shown to Bob. Hence, Alice has to produce some witnesses to convince Bob of that c or c/g is indeed in the subgroup $\mathbb{G}_q$.

3.2 How to Understand the Phrase of "Common Reference String"

The notion of "common reference string" used in NIZK model can be traced back to [5]. It had stressed that

> The moral is that one must be careful when using the same set-up, i.e., common reference string, and the same pair (x, y), to prove an "unlimited" number of formulae to be satisfiable.

Apparently, *"common reference string" represents the same set-up known to the prover and the verifier.* But it does not specify whether or not there is any trapdoor key related to the common reference string.

Recalling Blum-Santis-Micali-Persiano proof system [5] and its like, we find they have not any trapdoor key at all. For readers' convenience, we now briefly relate Blum-Santis-Micali-Persiano proof system as follows.

Common reference string. The random string is $\rho = \rho_1 \rho_2 \cdots \rho_{n^2}$, each ρ_i has length n.

Statement. The odd number $x < n$ is a composite of two different primes p, q. Assume that $|J_x^{+1}| = |J_x^{-1}|$, where

$$J_x^{+1} = \left\{ y \in \mathbb{Z}_x^* \,|\, \text{Jacobi symbol} \left(\frac{y}{x}\right) = 1 \right\},$$
$$J_x^{-1} = \left\{ y \in \mathbb{Z}_x^* \,|\, \text{Jacobi symbol} \left(\frac{y}{x}\right) = -1 \right\}$$

and $\mathbb{Z}_x^* = \{1, 2, \cdots, x-1\}$. Alice knows p, q and wants to convince Bob of this fact while preventing Bob from knowing p, q.

Proof. Alice picks $y < x$ such that $\left(\frac{y}{x}\right) = 1$ and y is not a quadratic residue of x. She then computes $\left(\frac{\rho_i}{x}\right)$ for $i = 1, 2, \cdots, n^2$. If $\left(\frac{\rho_i}{x}\right) = 1$, compute s_i such that $s_i^2 = \rho_i \bmod x$ or $s_i^2 = y\rho_i \bmod x$. Send these s_i and x, y to Bob.

Verification. Bob checks that x is not a perfect square. Verify that $\left(\frac{y}{x}\right) = 1$ and the number of s_i is greater than $3n$. He then checks that each $\left(\frac{\rho_i}{x}\right) = 1$ and $s_i^2 = \rho_i \bmod x$ or $s_i^2 = y\rho_i \bmod x$.

It is easy to find that in [5] there is not any trapdoor key related to the setup. We refer to Table 1 for the big differences between Blum-Santis-Micali-Persiano proof system and Groth-Ostrovsky-Sahai proof system.

Clearly, Blum-Santis-Micali-Persiano proof system needs only *a very simple common reference string*, and Alice has to make use of her private key (the factors of x) to generate witnesses. To the contrary, Groth-Ostrovsky-Sahai proof system needs *a very complicated common reference string associated with a trapdoor key*.

The model introduced by Blum *et al.* is more suitable to practical applications because *it does not require any extra trust*. But the model considered by Groth *et al.* entails the verifier to trust that the related trapdoor key cannot be accessed to the prover (see the discussion in the following sections). The requirement does contradict the general assumptions for zero-knowledge proof.

3.3 Alice is not Compelled to Invoke Any Trapdoor Key

It is easy to find that Alice does not invoke the trapdoor key (p, q) to generate witnesses. Besides, the system does not specify who is responsible for generating the common reference string. So, it is reasonable to assume that there is a third-party, Cindy, who generates the common reference string. Note that Cindy is not fully trustable and she knows the trapdoor key. Otherwise, the presence of a fully trustable party is indeed incompatible with the primitive of zero-knowledge proof system.

3.4 Alice and Cindy Can Conspire to Cheat Bob

Can Cindy form an alliance with Alice? If so, we now show that Alice and Cindy can conspire to cheat Bob to accept a false claim.

Suppose that Alice picks an integer r and sets

$$\begin{aligned} \pi_1 &= h^r \\ \pi_3 &= g^r \\ c &= g^{\alpha_1} h^{\alpha_2} \\ \pi_2 &= g^{\beta_1} h^{\beta_2}, \end{aligned}$$

where $\alpha_1, \alpha_2, \beta_1, \beta_2$ are to be determined. Since

$$\begin{aligned} \hat{e}(c, cg^{-1}) &= \hat{e}(g^{\alpha_1} h^{\alpha_2}, g^{\alpha_1 - 1} h^{\alpha_2}) \\ &= \underline{\hat{e}(g,g)^{\alpha_1(\alpha_1-1)}} \hat{e}(g,h)^{\alpha_1\alpha_2 + \alpha_2(\alpha_1-1)} \\ &\qquad \cdot \hat{e}(h,h)^{\alpha_2^2}, \\ \hat{e}(\pi_1, \pi_2) &= \hat{e}(h^r, g^{\beta_1} h^{\beta_2}) \\ &= \hat{e}(h,g)^{r\beta_1} \hat{e}(h,h)^{r\beta_2}, \end{aligned}$$

it suffices for Alice to solve the following equations

$$\left\{ \begin{array}{rcl} \alpha_1(\alpha_1 - 1) & = & 0 \bmod n \\ 2\alpha_1\alpha_2 - \alpha_2 & = & r\beta_1 \bmod n \\ \alpha_2^2 & = & r\beta_2 \bmod n \end{array} \right. \quad (1)$$

for those exponents. *The authors [17] mistakenly thought that α_1 in the equations (1) has to take 0 or 1.*

In fact, armed with the trapdoor key p, q, Alice can obtain k, ℓ using extended Euclid algorithm such that

$$kq - \ell p = 1.$$

She sets $\underline{\alpha_1 = kq}$. Clearly,

$$\alpha_1(\alpha_1 - 1) = kq(kq - 1) = kq\ell p \equiv 0 \bmod n.$$

She then picks $\underline{\beta_1 < n}$ and computes

$$\begin{aligned} \alpha_2 &= r\beta_1(2kq - 1)^{-1} \bmod n \\ \beta_2 &= \alpha_2^2 r^{-1} \bmod n. \end{aligned}$$

It is easy to check that the above values c, π_1, π_2, π_3 pass the original verification.

Obviously, $\alpha_1 = kq \neq 0, 1$. Besides,

$$(g^{\alpha_1})^q = (g^{kq})^q = (g^{\ell p + 1})^q = g^q \neq 1, \text{ i.e., } g^{\alpha_1} \notin \mathbb{G}_q.$$

Thus, there does not exist an integer α' such that

$$g^{\alpha_1} = h^{\alpha'}.$$

That means $c = g^{\alpha_1} h^{\alpha_2}$ *cannot be eventually expressed as* h^{w_1} *or* gh^{w_2}. Therefore, the adversary can cheat Bob to accept the false claim $c = g^{\alpha_1} h^{\alpha_2}$, where $\alpha_1 \neq 0$ or 1.

4 Conclusion

We remark that the Groth-Ostrovsky-Sahai proof system adopts an artificial security model due to the existence of trapdoor key related to the common reference string. Under the strong assumption that the adversary cannot access to the trapdoor key, the proof system seems secure. But the assumption is ultimately incompatible with the general primitive of zero-knowledge proof *which does not require any extra trust*, and makes the system itself unsuitable to more broader applications.

We would like to stress that the first thing for designing a cryptographic scheme is to consider what is trusted or untrusted. Otherwise, an assumption for extra trust suffices to ruin the whole system.

Acknowledgements

We thank the National Natural Science Foundation of China (#61411146001), and Open Foundation of State key Laboratory of Networking and Switching Technology (#SKLNST-2016-2-03, Beijing University of Posts and Telecommunications). We gratefully acknowledge the reviewers for their valuable suggestions.

Table 1: Two kinds of common reference strings

	Blum-Santis-Micali-Persiano	Groth-Ostrovsky-Sahai
Common reference string	A random string $\rho = \rho_1\rho_2\cdots\rho_{n^2}$, where each ρ_i is of length n.	$(n, \mathbb{G}, \mathbb{G}_1, \hat{e}, g, h)$ where $n = pq$.
[trapdoor key]	NO	(p, q)
Statement	Knowing the factors of the integer x.	c is of the structure $g^m h^w$ with $(m, w) \in \{0, 1\} \times \mathbb{Z}$.
Proof	$x; y, \{s_i\}$	$c; \pi_1, \pi_2, \pi_3$
Verification	$\left(\frac{\rho_i}{x}\right) = 1$, and $s_i^2 = \rho_i \bmod x$ or $s_i^2 = y\rho_i \bmod x$	$\hat{e}(c, cg^{-1}) = \hat{e}(\pi_1, \pi_2)$, and $\hat{e}(\pi_1, g) = \hat{e}(h, \pi_3)$

References

[1] D. Abbasinezhad-Mood, A. Ostad-Sharif, and M. Nikooghadam, "Design of an anonymous lightweight communication protocol for smart grid and its implementation on 8-bit avr and 32-bit arm," *International Journal of Network Security*, vol. 21, no. 4, pp. 607–617, 2019.

[2] S. Bayer and J. Groth, "Efficient zero-knowledge argument for correctness of a shuffle," in *Proceedings of 31st Annual International Conference on the Theory and Applications of Cryptographic Techniques (EUROCRYPT'12)*, pp. 263–280, 2012.

[3] S. Bayer and J. Groth, "Zero-knowledge argument for polynomial evaluation with application to blacklists," in *Proceedings of 32st Annual International Conference on the Theory and Applications of Cryptographic Techniques (EUROCRYPT'13)*, pp. 646–663, 2013.

[4] M. Blum, P. Feldman, and S. Micali, "Non-interactive zero-knowledge and its applications," in *Proceedings of the 20th Annual ACM Symposium on Theory of Computing (STOC'88)*, pp. 103–112, 1988.

[5] M. Blum and *et al.*, "Noninteractive zero-knowledge," *SIAM Journal on Computing*, vol. 20, no. 6, pp. 1084–1118, 1991.

[6] D. Boneh, E. Goh, and K. Nissim, "Evaluating 2-dnf formulas on ciphertexts," in *Proceedings of 2nd Theory of Cryptography Conference (TCC'05)*, pp. 325–341, Feb. 2005.

[7] J. Bootle and *et al.*, "Efficient zero-knowledge arguments for arithmetic circuits in the discrete log setting," in *Proceedings of 35th Annual International Conference on the Theory and Applications of Cryptographic Techniques (EUROCRYPT'16)*, pp. 327–357, 2016.

[8] J. Bootle and *et al.*, "Efficient zero-knowledge proof systems," in *Proceedings of Foundations of Security Analysis and Design (FOSAD'16)*, pp. 1–31, June 2016.

[9] R. Challa and V. Gunta, "Additively lwe based homomorphic encryption for compact devices with enhanced security," *International Journal of Network Security*, vol. 21, no. 3, pp. 378–383, 2019.

[10] U. Feige, D. Lapidot, and A. Shamir, "Multiple non-interactive zero knowledge proofs under general assumptions," *SIAM Journal on Computing*, vol. 29, no. 1, pp. 1–28, 1999.

[11] J. Garay, P. MacKenzie, and K. Yang, "Strengthening zero-knowledge protocols using signatures," *Journal of Cryptology*, vol. 19, no. 2, pp. 169–209, 2006.

[12] C. Gentry and *et al.*, "Using fully homomorphic hybrid encryption to minimize non-interative zero-knowledge proofs," *Journal of Cryptology*, vol. 28, no. 4, pp. 820–843, 2015.

[13] O. Goldreich, A. Sahai, and S. Vadhan, "Can statistical zero knowledge be made non-interactive? or on the relationship of szk and niszk," in *Proceedings of 19th Annual International Cryptology Conference (CRYPTO'99)*, pp. 467–484, Aug. 1999.

[14] J. Groth, "Non-interactive zero-knowledge arguments for voting," in *Proceedings of 3rd International Conference on Applied Cryptography and Network Security (ACNS'05)*, pp. 467–482, June 2005.

[15] J. Groth, "Linear algebra with sub-linear zero-knowledge arguments," in *Proceedings of 29th Annual International Cryptology Conference (CRYPTO'09)*, pp. 192–208, Aug. 2009.

[16] J. Groth, "Short pairing-based non-interactive zero-knowledge arguments," in *Proceedings of 16th International Conference on the Theory and Application of Cryptology and Information Security (ASIACRYPT'10)*, pp. 321–340, Dec. 2010.

[17] J. Groth, R. Ostrovsky, and A. Sahai, "Perfect non-interactive zero knowledge for np," in *Proceedings of 25th Annual International Conference on the Theory and Applications of Cryptographic Techniques (EUROCRYPT'06)*, pp. 339–358, May 2006.

[18] J. Groth, R. Ostrovsky, and A. Sahai, "New techniques for noninteractive zero-knowledge," *Journal of ACM*, vol. 59, no. 3, pp. 1–35, 2012.

[19] J. Groth and A. Sahai, "Efficient noninteractive proof systems for bilinear groups," *SIAM Journal on Computing*, vol. 41, no. 5, pp. 1193–1232, 2012.

[20] M. S. Hwang, C. C. Lee, and S. T. Hsu, "An elgamal-like secure channel free public key encryption with keyword search scheme," *International Journal of Foundations of Computer Science*, vol. 30, no. 2, pp. 255–273, 2019.

[21] J. Kilian and E. Petrank, "An efficient noninteractive zero-knowledge proof system for np with general assumptions," *Journal of Cryptology*, vol. 11, no. 1, pp. 1–27, 1998.

[22] A. Sahai and S. Vadhan, "A complete problem for statistical zero knowledge," *Journal of ACM*, vol. 50, no. 2, pp. 196–249, 2003.

[23] A. Santis, G. Crescenzo, and G. Persiano, "Randomness-optimal characterization of two np proof systems," in *Proceedings of 6th International Workshop on Randomization and Approximation Techniques (RANDOM'02)*, pp. 179–193, Sep. 2002.

[24] A. Santis and *et al.*, "Robust non-interactive zero knowledge," in *Proceedings of 21st Annual International Cryptology Conference (CRYPTO'01)*, pp. 566–598, Aug. 2001.

[25] C. Y. Tsai, P. F. Ho, and M. S. Hwang, "A secure group signature scheme," *International Journal of Network Security*, vol. 20, no. 2, pp. 201–205, 2018.

[26] C. Y. Tsai and *et al.*, "A publicly verifiable authenticated encryption scheme based on factoring and discrete logarithms," *International Journal of Network Security*, vol. 19, no. 3, pp. 443–448, 2017.

[27] Y. L. Wang, J. J. Shen, and M. S. Hwang, "A novel dual image-based high payload reversible hiding technique using lsb matching," *International Journal of Network Security*, vol. 20, no. 4, pp. 801–804, 2018.

Zhengjun Cao is an associate professor with the Department of Mathematics at Shanghai University. He received his Ph.D. degree in applied mathematics from Academy of Mathematics and Systems Science, Chinese Academy of Sciences. He served as a post-doctor in Computer Sciences Department, Université Libre de Bruxelles, from 2008 to 2010. His research interests include cryptography, discrete logarithms and quantum computation.

Xiqi Wang is currently pursuing his M.S. degree from Department of Mathematics, Shanghai University. His research interests include information security and cryptography.

Lihua Liu is an associate professor with the Department of Mathematics at Shanghai Maritime University. She received her Ph.D. degree in applied mathematics from Shanghai Jiao Tong University. Her research interests include combinatorics, cryptography and information security.

A Perceptual Hash-based Approach to Detect Covert Timing Channels

Linfan Wang and Yonghong Chen
(Corresponding author: Yonghong Chen)

College of Computer Science, Technology, Huaqiao University
No.668 Jimei Avenue, Xiamen, Fujian, China 361021
(Email: iamcyh@hqu.edu.cn)

(Received Apr. 12, 2019; Revised and Accepted Oct. 4, 2019; First Online Dec. 7, 2019)

Abstract

Covert timing channels have received intensive interests in recent years, which are integrated into the existing resources of network systems. This means that the traditional security policy, such as firewalls and intrusion detection system, can not capture them effectively. However, there is still not a generic mechanism used to detect a large variety of covert timing channels. Thus, it is a challenging task to detect and disrupt them. In this paper, we introduce a perceptual hash-based approach to detect covert communications. The proposed approach utilizes the extracted perceptual features from the network traffic and the designed perceptual hash functions to classify the traffic as covert or overt. We extracted the perceptual features from four typical and different covert timing channels and tested each of them independently. The experimental results verify that the perceptual hash-based approach is effective and prove that the perceptual hash technique has great potential for blind detection of covert timing channels.

Keywords: Covert Timing Channels; Network Security; Perceptual Features; Perceptual Hash

1 Introduction

Covert timing channels (CTCs) transmit the covert information using the elements of existing network resources that were not designed for communication. This makes them to evade the detection of network security mechanisms like firewalls and intrusion detection systems (IDS) [5,16,25]. By manipulating the time or ordering of network events (*e.g.*, data packets), CTCs transmit the secret information from a network system with higher security privilege to the Internet. A scenario where CTCs are used is illustrated in Figure 1. Unfortunately, CTCs are used for harmful intentions and that are the ways of damaging network security, which proposes a challenging task to detect a large variety of CTCs.

There are many existing detection methods for CTCs, while these methods have certain limitations. Some detection methods [1, 2, 9] only detect a specific CTCs algorithm which lack a common mechanism. The other methods [7,17,22] could detect most of CTCs algorithms. However, due to the high-speed network environment, these detection methods lack certain detection robustness to capture CTCs. In summary, the existing detection methods for detecting CTCs has some disadvantages in a high-speed network environments. The details will be described in Section 2.

For the shortcomings of existing detection methods above, we introduce a perceptual hash-based approach to detect CTCs. Perceptual hash [6,15,23] is a new technique in multimedia information security. It solves the problem that the traditional hash doesn't have perceptual robustness. Perceptual hash technique is commonly used in fields such as the image [10,19,24] and voice authentication [14,20,26], while has never been applied into the detection of CTCs in network systems. We observe that the extracted perceptual features of CTCs are different from the features of the original process, but there is a striking similarity among the CTCs features of the same class. The similarity and discrimination of perceptual features give new ideas for the detection of CTCs. Therefore, we have studied the way of appling the robustness and discrimination of perceptual hash into the detection of CTCs.

More specifically, we convert the network traffic into the corresponding time series. The discrete wavelet transform (DWT) is designed to extract the perceptual features of time series in frequency-domain, and the information entropy is designed to extract the perceptual features of time series in time-domain. The perceptual features are then transformed into the perceptual sequences by the designed perceptual hash functions. The purpose is to preserve the similarity and discrimination of the perceptual features and reduce data amount by the robustness and summary of perceptual hash. We last perform the detection threshold estimation of perceptual sequences for both covert and legitimate traffic, and determine whether

the covert traffic is accurately detected. To evaluate the proposed approach, we conducted a series of experiments to verify whether it is effective to detect multiple CTCs. The experimental results show that the perceptual hash-based approach has solved the following problems. In the case of samll sample size, the true-positive rates of CTCs is improved by the good discrimination of perceptual features. At the same time, the proposed approach has a good detection robustness compared with other approaches mentioned in this paper in the case of network interference, such as jitter, packets loss, *etc.*

The rest of this paper is structured as follows: Section 2 introduces the background, related works and existing detection approaches of CTCs. In Section 3, the principle and algorithm of the perceptual hash-based approach are described. Section 4 passes the detection experiments of legitimates and four typical CTCs to validate the effectiveness of the proposed approach in this paper. The content of this paper is summarized and our future work direction is discussed in Section 5.

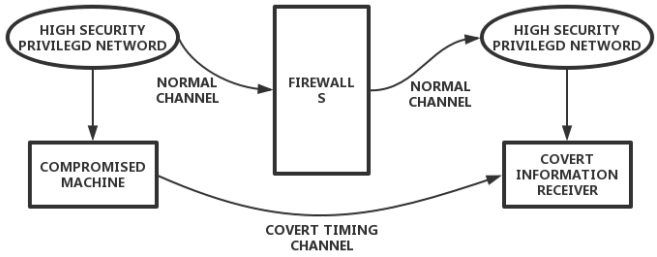


Figure 1: CTCs scenario

2 Background and Related Works

2.1 Threat Analysis of CTCs

There are two classic kinds of CTCs: Active and passive. Active requires communication to generate a additional traffic that is not designed in the network system. Passive transmits the covert information by manipulating the existing traffic in the network system. A scenario in which active and passive CTCs are used is shown in Figure 2. In order to better detect active and passive CTCs, we chose two typical active CTCs, IP - Covert Timing Channel (IPCTC), Time-Replay Covert Timing Channel (TRCTC) and two typical passive CTCs, Distribution-Matching Covert Timing Channel (DMCTC), JitterBug to test the reliability of our approach. The details are introduced in the next sections.

2.1.1 IPCTC

Cabuk *et al.* [3] designed the first IPCTC that is a covert timing channel operating at the IP layer. IPCTC encodes a 1-bit by transmitting a data packet during a inter-packet delay t and encodes a 0-bit by not transmitting packets during a inter-packet delay t. The receiver receives

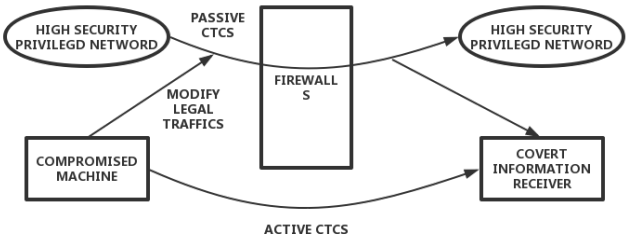


Figure 2: Active and passive CTCs scenario

the size of inter-packet delays to recover the covert information. For the multi-band channels, IPCTC uses the multiple inter-packet delays, each inter-packet delay corresponding to one code. The characteristic of IPCTC is that if the inter-packet delay t is set to a fixed value, the distribution of IPCTC would be close to the geometirc distribution. To avoid the phenomenon, IPCTC changes the value of t according to a certain inter-packet delay set. The communication parties share the set.

2.1.2 TRCTC

Cabuk *et al.* [1] later designed a more advanced covert timing channel - TRCTC, which generates a covert channel by replaying the inter-packet delays of legitimate traffic. TRCTC collects the legitimate traffic as the existing data, and obtains a set S_{in} by sorting the inter-packet delays of legitimate traffic. By finding the intermediate value t_c from S_{in}, S_{in} is divided into two subsets S_0 and S_1. TRCTC encodes a 1-bit by randomly replaying a inter-packet delay t_x from S_0 and encodes a 0-bit by randomly replaying a inter-packet delay T_y from S_1. The communication parties share t_c, so the receiver recovers the covert information by receiving the inter-packet delays and t_c. The reason why it is difficult to detect TRCTC is that its statistical characteristics are close to that of the legitimate traffic.

2.1.3 DMCTC

Liu *et al.* [11] presented a covert timing channel with DMCTC, which is designed to counter the statistical-based detection approaches. DMCTC collects N inter-packet delays of legitimate traffic, and the N inter-packet delays are recorded in L intervals. The boundary block α is found so that the number of inter-packet delays on the left side and on the right side of α are as equal. DMCTC encodes a 1-bit by randomly replaying a inter-packet delay from the left interval of α and encodes a 0-bit by randomly replaying a inter-packet delay from the right interval of α. The receiver calculates the boundary block α of L intervals by the same way, and recovers the covert information according to the N inter-packet delays. DMCTC is similar to TRCTC in that they utilize the legitimate traffic to construct a covert channel. Therefor, the reason why it also is difficult to detect DMCTC is that its distribution characteristic are close to that of the legitimate traffic.

2.1.4 JitterBug

Shah *et al.* [21] designed a covert timing channel called JitterBug. The legitimate package is performed a short delay to construct JitterBug and the communication parties of JitterBug share a value w. JitterBug encodes a 1-bit by increasing an packet delay to a value modulo w and encodes a 0-bit by increasing a packet delay to a value modulo $[w/2]$. The receiver recovers the covert information according to the shared value w. For small values of w, the distribution of JitterBug is similar to that of the legitimate traffic.

2.1.5 Others

Model-Base Covert Timing Channel (MBCTC) [8] is a covert timing channels based on distribution fitting methods. MBCTC counters detection by fitting the statistical distribution of the inter-packet delays of legitimate traffic. The on/off channel [27] is proposed from a stand-alone system, which is a classic binary covert timing channel. TCPScript [12] is a passive network covert timing channels established at the TCP layer. The senders confirm the correctness of the hidden information by observing ACK packet of the receiver, and this method is suitable for constructing a multi-ary channels.

2.2 Detection Tests

There are two broad types of detection approaches: special and generic detection approaches. The special detection approaches are only applied to detect a CTCs algorithm, and the limitation of these approaches is large. The generic detection approaches are able to detect different CTCs algorithms. However, due to the high-speed network environment, they are ineffective to capture CTCs.

2.2.1 Special Detection Tests

Cabuk *et al.* [1] proposed a statistical-based detection approach of TRCTC. Hypothesis Test is used to test whether the set of time intervals of TRCTC and legitimate traffic follow the same distribution. The premise is that it is difficult to get a sorted set of TRCTC, and the set needs to satisfy a certain distribution.

Cabuk *et al.* [9] proposed a detection approach based on laws called *"β-Similarity"* for detecting IPCTC. The authors find that the inter-packet delays variance of legitimate traffic always changes. However, the covert channels don't change the variance of the inter-packet delays when the coding method is unchanged. Later Cabuk *et al.* [2] proposed another laws-based approach for detecting IPCTC called *"Compressibility"*. However, it is not suitable for online detection due to the slow true-positive rates.

2.2.2 Generic Detection Tests

The Kolmogorov-Smirnov Test (K-S test) [17] is a statistical-based detection approach. The approach needs to calculate the distance of the empirical distribution functions of test sample and training sample. To judge whether the inter-packet delays of different channels follow the same distribution according to the distance, thereby the author judges whether the test sample is covert timing channels. Although the K-S test is a generic detection method, it's ineffective for capturing CTCs in a complex and varied network environment. Moreover, the approach can not detect TRCTC, JitterBug and DMCTC.

Gianvecchio *et al.* [7] proposed the most effective detection method currently known as entropy detection. Gianvecchio believe the detection of CTCs is mainly divided into two categories: shape-based detection and rule-based detection. The shape of the inter-packet delays can be described by first-order statistics such as mean, standard deviation and distribution. The regularity of the inter-packet delays can be described by high-order statistics, such as the correlation between data. Therefore, Gianvecchio uses the information entropy and the corrected conditional entropy to describe the shape and regularity of the inter-packet delays. Although two entropy methods could effectively detect most of CTCs, unfortunately they did not detect latest DMCTC.

Shrestha *et al.* [22] proposed a support vector machine framework (SVM test) for reliable detection of covert timing channels. The authors extracted multiple fingerprints (*e.g.* K-S test [17], Entropy test and Corrected Conditional Entropy [7] scores) from the network traffic and used them as features to train the support vector machine. Through the classifier of the SVM after training to distinguish whether the traffic is overt or covert. Although the SVM test could detect most of covert timing channels, it needs a lot of features and time to train the classifier of framework. In the high-speed network environment, the features of covert timing channels would change due to the influence (*e.g.* network jitter, packet loss) of the network environment, and the classifier needs to make corresponding changes. This means that the SVM test would lack certain detection robustness in the high-speed network environment, that is, it is easy to be affected by the network environment.

3 Perceptual Hash Measures

Perceptual hash [6, 15, 23], unlike traditional hash, is also known as robust hash and digital fingerprinting. It is a one-way mapping from the digital representation of multimedia information to the perceptual digest. The robustness, discrimination and reliability of perceptual hash are of great importance in the field of information security and communication. Over the last decades, perceptual hash was originally applied in image recognition and authentication [10, 19, 24], and later applied in multimedia information such as audio and video [14, 20, 26]. However, few of perceptual hash technique is applied to the detection of CTCs in the network system. In this sec-

tion, the detail description of the perceptual hash-based approach designed and the approach used to accomplish the detection of covert traffic are provided.

3.1 System Model

Figure 3 shows the model representation of the detection framework. The model is a essential network monitor that has access to the network traffic which is attempting to detect. It could be designed to simply tap into all legitimate network traffic as shown in figure. The model consists of four primary units-a traffic filter, a perceptual feature extractor, the perceptual hash functions and a perceptual hash matching detector. The traffic filter selects network traffic for the perceptual feature extraction. The perceptual feature extractor derives the perceptual features from the traffic selected by the traffic filter. The extracted features are then transformed into the perceptual hash sequences by the perceptual hash functions. By performing the detection threshold estimation, it is ready for implementation on a high-speed network for detection of CTCs.

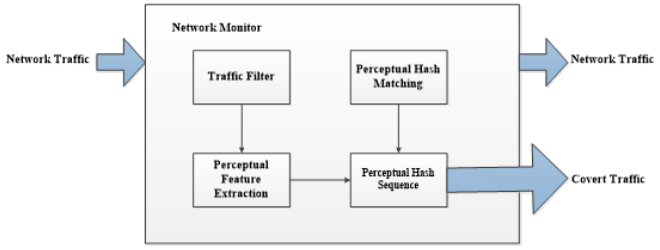


Figure 3: A perceptual hash-based system model for detecting CTCs

3.2 The System Model-based Perceptual Hash Design Process

In this section, we demonstrate the way of designing each unit of the system model. The specific content is as follows:

The perceptual features extraction: First, the network traffic of CTCs is converted into the time series signal $A(t)$, t is the serial number of inter-packet delay. We then extract the perceptual features of $A(t)$ in time-domain and in frequency-domain. The extracted features in frequency-domain are defined as matrix T and the extracted features in time-domain are defined as matrix T_1.

The perceptual hash functions design: H_i=$PH_{gen}(T)$, where PH_{gen} represents a perceptual hash generation function and H_i represents a perceptual hash sequence generated by the matrix T and T_1.

The perceptual matching algorithm design: PD=$PH_{match}(H_i, H_j)$, where PH_{match} is the matching function, H_i and H_j represent the perceptual hash sequences of unknown traffic and sample traffic. PD is the perceptual distance between two sequences that is used to identify the perceptual hash value. It is determined whether H_i and H_j are the same channels by evaluating whether PD is within the detection threshold (β) estimated by the sample traffic. The design process of perceptual hash-approach is shown in Figure 4.

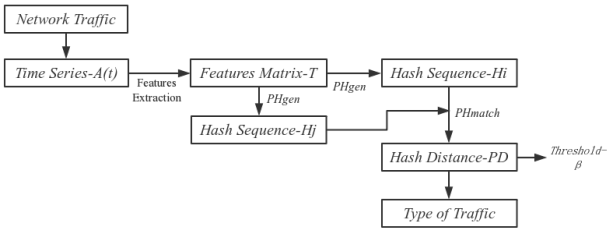


Figure 4: The design of perceptual hash process

3.2.1 The Perceptual Features Extraction Process

Since the law of CTCs in time-domain and the features in frequency-domain are different from the legitimate traffic, we use the discrete wavelet transform (DWT) [18] and the information entropy [4, 13] to extract the perceptual features in frequency-domain and in time-domain, respectively. The extraction process is as follows:

Step 1. DWT analysis: The time series $A(t)$ of CTCs is performed to the global DWT to obtain the high frequency coefficients Hg={Hg_i | i=1, 2, ..., n} and the low frequency coefficients Lh={Lh_j | j=1, 2, ..., m}. Where n and m are the length of high frequency and low frequency coefficients, respectively.

Step 2. Blocking the coefficients: Hg and Lh are divided into non-overlapping range blocks of fixed size. The block length of Hg and Lh are N and M, respectively. The block number of Hg and Lh are S. The obtained matrix T is shown in Equation (1):

$$T = \begin{bmatrix} Lh^1 & Lh^2 & \cdots & Lh^M \\ Lh^{M+1} & Lh^{M+2} & \cdots & Lh^{2\times M} \\ \vdots & \vdots & \vdots & \vdots \\ Lh^{(S-1)\times M+1} & Lh^{(S-1)\times M+2} & \cdots & Lh^{S\times M} \\ Hg^1 & Hg^2 & \cdots & Hg^N \\ Hg^{N+1} & Hg^{N+2} & \cdots & Hg^{2\times N} \\ \vdots & \vdots & \vdots & \vdots \\ Hg^{(S-1)\times N+1} & Hg^{(S-1)\times N+2} & \cdots & Hg^{S\times N} \end{bmatrix} \quad (1)$$

Step 3. The perceptual features extraction in frequency-domain: The standard deviation of each column of the matrix T is calculated, as shown in Equation (2), where v is the mean of each column value of T. The

features parameter vector H_1={$std(k)$ | k=1, 2, ..., N+M} is generated by compressing T.

$$H_1 = \begin{bmatrix} std(1) \\ std(2) \\ \vdots \\ std(N+M) \end{bmatrix}$$

$$std(k) = \sqrt{\frac{1}{S}\sum_{m=1}^{S}(T(m,k)-v)^2} \quad (2)$$

Step 4. The corrective perceptual features extraction: The short-term energy is used to compensate for some frequency domain features due to segmentation loss. The energy value of each row of T is calculated as show in Equation (3). The calculation result is generated as the corrective features parameter vector H_2={$g(k)$ | k=1, 2, ..., S}.

$$H_2 = \begin{bmatrix} g(1) \\ g(2) \\ \vdots \\ g(S) \end{bmatrix}$$

$$g(k) = 10\log\sum_{m=1}^{N+M} T(k,m). \quad (3)$$

Step 5. The perceptual features extraction in time-domain: In order to extract the features in time-domain, the time series $A(t)$ is reasonably segmented. The segment length is C and the number of segments is D. The obtained matrix T_1 is shown in Equation (4).

$$T_1 = \begin{bmatrix} X_1^1 & X_1^2 & \cdots & X_1^C \\ X_2^1 & X_2^2 & \cdots & X_2^C \\ \vdots & \vdots & \vdots & \vdots \\ X_D^1 & X_D^2 & \cdots & X_D^C \end{bmatrix} \quad (4)$$

We calculate the information entropy value G of each row of matrix T_1 and obtain the perceptual features in time-domain. The obtained features parameter vector H_3 is shown in Equation (5).

$$H_3 = \begin{bmatrix} G_1 & G_2 & \cdots & G_D \end{bmatrix} \quad (5)$$

3.2.2 The Perceptual Hash Functions Design

Through the features extraction, we obtain three parameter vectors H_1, H_2 and H_3 respectively. The binary perceptual hash construction is performed on H_1 and H_3, the designed perceptual hash function in Equation (6) is shown. A perceptual sequence $ph_1(k)$={$ph_1(k)$ | k=1, 2, ..., N+M} in frequency-domain and a sequence $ph_3(k)$={$ph_3(k)$ | k=1, 2, ..., D} in time-domain are generated. The three binary hash construction is performed on H_2, The designed perceptual hash function is shown in Equation (7). A perceptual sequence $ph_2(k)$={$ph_2(k)$ | k=2, 3, ..., S-1} corresponding to the corrective features is generated.

$$b_i = \begin{cases} 0 & \text{if } H(k) \le H_{Mean} \\ 1 & \text{otherwise} \end{cases} \quad (6)$$

$$ph(k) = \begin{cases} 1 & \text{if } H(k)^2 - H(k-1)\times H(k+1) > 0 \\ 0 & \text{else if } H(k) - H(k-1) > 0 \\ -1 & \text{otherwise} \end{cases} \quad (7)$$

3.2.3 The Perceptual Hash Matching Algorithm Design

In this paper, the uniformly Hamming distance $D(:,:)$ is used as the method of calculating PD, which is the Bit Error Rate (BER). The calculation formula is as shown in Equation (8), where {i | i=*Normal*, *IPCTC*, *TRCTC*, *JitterBug*, *DMCTC*}, A is the time series of unknown traffic.

$$BER = D(ph(A), ph(A_i)) = \frac{\sum_{j=1}^{3N}|ph_A(j) - ph_{A_i}(j)|}{3N} \quad (8)$$

The perceptual hash sequences of unknown traffic, legitimate traffic, IPCTC, TRCTC, DMCTC and JitterBug sample traffic are generated by the perceptual hash functions. The results are represented as ph, ph_N, ph_{IP}, ph_{TR}, ph_J, ph_{DM}, respectively. The BER of ph to ph_N, ph_{IP}, ph_{TR}, ph_J, ph_{DM} is calculated according to Equation (8). The results are expressed as BER_N, BER_{IP}, BER_{TR}, BER_J, BER_{DM}, respectively. The detection threshold is estimated as β by calculating BER between the homogeneous channels, where {β | β_N, β_{IP}, β_{TR}, β_{DM}, β_J}. Thus, the matching and detection process of unknown traffic is as in Algorithm 1.

Algorithm 1 The matching and detection process

1: Begin
2: **if** $BER_N < \beta_N$ and $\max\{BER_{IP}, BER_{TR}, BER_J, BER_{DM}\} < \min\{\beta_{IP}, \beta_{TR}, \beta_{DM}, \beta_J\}$ **then**
3: $ph \in ph_N$.
4: **else if** $BER_{IP} < \beta_{IP}$ and $\max\{BER_N, BER_{TR}, BER_J, BER_{DM}\} < \min\{\beta_N, \beta_{TR}, \beta_{DM}, \beta_J\}$ **then**
5: $ph \in ph_{IP}$.
6: **else if** $BER_{TR} < \beta_{TR}$ and $\max\{BER_{IP}, BER_N, BER_J, BER_{DM}\} < \min\{\beta_{IP}, \beta_N, \beta_{DM}, \beta_J\}$ **then**
7: $ph \in ph_{TR}$
8: **else if** $BER_J < \beta_{DM}$ and $\max\{BER_{IP}, BER_{TR}, BER_N, BER_{DM}\} < \min\{\beta_{IP}, \beta_{TR}, \beta_N, \beta_J\}$ **then**
9: $ph \in ph_J$
10: **else if** $BER_{DM} < \beta_J$ and $\max\{BER_{IP}, BER_{TR}, BER_J, BER_N\} < \min\{\beta_{IP}, \beta_{TR}, \beta_{DM}, \beta_N\}$ **then**
11: $ph \in ph_{DM}$
12: **end if**
13: If the matching result does not meet the above conditions, we could determine that the unknown traffic is unrecognizable.
14: End

4 Experimental Evaluation

In this section, we verity whether the perceptual hash-based approach is valid through a series of experiments. The perceptual hash-based approach is tested against four typical CTCs: IPCTC [3], TRCTC [1], DMCTC [11], JitterBug [21]. Furthermore, we compare the perceptual hash-based approach (PER-H-test) with three detection tests: the information entropy test (EN-test) [7], the corrective conditional entropy test (CCE-test) [7] and the SVM test (SVM-test) [22].

More specifically, we evaluate the discrimination between unknown traffic and legitimate traffic, four CTCs traffic by calculating *BER*, where $\{BER \mid BER_N, BER_{IP}, BER_{TR}, BER_{DM}, BER_J\}$. The similarity between unknown traffic and legitimate traffic, four CTCs traffic is evaluated by calculating 100-*BER*. The evaluation results of similarity and discrimination will be determined weather the unknown traffic is accurately detected by the true-positive rates. The true-positive rates represent the ratio at which covert timing channels are accurately detected.

4.1 Experimental Results

In the following, we show our experimental results in detail. Four typical CTCs are IPCTC, TRCTC, DMCTC and JitterBug. The experiments are organized by detecting the difficulty of covert timing channels.

4.1.1 IPCTC

Our first set of experiments is test on IPCTC. IPCTC is the simplest and easily detectable covert channel among four CTCs, because its perceptual features exhibit abnormality in both time-domain and frequency-domain. The abnormality features in frequency-domain of IPCTC are caused by the encoding way. If the time series of IPCTC is random, then we treat the time series as a series of Bernoulli trials. Therefore, the inter-packet delays of the time series are approximate to the Geometric distribution. The abnormality features in time-domain is due to the lack of obvious correlations between the inter-packet delays of IPCTC. This means, the inter-packet delays are determined by the covert information being encoded, not by the foregoing packet delays.

We conducted 100 times the detection test for samples of unknown traffic, legitimate traffic and four CTCs traffic. The similarity (100-*BER*) estimation results between the perceptual sequence of unknown traffic (IPCTC) and IPCTC, TRCTC, DMCTC, JitterBug and the legitimate sample traffic are shown in Figure 5, where $\{BER \mid BER_N, BER_{IP}, BER_{TR}, BER_{DM}, BER_J\}$. We could obverse that the similarity between the unknown traffic (IPCTC) and the perceptual sequence of IPCTC is much higher than that of other traffic samples. This means that the generated perceptual sequences of IPCTC samples could be used to accurately distinguish the traffic of IPCTC from the legitimate traffic and three other CTCs. In addition, the reason why 100-*BER* of legitimate traffic and three other CTCs traffic dose not change is that the encoding method and regularity of IPCTC are quite different from that of four other channels. The distinguishing matrix for detection IPCTC is shown in Table 1. It is seen that the identification accuracy of IPCTC and the legitimate traffic was 100 percent, respectively, when working with traffic sizes of 2,000 and 500 samples. At the same time, the experimental results on IPCTC verify the good discrimination of the perceptual hash-based approach.

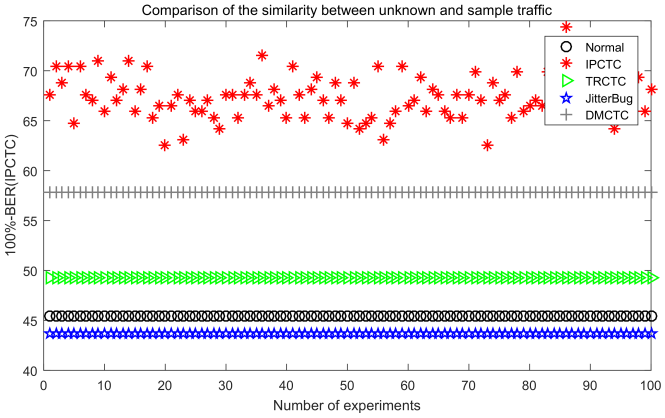


Figure 5: The similarity estimation between IPCTC and other traffic samples

Table 1: The true-positive rates of IPCTC

	Sample Size=2000		Sample Size=500	
	Overt	Covert	Overt	Covert
Overt	100	0	98	2
Covert	0	100	0	100

In order to verify the robustness and practicality of the proposed approach, we run each 100 times test on IPCTC with 10 percent noise (network jitter and 10 percent packet loss) and 30 percent noise (network jitter and 30 percent packet loss) respectively. The robustness estimation results are shown in Figure 6. We could see from Figure 5 that *BER* of DMCTC (45) is closest to IPCTC except IPCTC itself, thus 45 is as the detection threshold β_{IP} of IPCTC. For noiseless and 10 percent noise IPCTC, all *BER* values below 45 mean that the noiseless and 10 percent noise IPCTC can be 100 percent identified. But for the 30 percent noise IPCTC, there is a part of BER value exceeds 45. This means that the 30 percent noise IPCTC could be recognized around 90 percent. In summary, the experimental results verify that the proposed approach for detecting IPCTC has a good detection robustness in a complex network environment.

4.1.2 TRCTC

The second set of experiments is investigated how the proposed approach performs against TRCTC. TRCTC is

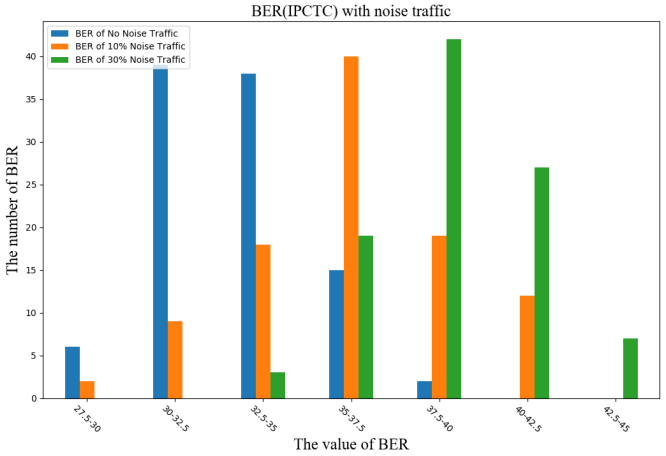


Figure 6: The robustness estimation of the proposed test for detecting IPCTC

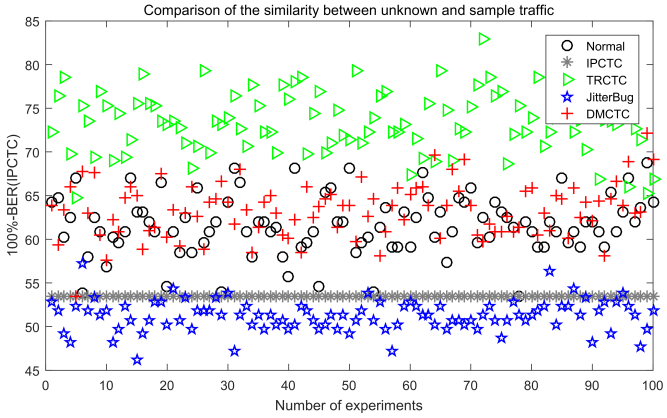


Figure 7: The similarity estimation between TRCTC and other traffic samples

a more advance CTCs, which approximates the behavior of legitimate traffic by replaying a set of legitimate inter-packet delays. Thus, TRCTC has the comparable perceptual features in frequency-domain as legitimate traffic, but the features in time-domain exhibit abnormal. The regularity of TRCTC, like IPCTC, is due to the lack of correlation between inter-packet delays, and the replayed delays still falsifies the regularity of the original process.

We conducted 100 times the detection test for samples of unknown traffic, legitimate traffic and four CTCs traffic. The similarity (100-BER) estimation results between the perceptual sequence of unknown traffic (TRCTC) and IPCTC, TRCTC, DMCTC, JitterBug and the legitimate sample traffic are shown in Figure 7, where $\{BER \mid BER_N, BER_{IP}, BER_{TR}, BER_{DM}\}$. Due to TRCTC traffic is similar to the legitimate traffic, the BER values of some points are below the legitimate traffic samples in the figure. However, the similarity between unknown traffic (TRCTC) and the perceptual sequence of TRCTC is generally higher than that of other traffic samples. We thus could use the generated perceptual sequence of TRCTC samples to distinguish the traffic of TRCTC well from the legitimate traffic and three other CTCs. In addition, the reason why BER of DMCTC traffic is close to the legitimate traffic is that DMCTC has the comparable perceptual features in frequency-domain as legitimate traffic. The distinguishing matrix for detection TRCTC is shown in Table 2. For a traffic size of 500 samples, the proposed approach was able to detect TRCTC with 100 percent. For a traffic size of 2000 samples, the approach was able to detect TRCTC with 81 percent.

We run each 100 times test on TRCTC with 10 percent noise (network jitter and 10 percent packet loss) and 30 percent noise (network jitter and 30 percent packet loss) respectively. The robustness estimation results are shown in Figure 8. We could see from Figure 7 that BER of legitimate traffic (37) is closest to TRCTC except TRCTC itself, thus 37 is as the detection threshold β_{TR} of TRCTC. For the noiseless TRCTC, the values of around 85 percent BER below 37 mean that our approach is able to detect TRCTC with around 85 percent. For the 10 percent and 30 percent noise TRCTC, there are values of around 75 percent BER and around 60 percent BER below 37, respectively. Thus, the proposed approach still has good robustness for detecting TRCTC in the poor network environment.

Table 2: The true-positive rates of TRCTC

	Sample Size=2000		Sample Size=500	
	Overt	Covert	Overt	Covert
Overt	98	2	83	17
Covert	0	100	19	81

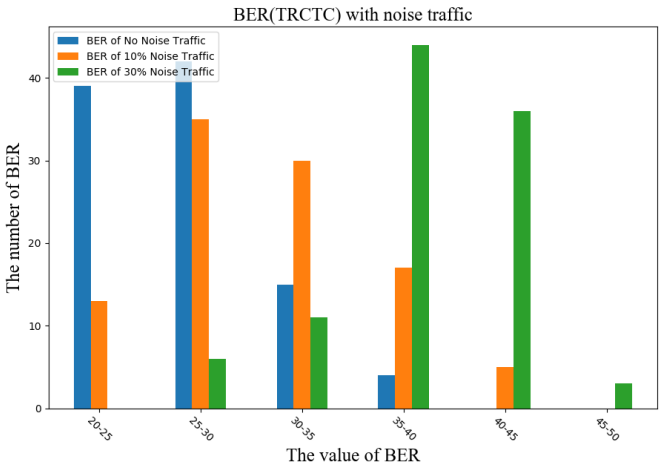


Figure 8: The robustness estimation of the proposed test for detecting TRCTC

4.1.3 DMCTC

Our third set of experiments tested for detecting DMCTC. DMCTC is a more advanced covert timing channel that matches the traffic distribution to imitate the legitimate traffic. By collecting the inter-packet delays of legitimate traffic, DMCTC fits the distribution and replays the inter-packet delays of legitimate traffic to confuse the

detector. Thus, DMCTC has the similar perceptual features to the legitimate traffic in frequency-domain, due to the distribution, and the similar perceptual features in time-domain, due to the packet replay.

We conducted 100 times the detection test for samples of unknown traffic, legitimate traffic and four CTCs traffic. The similarity (100-*BER*) estimation results between the perceptual sequence of unknown traffic (DMCTC) and IPCTC, TRCTC, DMCTC, JitterBug and the legitimate sample traffic are shown in Figure 9, where $\{BER \mid BER_N, BER_{IP}, BER_{TR}, BER_{DM}\}$. We could see it that the similarity between unknown traffic (DMCTC) and the perceptual sequence of DMCTC is generally higher than that of other traffic samples. Since the shape of DMCTC, like TRCTC, simulates the features of legitimate traffic in frequency-domain, the *BER* values of some points are below the legitimate traffic samples in the figure. But the generated perceptual sequence of DMCTC samples could still be used to distinguish the traffic of DMCTC well from the legitimate traffic and three other CTCs. The distinguishing matrix for detection DMCTC is shown in Table 3. For a traffic size of 500 samples, the proposed approach is able to detect DMCTC with 91 percent. For a traffic size of 2000 samples, the approach is able to detect DMCTC with 82 percent.

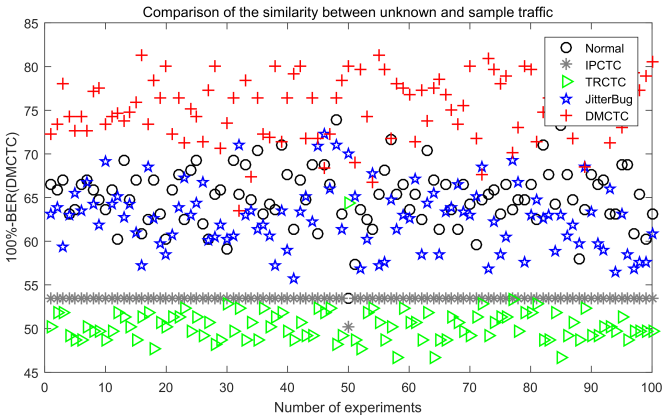


Figure 9: The similarity estimation between DMCTC and other traffic samples

Table 3: The true-positive rates of DMCTC

	Sample Size=2000		Sample Size=500	
	Overt	Covert	Overt	Covert
Overt	79	21	88	12
Covert	18	82	09	91

We run each 100 times test on DMCTC with 10 percent noise (network jitter and 10 percent packet loss) and 30 percent noise (network jitter and 30 percent packet loss) respectively. The robustness estimation results are shown in Figure 10. We could see from Figure 9 that *BER* of legitimate traffic (28) is closest to DMCTC except DMCTC itself, thus 28 is as the detection threshold β_{DM} of DMCTC. For the noiseless DMCTC, the values of 91 percent *BER* below 28 mean that our approach is able to detect DMCTC with 91 percent. For the 10 percent and 30 percent noise DMCTC, there are values of around 70 percent *BER* and 50 percent *BER* below 37, respectively. Since DMCTC is extremely difficult to detect, in the 30 percent noise of cases, there is still a true-positive rate of around 50 percent. Thus, the proposed approach has good detection robustness for detecting DMCTC in the poor network environment.

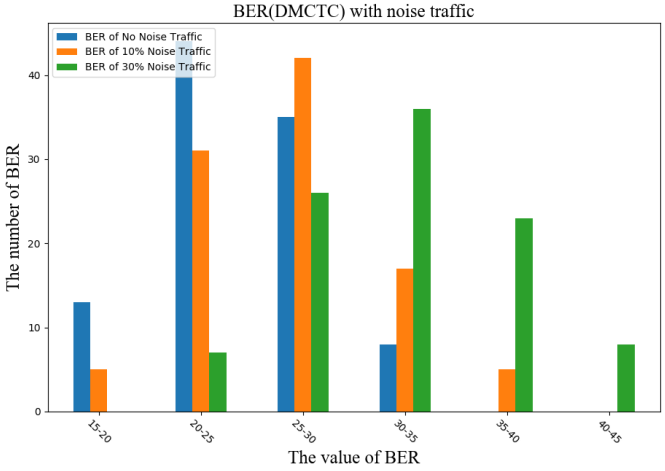


Figure 10: The robustness estimation of the proposed test for detecting DMCTC

4.1.4 JitterBug

The fourth set of experiments is investigated how the proposed approach performs against JitterBug. JitterBug is a passive CTCs, which dose not generate an additional traffic to transmit the covert information. Thus, due to having legitimate traffic as the base and only slightly adding the packet delay, JitterBug preserves partial correlation of the original process. Therefore, JitterBug has similar features in time-domain and in frequency-domain to legitimate traffic. Based on the above reasons, JitterBug is very difficult to detect. Considering that adding the network jitter may affect the channel capacity, we choose the value of w as 15ms to construct JitterBug.

We conducted 100 times the detection test for samples of unknown traffic, legitimate traffic and four CTCs traffic. The similarity (100-*BER*) estimation results between the perceptual sequence of unknown traffic (JitterBug) and IPCTC, TRCTC, DMCTC, JitterBug and the legitimate sample traffic are shown in Figure 11, where $\{BER \mid BER_N, BER_{IP}, BER_{TR}, BER_{DM}\}$. We could see it that the similarity between unknown traffic (JitterBug) and the perceptual sequence of JitterBug is generally higher than that of other traffic samples. Since the shape and regularity of JitterBug perceptual features, is based on the features of legitimate traffic in frequency-domain and time-domain, the *BER* values of some points are below the legitimate traffic samples in the figure. However, the generated perceptual sequence of JitterBug samples could

be used to distinguish the traffic of JitterBug well from the legitimate traffic and three other CTCs. The distinguishing matrix for detection JitterBug is shown in Table 4. For a traffic size of 500 samples, the proposed approach is able to detect JitterBug with 83 percent. For a traffic size of 2000 samples, the approach is able to detect JitterBug with 100 percent.

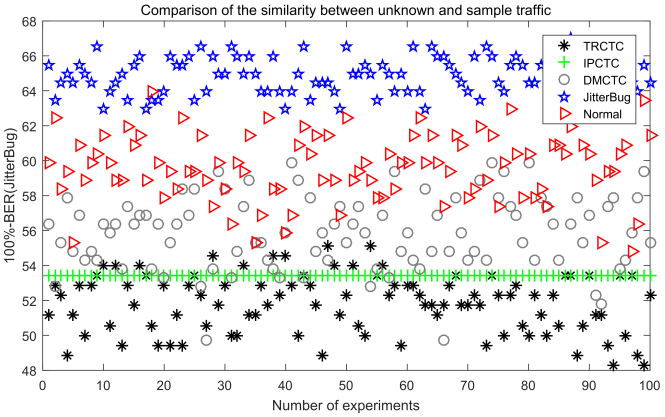


Figure 11: The similarity estimation between JitterBug and other traffic samples

Table 4: The true-positive rates of JitterBug

	Sample Size=2000		Sample Size=500	
	Overt	Covert	Overt	Covert
Overt	97	3	80	20
Covert	0	100	17	83

The 100 times test is performed on detecting JitterBug with 10 percent noise (network jitter and 10 percent packet loss) and 30 percent noise (network jitter and 30 percent packet loss) respectively. The robustness estimation results are shown in Figure 12. We could see from Figure 11 that the *BER* value of legitimate traffic (39) is closest to JitterBug except JitterBug itself, thus 39 is as the detection threshold β_J of JitterBug. For the noiseless JitterBug, the values of around 90 percent *BER* below 39 mean that our approach is able to detect JitterBug with around 90 percent. For the 10 percent and 30 percent noise JitterBug, there are values of around 75 percent *BER* and around 65 percent *BER* below 39, respectively. Since JitterBug is extremely difficult to detect, in 30 percent noise of cases, there is still a detection rate of around 65 percent. Thus, the proposed approach has good robustness for detecting JitterBug in the poor network environment.

4.2 Four Covert Timing Channels Traffic-Variable Sample Size

The last set of experiments is performed to investigate how the perceptual hash-based approach detects with different sample sizes against all four CTCs IPCTC,

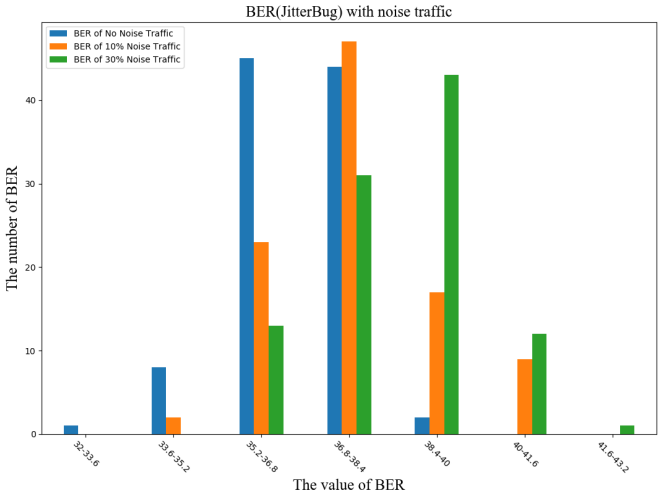


Figure 12: The robustness estimation of the proposed test for detecting JitterBug

TRCTC, DMCTC and JitterBug. The sample size is essential because it is determined the amount of time the detection tests take to detect a covert timing channel. This means that we need to detect CTCs before they transmit as little covert information as possible. Of course, if we detect CTCs with the smallest possible sample size, they will transmit less covert information prior to detection. In general, although the smaller sample size means faster detection, it is often less accurate than the larger sample size. Therefore, we need a compromise between a small sample size and the accuracy of the detection. In this section, we vary sample sizes from 500 to 2250 inter-packet delays for the perceptual hash-based approach (PER-H test), the information entropy approach (EN test) [7], the corrected conditional entropy approach (CCE test) [7] and the support vector machine approach (SVM test) [22].

The true-positive rates for PER-H test, EN test, CCE test and SVM test against IPCTC, TRCTC, DMCTC and Jitterbug with 500 to 2250 inter-packet delays are shown in Figure 13. For these four detection tests, the true-positive rates decreases with the decrease of sample size in different covert timing channels. For IPCTC (Figure 13(a)), there is no decrease in true-positive rates expect CCE test. The regularity of IPCTC is obvious, thus the detection tests doesn't need a large amount of data to detect IPCTC expect CCE test. For TRCTC (Figure 13(b)), En test is unable to detect TRCTC, thus its true-positive rate remain close to 0. In addition, the true-positive rates of the other three approaches all decrease at different rates with the decrease of the sample size, and the declining trend of our approach (PER-H test) is more gentle. Therefore, in the case of small sample size, PER-H test is superior to the other three approaches on detecting TRCTC. For DMCTC (Figure 13(c)), the DMCTC true-positive rates of the approaches mentioned in this section except PER-H test degrade at different rates with the decrease of sample. The DMCTC true-positive rates of PER-H test elevate with the decrease of sample size. Thus, on detecting DMCTC, PER-H test is superior

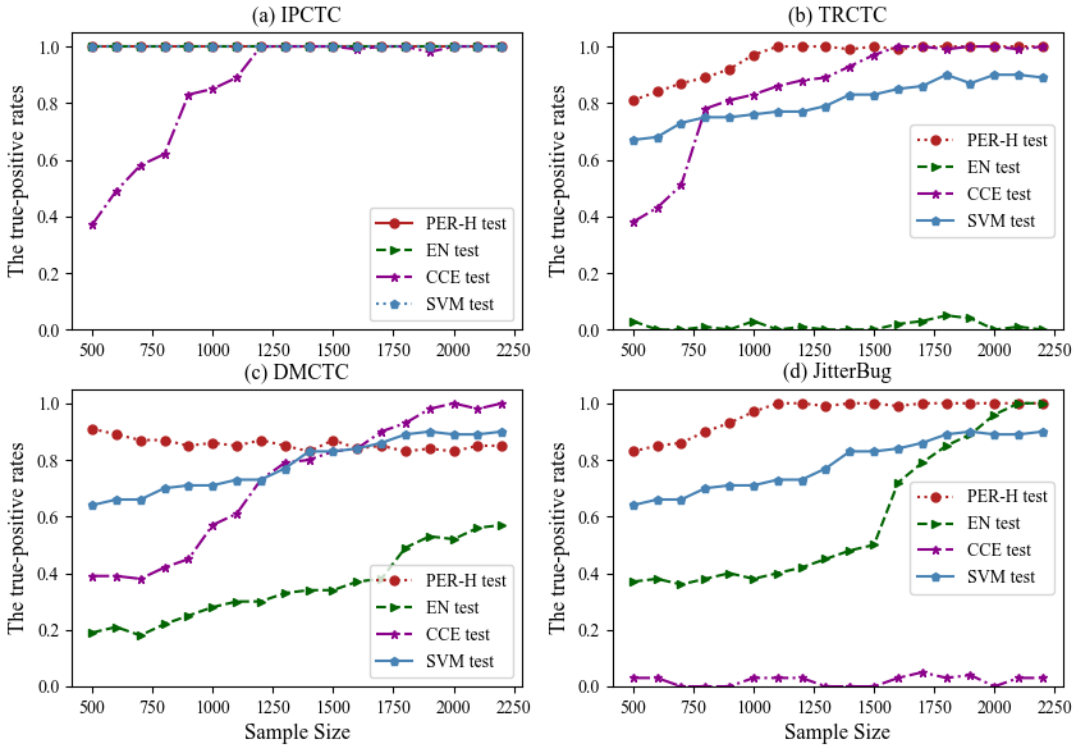


Figure 13: The true-positive rates of all channels-variable sample size

to three other approaches in the case of small sample size and is inferior to three other approaches in the case of large sample size.

Compare with IPCTC, TRCTC and DMCTC, JitterBug (Figure 13(d)) is relatively difficult to detect. It adds only small delays and does not affect the channel capacity. Therefore, it is difficult to distinguish JitterBug from the inter-packet delays of legitimate. CCE test is unable to detect JitterBug, thus its true-positive rate remain close to 0. In the case of small sample size, the true-positive rates of EN test and SVM test is lower than PER-H test. This is because the proposed approach combines the perceptual features of time-domain and frequency-domain and uses the analysis of perceptual hash to find out the weak discrimination between JitterBug and the legitimate traffic. Therefore, PER-H test is superior to EN test and SVM test on detecting JitterBug in both the case of small and large sample size.

In general, we could observe that IPCTC and TRCTC are easier to be detected than DMCTC and JitterBug. The approach in this paper is able to accurately detect IPCTC, TRCTC and JitterBug at the true-positive rates of 1.0 with a small sample size. DMCTC is much more difficult to detect than three other covert timing channels, the season is attributed to the fact that DMCTC not only imitates the regularity and shape of legitimate traffic, but also has a striking similarity with TRCTC and JitterBug. Although the proposed approach is unable to detect DMCTC at the true-positive rate of 1.0, it is superior to EN test, CCE test and SVM test in the case of small sample size.

4.3 Discussion

Our approach is able to detect multiple covert timing channels under certain conditions. For the previous detection approaches, they are unable to detect most of the tested covert timing channels in the high-speed network environment. The main reason is that the detection approaches lose the robustness of detection. For example, the information entropy and corrected conditional entropy approach are incapable of recognizing most covert timing channels in a poor network environment (including network packet loss, jitter, packet injection, *etc.*).

Another reason is that the detection methods need to adapt to the high-speed network environment. This means that the detection methods need detect them under the premise that covert timing channels transmit as little covert information as possible. However, the previous detection approaches are not ideal in the case of small sample size, like the approaches mentioned in Section 4.2.

Our method is proved to be more efficient than the previous detection methods. The proposed approach uses the robustness of perceptual hash to solve the problem that the detection robustness is loss in the case of network interference. At the same time, our method combines the perceptual features in time-domain and in frequency-domain, and solves the problem that the true-positive rates is low in the case of small sample size by the discrimination of the perceptual hash. Thus, the perceptual hash-based approach has effectively passed the test of detection covert timing channels.

5 Conclusion and Future Works

Existing detection methods could detect most of covert timing channels, while lack certain detection robustness. At the same time, the true-positive rates of existing methods are not ideal in the case of small sample size. For the above shortcomings of methods, we proposed a perceptual hash-based approach to detect covert timing channels. The proposed approach improves the discrimination between CTCs traffic and the legitimate traffic by combining the perceptual features in time-domain and in frequency-domain of CTCs. This makes the true-positive rates of proposed approach on detecting CTCs is superior to others in the case of small sample size. In the meantime, the proposed approach utilizes the robustness of perceptual hash to preserve the main of features. This means that our approach could still identify covert timing channels well in the case of network interference (network jitter, data packet loss *e.g.*). Experimental results confirm that the proposed approach not only has a good true-positive rate in the case of small sample size, but also has good detection robustness in the case of network interference.

In future, we plan to optimize the extracted perceptual features and improve the designed perceptual hash function. In addition, we plan to utilize the perceptual hash-based approach against the new covert timing channels in the future.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (NO.61370007), the Program for New Century Excellent Talents of Fujian Provincial (NO.2014FJ-NCET-ZR06), and the Subsidized Project for Postgraduates' Innovative Fund in Scientific Research of Huaqiao University (No.17014083018).

References

[1] S. Cabuk, Network Covert Channels: Design, Analysis, Detection, and Elimination, 2006. (`https://core.ac.uk/download/pdf/21173179.pdf`)

[2] S. Cabuk, C. E. Brodley, and C. Shields, "Ip covert channel detection," *Acm Transactions on Information & System Security*, vol. 12, no. 4, pp. 1–29, 2009.

[3] S. Cabuk, C. E. Brodley, and C. Shields, "IP covert timing channels: Design and detection," in *Acm Conference on Computer & Communications Security*, pp. 178–187, 2004.

[4] Y. Chen, N. Zhang, H. Tian, T. Wang, and Y. Cai, "A novel connection correlation scheme based on threshold secret sharing," *IEEE Communications Letters*, vol. 20, no. 12, pp. 2414–2417, 2016.

[5] Y. Chen, X. Ma, and X. Wu, "DDoS detection algorithm based on preprocessing network traffic predicted method and chaos theory," *IEEE Communications Letters*, vol. 17, no. 5, pp. 1052–1054, 2013.

[6] L. Chen, Z. Li, and J. F. Yang, "Compressive perceptual hashing tracking," *Neurocomputing*, vol. 239, pp. 69–80, 2017.

[7] S. Gianvecchio and H. Wang, "An entropy-based approach to detecting covert timing channels," *IEEE Transactions on Dependable & Secure Computing*, vol. 8, no. 6, pp. 785–797, 2011.

[8] S. Gianvecchio, H. Wang, D. Wijesekera, and S. Jajodia, "Model-based covert timing channels: Automated modeling and evasion," in *International Symposium on Recent Advances in Intrusion Detection*, pp. 211–230, 2008.

[9] C. G. Girling, "Covert channels in Lan's," *IEEE Transactions on Software Engineering*, vol. SE-13, no. 2, pp. 292–296, 1987.

[10] J. Ji, X. Lü, L. Han, and C. Zhang, "Fast and adaptive region merging based on perceptual hashing via multi-thresholding for SAR image segmentation," *Remote Sensing Letters*, vol. 7, no. 12, pp. 1199–1208, 2016.

[11] G. Liu, J. Zhai, and Y. Dai, "Network covert timing channel with distribution matching," *Telecommunication Systems*, vol. 49, no. 2, pp. 199–205, 2012.

[12] X. Luo, E. W. W Chan, and R. K. C Chang, "TCP covert timing channels: Design and detection," in *IEEE International Conference on Dependable Systems & Networks with Ftcs & Dcc*, 2008. DOI: 10.1109/DSN.2008.4630112.

[13] X. Ma and Y. Chen, "DDoS detection method based on chaos analysis of network traffic entropy," *IEEE Communications Letters*, vol. 18, no. 1, pp. 114–117, 2014.

[14] R. D. Major, *Pre-Distribution Identification of Broadcast Television Content using Audio Fingerprints*, US20180359540A1, 2014. (`https://patents.google.com/patent/US20180359540A1/en`)

[15] A. Neelima and K. M. Singh, "Perceptual hash function based on scale-invariant feature transform and singular value decomposition," *Computer Journal*, vol. 59, no. 9, pp. 1275–1281, 2016.

[16] 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 & Information Engineering*, vol. 3, no. 1, pp. 10–18, 2015.

[17] P. Peng, P. Ning, and D. S. Reeves, "On the secrecy of timing-based active watermarking trace-back techniques," in *IEEE Symposium on Security & Privacy*, 2006. (`http://discovery.csc.ncsu.edu/pubs/Oakland06.pdf`)

[18] J. Qin, R. Sun, X. Xiang, H. Li, and H. Huang, "Anti-fake digital watermarking algorithm based on QR codes and DWT," *International Journal of Network Security*, vol. 18, no. 6, pp. 1102–1108, 2016.

[19] S. B. Qiao, Q. Y. Zhang, T. Zhang and D. F. Wu, "Spectrogram-based efficient perceptual hashing scheme for speech identification," *International Journal of Network Security*, vol. 21, no. 2, pp. 259–268, 2019.

[20] N. Saikia, and P. K. Bora, "Perceptual hash function for scalable video," *International Journal of Information Security*, vol. 13, no. 1, pp. 81–93, 2014.

[21] G. Shah, A. Molina, and M. Blaze, "Keyboards and covert channels," in *Conference on Usenix Security Symposium*, vol. 15, no. 5, 2006.

[22] P. Shrestha, M. Hempel, F. Rezaei, and H. Sharif, "A support vector machine-based framework for detection of covert timing channels," *IEEE Transactions on Dependable & Secure Computing*, vol. 13, no. 2, pp. 1–1, 2016.

[23] X. Wang, K. Pang, X. Zhou, Z. Yang, L. Lu, and J. Xue, "A visual model-based perceptual image hash for content authentication" *IEEE Transactions on Information Forensics & Security*, vol. 10, no. 7, pp. 1336–1349, 2015.

[24] F. Wen, H. M. Hu, Z. Hu, S. Liao, and L. Bo, "Perceptual hash-based feature description for person re-identification," *Neurocomputing*, vol. 272, pp. 520–531, 2017.

[25] X. Wu and Y. Chen, "Validation of chaos hypothesis in nada and improved ddos detection algorithm," *Communications Letters IEEE*, vol. 17, no. 12, pp. 2396–2399, 2013.

[26] G. Yang, X. Chen, and D. Yang, "Efficient music identification by utilizing space-saving audio fingerprinting system," in *IEEE International Conference on Multimedia & Expo*, 2014. https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6890236

[27] S. Zander, G. Armitage, and P. Branch, "Stealthier inter-packet timing covert channels," in *International Ifip Tc 6 Conference on Networking*, pp. 458–470, 2011.

Biography

Linfan Wang was born in Shanxi, China in 1995. He received the B.S. Degree from Hubei University of Mediciney, Hubei, China in 2017. He is currently pursuing the M.S. Degree in Huaqiao University. His research interests include Covert Channels Detection and Perceptual Hash and Blockchain and Application.

Yonghong Chen received the Ph.D. degree from Chongqing University, Chongqing, China, in 2005. He is a Professor in Huaqiao University of China. His current interests include Network and Information Security, Network intrusion detection, Digital Watermarking and Property Protection and Blockchain and Application.

An Efficient Mobile Location-based Security Service Framework for Resource-constrained Devices

Yinghui Zhang[1,2], Xinwei Ma[1], Axin Wu[3], Fangyuan Ren[1], and Dong Zheng[1,2]
(Corresponding author: Yinghui Zhang)
School of Cyberspace Security, Xi'an University of Posts and Telecommunications[1]
Xi'an 710121, China
Westone Cryptologic Research Center, Beijing 100070, China[2]
College of Cybersecurity, Jinan University, Guangzhou 510632, China[3]
(Email: yhzhaang@163.com)
(Received June 5, 2019; Revised and Accepted Sept. 4, 2019; First Online Sept. 21, 2019)

Abstract

In this paper, we propose a mobile location-based security service framework supporting Cloud-side Matching and Preliminary Decryption, which is called CMPD. A user only needs to send part of his attribute key to the cloud server and doesn't need any computation in the matching phase. Then the cloud service provider (CSP) tests whether the attributes of a user meet the corresponding access policy. After passing the matching phase, the user needs to send his attribute key to the CSP for decryption. The CSP sends the result of the computation to the user after having completed the preliminary decryption. The user can get the plaintext through a few decryption operations. CMPD neither leakages the attributes and geographic coordinate of the user, nor reveals the information of access policies. Besides, CMPD achieves the confidentiality of the mobile location-based service data. Finally, security and performance analysis show that our proposed framework has achieved privacy protection, data security and efficiency improvement for resource-constrained mobile devices.

Keywords: Attribute-based Encryption; Cloud Computing; Fast Decryption; Mobile Location-based Service; Privacy Protection

1 Introduction

Thanks to the progress of cloud storage and computing, numerous cloud service providers have emerged, i.e., Huawei Cloud, Microsoft Azure, Google Cloud, Amazon Web Services and Tencent Cloud. They can provide cloud services to the government, companies, and individuals. Cloud service users can send their data files to the cloud server.

The CSP can not only store the data files from users but also compute and process it. Mobile location-based service (MLBS) has been further developed under the technology of cloud storage and computing. MLBS includes two phases. First, the MLBS provider receives geographic coordinate from users. Mobile devices can get geographic coordinate through GPS, mobile networks and WIFI. Second, the MLBS provider will provide kinds of information services according to the geographic coordinate.

For example, finding friends around your location on the chat software, finding nearby restaurants and hotels, checking in at a meeting or conference. However, more and more people are concerned about cloud security [1] and care about security and privacy of the data. When MLBS data files are outsourced to the cloud server, the MLBS provider hopes that no one can access the files except the authorized users. Besides, flexible access control and fine-grained framework are required because MLBS is a model [9] which provides data services. Some location-based service frameworks [5, 14] employ attribute-based encryption to solve the above problems.

Anonymous ciphertext-policy attribute-based encryption (CP-ABE) was proposed in [10]. Data owners can formulate an access policy and encrypt the plaintext according to the access policy using the encryption algorithm. Access policy is encrypted in the ciphertext. A person can obtain the plaintext from the encrypted ciphertext when the attributes of this person meet the access policy and this person cannot guess what the access policy is. However, it is very complex for the resource-constraint mobile devices, because heavy computation is needed during the decryption in CP-ABE algorithm. In existing ABE algorithms, a user can only know whether he or she has permission to access data after fully decrypting the ciphertext. Hence, the existing ABE algorithms

is a bit inefficient, because even when a user does not have an access right, he or she can know the fact after decrypting the ciphertext. However, most mobile devices are resource-constraint, the above ABE algorithms are not suitable for these devices.

Recently, Zhang *et al.* [27] proposed a scheme of anonymous CP-ABE, which includes a matching phase and a decryption phase. The proposed scheme is an improved one of the preliminary version [26]. The user's attributes first need to satisfy access policy in the match phase and then the user can perform the decryption phase. However, these two phases are performed on the user side. To further reduce the computing cost of resource-constraint mobile devices and improve the efficiency of MLBS, we modify and innovate the basic scheme of Zhang *et al.* [27] and then apply it to MLBS. In our scheme, the match and preliminary decryption are operated and computed by the CSP. Because of the application of MLBS, we add a geographic coordinate match in the match phase. Users can complete decryption with only a few computation operations.

In this paper, we propose a mobile location-based security service framework supporting Cloud-side Matching and Preliminary Decryption, which is called CMPD. The following are our contributions to this paper.

In our framework, which is shown in Figure 1, the user only needs to send part of his attribute key which is used to match and geographic coordinate to the cloud server and doesn't need any computation in the match phase. Then the cloud service provider tests whether the attributes of a certain person meet the corresponding access policy. After passing the match phase, the user needs to send an attribute key which is used to decrypt to the cloud server. Then the CSP sends the result of the computation to the user after having completed the preliminary decryption. The user can get the plaintext through a few decryption operations.

To make our CP-ABE more suitable for MLBS, we add a geographic coordinate attribute to the algorithm, which is used to enable users to obtain corresponding MLBS. The proposed framework neither leakages the attributes and geographic coordinate of the user, nor reveals the information of access policy. Besides, the proposed framework achieves the confidentiality of the mobile location-based service data. We guarantee data security, user privacy security and access policy security. At the same time, our proposed framework improves the efficiency of the scheme and reduces the computation cost for the resource-constraint mobile devices.

Now, we introduce the structure of the rest of our paper. We describe the related work in the next section. We describe the system model, security model and design goals in Section 3. Some preliminaries are given in Section 4. We introduce our CP-ABE scheme in Section 5. We introduce our CMPD framework in Section 6. Finally, we analyze the security and performance of our CMPD framework in Sections 7 and 8 respectively.

2 Related Work

Since Sahai *et al.* [20] proposed the ABE scheme, many types of research have been studied on the schemes of various ABE. This cryptographic algorithm has two forms, where key-policy ABE (KP-ABE) and key-ciphertext (CP-ABE). KP-ABE and CP-ABE were both proposed in [8]. In the KP-ABE scheme, the access policy is connected with the private key and hidden in it. In the CP-ABE scheme, the access policy is connected with ciphertext and hidden in it. Goyal *et al.* [8] first presented the KP-ABE algorithm and the algorithm realized the goal of monotonic access policy structure. A flexible access policy was presented by Ostrovsky *et al.* [18] which realizes the goal of non-monotonic access policy structure.

The first CP-ABE scheme was presented by Bethencourt *et al.* [2]. He only proved secure of the scheme with the condition of a generic group model. A more secure scheme was presented by Cheung *et al.* [7] which is proved secure with the condition of the standard model. CP-ABE [2] encryption allows a person to establish an access policy and the plaintext is encrypted according to it. Encryption could be completed if the attributes meet the access policy. The CP-ABE algorithm could realize a fine-grained access control framework. Lee *et al.* [12] proposed a new convertible encryption scheme based on the ElGamal algorithm. To protect attribute privacy and renew access policy, Zhang *et al.* [28] proposed an anonymous CP-ABPRE framework, in which a matching phase is added before the proxy re-encryption phase. Liu *et al.* [17] proposed an online and offline CP-ABE scheme in an electronic health record system. Zhang *et al.* [29] proposed an attribute-based data sharing system which realizes the function of offline key generation and encryption. Li *et al.* [13] proposed a lightweight protocol without public-key encryption and decryption. Zhang *et al.* [30] proposed a policy-hiding CP-ABE scheme and used it to design smart health security and privacy system. Based on the behavior of the receiver, Cao *et al.* [6] analyzed attribute-based encryption. Attribute-based encryption evolved from identity-based encryption. Liu *et al.* [16] improved an anonymous identity-based encryption scheme by removing one decrypting helper and the strong simulator.

Khoshgozaran *et al.* [11] presented a privacy-preserving location-based service (LBS) framework which needs a trusted institution to convert the original geographic coordinate into a new space. Avoiding a trusted institution keeping the location privacy, Paulet *et al.* [19] presented a novel method in which location privacy is obtained by retrieving private information. In order to avoid a large amount of computing consumption of mobile terminals, Lien *et al.* [15] presented a private circular query protocol that solves the problem of privacy and accuracy for privacy-preserving LBS. However, when the number of points becomes enormous, the scheme isn't appropriate for resource-constraint mobile Internet of Things devices. Jung *et al.* [14] proposed a privacy-preserving LBS scheme

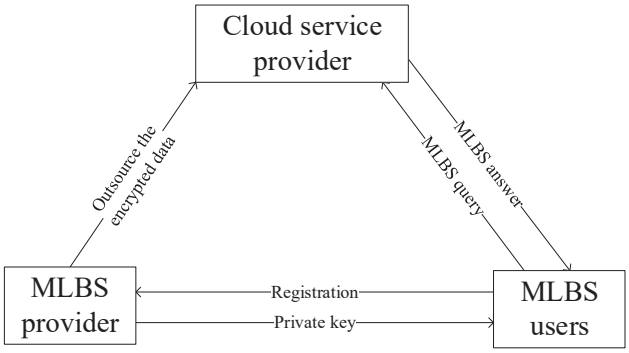


Figure 1: The system model of our proposed framework

according to the ABE algorithm, which doesn't keep the access policy secret. Yu *et al.* [25] first presented a framework that achieves fine-grained access control, scalability and data security in cloud computing. Wang *et al.* [22] presented a framework which achieves practicability and privacy protection search in the cloud environment. But the scheme is not suitable for LBS. Ye *et al.* [24] presented a novel location privacy-preserving framework according to l-queries for continuous LBS using the Paillier public-key cryptosystem. Zhao *et al.* [31] presented a secure and highly efficient LBS framework in which users can retrieve information related to the current geographic coordinate without revealing the geographic coordinate privacy to the CSP. Shao *et al.* [21] presented an LBS framework for mobile devices according to the CP-ABE algorithm, which realizes a fine-grained scheme and privacy-preserving.

3 Model and Design Goals

3.1 System Model

Our system model is the same as the model [9] which provides data services. The system model of our proposed CMPD framework consists of three parts: the MLBS provider, plenty of MLBS users and the CSP which are shown in Figure 1. The following is a detailed description of these three parts.

MLBS provider computes the system public key and system master key. MLBS provider also produces the MLBS data files which have the information of MLBS. The MLBS data files are sent to the cloud server after being encrypted by the MLBS provider. The MLBS users register at the MLBS provider and obtain the attribute private keys from it.

The MLBS users are resource-constraint who want to obtain the MLBS according to his or her attribute private key and geographic coordinate. And our proposed framework doesn't reveal any information about attributes and geographic coordinate.

The cloud service provider storages and computes the data files outsourced from the MLBS provider. We suppose that the cloud service provider could store a large amount of data files and perform fast computing and is always online.

3.2 Security Model

In our proposed framework, we assume that the CSP is honest and curious which is stated in [21]. Specifically, the cloud service provider will faithfully obey our proposed framework, but it can attack the framework as much as possible to get private information. The CSP may try to obtain the plaintext or access policy of the encrypted MLBS data files and the geographic coordinate of the MLBS user by colluding malicious users. But the cloud service provider will store data correctly, perform computing correctly, and send data to the users correctly. The users would access the data files without permission or obtaining the access rights of data files. And the users may attack the framework independently or cooperatively to obtain private information.

3.3 Design Goals

The design goals of our proposed framework are as follows.

1) MLBS data files should be secret for anyone who doesn't have access permissions. Unauthorized MLBS users cannot obtain any information from the ciphertext of the encrypted MLBS data files.

2) Our proposed framework should have the capacity of the collusion-resistance in which the cloud service provider and the MLBS users may collude to obtain the information from the ciphertext of the encrypted MLBS data files when the attributes of a user don't meet the access policy.

3) The privacy of the MLBS users and MLBS data should be protected, where anyone including the cloud service provider cannot obtain any information of user's attributes, geographic coordinate and access policy, even when the cloud service provider is in the match and preliminary decryption phase.

4) High Efficient match and preliminary decryption are needed, because of resource-constraint mobile devices. Since the match and preliminary decryption phases are at the cloud, users only need a small and constant computation before obtaining the plaintext.

5) The CSP should send accurate MLBS data files to the user after it received the user's query which is composed of the geographic coordinate and part of the attribute private key.

6) A flexible access control and fine-grained framework are required because MLBS is a model that provides data services, where data files with different access policies should be decrypted by the MLBS users with different attributes.

4 Preliminaries

In this section, we briefly describe some concepts of cryptogram used in our paper, including the bilinear pairing and some complexity assumptions and access structure.

4.1 Bilinear Pairings

We assume that $\mathbb{G}$ and $\mathbb{G}_T$ are two multiplicative cyclic groups which have a same large prime order p. Assume the generator of $\mathbb{G}$ is g. Suppose the identity of $\mathbb{G}_T$ is $1_{\mathbb{G}_T}$. There is a bilinear pairing if $\mathbf{e} : \mathbb{G} \times \mathbb{G} \to \mathbb{G}_T$ is a map which satisfies the following properties:

1) Bilinear: $\mathbf{e}(\beta_1^a, \beta_2^b) = \mathbf{e}(\beta_1, \beta_2)^{ab}$ for all $\beta_1, \beta_2 \in \mathbb{G}$ and $a, b \in \mathbb{Z}_p$.

2) Non-degenerate: There exists $\beta_1, \beta_2 \in \mathbb{G}$ such that $\mathbf{e}(\beta_1, \beta_2) \neq 1_{\mathbb{G}_T}$.

3) Computable: There exists an efficient algorithm to calculate $\mathbf{e}(\beta_1, \beta_2)$ for all $\beta_1, \beta_2 \in \mathbb{G}$.

4.2 Complexity Assumptions

The Discrete Logarithm (DL) assumption: The DL assumption is right if for any probabilistic polynomial-time (PPT) algorithm, we can obtain the a from the β^a with non-negligible advantage, where unknown a is randomly selected from $\mathbb{Z}_p$ and β is randomly selected from $\mathbb{G}$.

The Decisional Diffie-Hellman (DDH) assumption: The DDH assumption is right if for any PPT algorithm, we can make a distinction between the tuple

$$[\beta, \beta^a, \beta^b, \beta^{ab}]$$

and the tuple

$$[\beta, \beta^a, \beta^b, \beta^z]$$

with non-negligible advantage, where a, b, z are randomly selected from $\mathbb{Z}_p$ and β is randomly selected from $\mathbb{G}$.

The Decisional Bilinear Diffie-Hellman (DBDH) assumption: The DBDH assumption [4] is right if for any PPT algorithm, we can make a distinction between the tuple

$$[\beta, \beta^a, \beta^b, \beta^c, \mathbf{e}(\beta, \beta)^{abc}]$$

and the tuple

$$[\beta, \beta^a, \beta^b, \beta^c, \beta^z]$$

with non-negligible advantage, where a, b, c, z are randomly selected from $\mathbb{Z}_p$ and β is randomly selected from $\mathbb{G}$.

The Decisional Linear Diffie-Hellman (D-Linear) assumption: The D-Linear assumption [3] is right if for all any PPT algorithm, we can make a distinction between the tuple

$$[\beta, \beta^{z_1}, \beta^{z_2}, \beta^{z_1 z_3}, \beta^{z_2 z_4}, \beta^{z_3 + z_4}]$$

and the tuple

$$[\beta, \beta^{z_1}, \beta^{z_2}, \beta^{z_1 z_3}, \beta^{z_2 z_4}, \beta^{z}]$$

with non-negligible advantage, where z, z_1, z_2, z_3, z_4 are randomly selected from $\mathbb{Z}_p$ and β is randomly selected from $\mathbb{G}$.

4.3 Access Structure

Suppose all the users' attributes set is $\mathbb{U} = \{U_1, U_2, \ldots, U_n\}$ in universe, where $|\mathbb{U}| = n$. Each attribute U_i has n_i multiple values, where $U_i = \{u_{i,1}, u_{i,2}, \ldots, u_{i,n_i}\}$ for $1 \leq i \leq n$. We assume the $L = [l_1, l_2, \ldots, l_n]$ is a user's attribute list. In our anonymous CP-ABE scheme, we suppose the access policy structure is a single AND-gate which supports multiple values and wildcards. Formally, there are a user's attribute list $L = [l_1, l_2, \ldots, l_n]$ and a access policy $W = [W_1, W_2, \ldots, W_n]$. For $1 \leq i \leq n$, $L \vDash W$ if $l_i \in W_i$ or $W_i = *$, otherwise $L \nvDash W$. The symbol $\vDash$ and $\nvDash$ mean that L meets or doesn't meet W respectively. And the wildcard $*$ means that the multiple values in this attribute of W is inconsequential.

For example, we assume that the attributes set is $\mathbb{U} = \{U_1, U_2, \ldots, U_6\}$ and the access policy is $W = [u_{1,2}, u_{2,4}, *, u_{4,1}, u_{5,4}, *]$ which the attributes U_3 and U_6 are inconsequential. The attributes satisfies access policy if the user has the multiple values $u_{1,2}$ for U_1, $u_{2,4}$ for U_2, $u_{4,1}$ for U_4, $u_{5,4}$ for U_5 and no matter what the attributes U_3 and U_6 are.

5 Our Anonymous CP-ABE

In this section, we describe our cloud-side matching and preliminary decryption scheme for anonymous CP-ABE, which is improved from the basic scheme of Zhang *et al.* [27]. The scheme protects the user's attributes of privacy and improves decryption efficiency for the resource-constraint mobile devices. More concretely, our scheme has two phases which need the cloud service provider and users to participate in.

In particular, the user needs to send part of the attribute private key and geographic coordinate to the cloud server and the final decryption cost is small and constant, no matter how many attributes exist and how complex the access policy is.

5.1 Proposed Scheme

Our anonymous CP-ABE scheme is as follows.

Setup (1^λ): We assume that $\mathbb{G}$ and $\mathbb{G}_T$ are two multiplicative cyclic groups which have a same large

prime order p and $\mathbf{e} : \mathbb{G} \times \mathbb{G} \to \mathbb{G}_T$ is a bilinear map. Assume the generator of $\mathbb{G}$ is g. Suppose the identity of $\mathbb{G}$ is $1_{\mathbb{G}}$. We define a hash function $H : \{0,1\}^* \to \mathbb{G}$. Suppose all the users' attributes set is $\mathbb{U} = \{U_1, U_2, \ldots, U_n\}$ in universe, where $|\mathbb{U}| = n$. And each attribute U_i has n_i multiple values, where $U_i = \{u_{i,1}, u_{i,2}, \ldots, u_{i,n_i}\}$ for $1 \le i \le n$. Then the $\mathbb{U} = \{u_{i,1}, u_{i,2}, \ldots, u_{i,n_i}\}$ is a set containing all the multiple values. And only MLBS provider knows the index order of elements in the set $\mathbb{U}$ which should be secret. MLBS randomly chooses $y, \in \mathbb{Z}_p$ and $\beta_1, \beta_2, \beta_3, \beta_4 \in \mathbb{G}$. Then MLBS provider computes $Y = \mathbf{e}(\beta_1, \beta_2)^y$. The system public key is $PK =< H, \beta_3, \beta_4 >$ and the system master key is $MK =< y, \beta_1, \beta_2 >$. MLBS provider public the multiple values of all attributes, but the index order of each multiple value in the $\mathbb{U} = \{u_{i,1}, u_{i,2}, \ldots, u_{i,n_i}\}$ is secret.

KeyGen (PK, MK, L): After obtaining a certain user's attributes, MLBS provider makes the user's attribute list $L = [l_1, l_2, \ldots, l_n]$ according to the index order of elements in $\mathbb{U}$. MLBS provider randomly chooses $r_1, r_2, \ldots, r_{n-1} \in \mathbb{Z}_p$ and computes $r_n = y - \sum_{i=1}^{n-1} r_i \mod p$. MLBS randomly chooses $r, \lambda, \hat{\lambda} \in \mathbb{Z}_p$ and $\hat{r}_1, \hat{r}_2, \ldots, \hat{r}_n \in \mathbb{Z}_p$ and computes $\hat{r} = \sum_{i=1}^{n} \hat{r}_i$. MLBS provider randomly chooses $a \in \mathbb{Z}_p$ and then computes $A = \mathbf{e}(\beta_1, \beta_2)^a$. MLBS provider computes $D_{\Delta,0} = \beta_1^r, \hat{D}_{\Delta,0} = \beta_2^{y-\hat{r}}, D_0 = \beta_2^{a\lambda}, \hat{D}_0 = \beta_1^{a\hat{\lambda}}$.

For $1 \le i \le n$, we assume the indexes satisfy $l_i = u_{i,t}$, MLBS provider computes

$$\begin{aligned} D_{i,t,\Delta} &= \beta_2^{\hat{r}_i} H(i \| u_{i,t})^r, \\ D_{i,t} &= \beta_1^{ar_i} H(0 \| i \| u_{i,t})^{a\lambda}, \\ \hat{D}_{i,t} &= \beta_2^{ar_i} H(1 \| i \| u_{i,t})^{a\hat{\lambda}} \end{aligned}$$

Then the attribute private key is $PK_L =< A, D_{\Delta,0}, \hat{D}_{\Delta,0}, D_0, \hat{D}_0, \{D_{i,t,\Delta}, D_{i,t}, \hat{D}_{i,t}\}_{1 \le i \le n} >$.

Encrypt (PK, M, W): MLBS provider encrypts a message $M \in \mathbb{G}_T$ under a ciphertext access policy $W = [W_1, W_2, \ldots, W_n]$. MLBS provider randomly chooses $s, s^{'}, s^{''} \in \mathbb{Z}_p$ and computes $C = MY^s, C_\Delta = Y^{s^{'}}, \hat{C}_0 = \beta_1^{s^{'}}, C_1 = \beta_2^{s^{''}}, \hat{C}_1 = \beta_1^{s-s^{''}}$.

Then for $1 \le i \le n$ and $1 \le t \le n_i$, MLBS provider randomly chooses $\kappa_{i,\Delta}, \kappa_{i,0}, \kappa_{i,1}$ such that $\prod_{i=1}^{n} \kappa_{i,\Delta} = \prod_{i=1}^{n} \kappa_{i,0} = \prod_{i=1}^{n} \kappa_{i,1} = 1_{\mathbb{G}}$, and computes $C_{i,t,\Delta}, C_{i,t}, \hat{C}_{i,t}$ as follows:

1) If $u_{i,t} \in W_i$ or $W_i = *$, then MLBS provider computes

$$\begin{aligned} C_{i,t,\Delta} &= \kappa_{i,\Delta} H(i \| u_{i,t})^{s^{'}}, \\ C_{i,t} &= \kappa_{i,0} H(0 \| i \| u_{i,t})^{s^{''}}, \\ \hat{C}_{i,t} &= \kappa_{i,1} H(1 \| i \| u_{i,t})^{s-s^{''}} \end{aligned}$$

2) If $u_{i,t} \notin W_i$, then MLBS provider randomly chooses $C_{i,t,\Delta}, C_{i,t}, \hat{C}_{i,t} \in \mathbb{G}$.

For simplicity, we denote $[L_x^{file}, L_y^{file}]$ as geographic coordinate of MLBS data files which have the information of location services. Then MLBS provider randomly chooses $z, z^{'} \in \mathbb{Z}_p$ computes $C_{n+1} = \beta_3^{z^{'}}, C_{n+1}^{'} = \beta_4^{z-z^{'}}$ and $C^{'} = \mathbf{e}(\beta_3, H(0\|\beta_3))^{z^{'} L_x^{file}} \cdot \mathbf{e}(\beta_4, H(1\|\beta_4))^{(z-z^{'}) L_y^{file}}$.

Then the ciphertext of M is $CT_W =< C, C^{'}, C_{n+1}, C_{n+1}^{'}, C_\Delta, \hat{C}_0, C_1, \hat{C}_1, \{\{C_{i,t,\Delta}, C_{i,t}, \hat{C}_{i,t}\}_{1 \le t \le n_i}\}_{1 \le i \le n} >$ which is encrypted by MLBS provider under the access policy W.

Decrypt (PK, CT_W, PK_L): The ciphertext CT_W is matched and preliminarily decrypted by cloud service provider. Then the final decryption is performed by the user. Process is as follow:

1) Cloud-side matching phase: The user computes $D_{n+1} = H(0\|\beta_3)^{L_x^{user}}, D_{n+1}^{'} = H(1\|\beta_4)^{L_y^{user}}$, where L_x^{user} and L_y^{user} are the user's geographic coordinate obtained from his or her mobile devices, i.e., mobile phone or smart watch. Then the user sends $D_{n+1}, D_{n+1}^{'}$ and part of the attribute private key PK_L except A to the cloud service provider.

   $L \vDash W$ if and only if

$$\mathbf{e}(\hat{C}_0, \hat{D}_{\Delta,0} \cdot \prod_{i=1}^{n} D_{i,t,\Delta}) = C_\Delta \cdot \mathbf{e}(\prod_{i=1}^{n} C_{i,t,\Delta}, D_{\Delta,0}) \quad (1)$$

   and

$$C^{'} = \mathbf{e}(D_{n+1}, C_{n+1}) \cdot \mathbf{e}(D_{n+1}^{'}, C_{n+1}^{'}). \quad (2)$$

   Otherwise, $L \nvDash W$. The cloud service provider chooses the $C_{i,t,\Delta}$ according to the indexes $[i,t]$ of the $D_{i,t,\Delta}$.

2) Cloud-side preliminary decryption phase: If $L \vDash W$, then cloud service provider computes

$$C^p = \frac{\mathbf{e}(C_1, \prod_{i=1}^{n} D_{i,t}) \cdot \mathbf{e}(\hat{C}_1, \prod_{i=1}^{n} \hat{D}_{i,t})}{\mathbf{e}(\prod_{i=i}^{n} C_{i,t}, D_0) \cdot \mathbf{e}(\prod_{i=1}^{n} \hat{C}_{i,t}, \hat{D}_0)}. \quad (3)$$

   Then the CSP sends C^p and C to the user. Similarly, the CSP chooses the $C_{i,t}$ and $\hat{C}_{i,t}$ according to the indexes $[i,t]$ of the $D_{i,t}$.

3) Final decryption: The user can obtain the plaintext M after computes

$$\frac{C \cdot A}{C^p} \to M. \quad (4)$$

5.2 Consistency of Proposed Scheme

Consistency of Formula (1):

$$\begin{aligned}
&\frac{\mathbf{e}(\hat{C}_0, \hat{D}_{\Delta,0} \cdot \prod_{i=1}^{n} D_{i,t,\Delta})}{\mathbf{e}(\prod_{i=1}^{n} C_{i,t,\Delta}, D_{\Delta,0})} \\
= &\frac{\mathbf{e}(\beta_1^{s'}, \beta_2^{y-\hat{r}} \cdot \prod_{i=1}^{n} \beta_2^{\hat{r}_i} H(i\|u_{i,t})^r)}{\mathbf{e}(\prod_{i=1}^{n} \kappa_{i,\Delta} H(i\|u_{i,t})^{s'}, \beta_1^r)} \\
= &\mathbf{e}(\beta_1^{s'}, \beta_2^{y-\hat{r}} \cdot \prod_{i=1}^{n} \beta^{\hat{r}_i}) \\
= &\mathbf{e}(\beta_1^{s'}, \beta_2^{y-\hat{r}} \cdot \beta^{\hat{r}}) \\
= &\mathbf{e}(\beta_1^{s'}, \beta_2^{y}) \\
= &C_\Delta
\end{aligned}$$

Consistency of Formula (2):

If the following equation holds:

$$\begin{aligned}
&\mathbf{e}(D_{n+1}, C_{n+1}) \cdot \mathbf{e}(D'_{n+1}, C'_{n+1}) \\
= &\mathbf{e}(H(0\|\beta_3)^{L_x^{user}}, \beta_3^{z'}) \cdot \mathbf{e}(H(1\|\beta_4)^{L_y^{user}}, \beta_4^{z-z'}) \\
= &\mathbf{e}(\beta_3, H(0\|\beta_3))^{z' L_x^{user}} \cdot \mathbf{e}(\beta_4, H(1\|\beta_4))^{(z-z')L_y^{user}} \\
= &C'
\end{aligned}$$

It means that the user's geographic coordinate is equal to the MLBS data file's.

Consistency of Formula (3):

$$\begin{aligned}
C^p = &\frac{\mathbf{e}(C_1, \prod_{i=1}^{n} D_{i,t}) \cdot \mathbf{e}(\hat{C}_1, \prod_{i=1}^{n} \hat{D}_{i,t})}{\mathbf{e}(\prod_{i=i}^{n} C_{i,t}, D_0) \cdot \mathbf{e}(\prod_{i=1}^{n} \hat{C}_{i,t}, \hat{D}_0)} \\
= &\frac{\mathbf{e}(\beta_2^{s''}, \prod_{i=1}^{n} \beta_1^{ar_i} H(0\|i\|u_{i,t})^{a\lambda})}{\mathbf{e}(\prod_{i=1}^{n} \kappa_{i,0} H(0\|i\|u_{i,t})^{s''}, \beta_2^{a\lambda})} \\
&\cdot \frac{\mathbf{e}(\beta_1^{s-s''}, \prod_{i=1}^{n} \beta_2^{ar_i} H(1\|i\|u_{i,t})^{a\hat{\lambda}})}{\mathbf{e}(\prod_{i=1}^{n} \kappa_{i,1} H(1\|i\|u_{i,t})^{s-s''}, \beta_1^{a\hat{\lambda}})} \\
= &\mathbf{e}(\beta_2^{s''}, \prod_{i=1}^{n} \beta_1^{ar_i}) \cdot \mathbf{e}(\beta_1^{s-s''}, \prod_{i=1}^{n} \beta_2^{ar_i}) \\
= &\mathbf{e}(\beta_1, \beta_2)^{y \cdot s \cdot a}
\end{aligned}$$

Consistency of Formula (4):

$$\frac{C \cdot A}{C^p} = \frac{MY^s \cdot \mathbf{e}(\beta_1, \beta_2)^a}{\mathbf{e}(\beta_1, \beta_2)^{y \cdot s \cdot a}} = \frac{MY^s}{Y^s} = M.$$

6 Our CMPD Framework

In this section, we describe our mobile location-based service framework supporting Cloud-side Matching and Preliminary Decryption. Now, we give the description of our CMPD framework as follows.

System Initialization: The MLBS provider chooses a security parameter λ with which a system public key PK and a system master key MK are generated by running the algorithm **Setup** (1^λ). Then the MLBS provider publics the PK and keeps MK secret.

MLBS Data Files Encryption: The MLBS provider produces MLBS data files and then encrypts the data files as follows.

The MLBS provider chooses a symmetric encryption algorithm and randomly chooses a secret key py from the key space. Different MLBS data files could have different secret keys. Then the MLBS provider encrypts the MLBS data file with the py using the symmetric encryption algorithm. We denote the ciphertext as C^{file} .

The MLBS provider defines an access policy W for an MLBS data file. Different MLBS data files could have different access policies. Then the MLBS provider encrypts the secret key py by running the algorithm **Encrypt** (PK, py, W). We denote the ciphertext as CT_W.

Finally, the format of each MLBS data file is (C_{file}, CT_W) as shown in Table 1. Then the MLBS provider outscores (C_{file}, CT_W) to the CSP.

Table 1: Format of a MLBS data file

MLBS data file	CT_W	C_{file}

Users Registration: The MLBS user who wants to obtain mobile location-based service sends his or her attributes to the MLBS provider. After running the algorithm **KeyGen** (PK, MK, L), the MLBS provider distributes corresponding the attribute private key PK_L to the MLBS user.

Users Query: We assume the geographic coordinate range within which the user wants to obtain MLBS is $[L_x^{R-}, L_x^{R+}]$ and $[L_y^{R-}, L_y^{R+}]$ as shown in Figure 2. There are many MLBS data files in this range. The MLBS user can obtain the corresponding MLBS data file if

$$L_x^{R-} \leq (L_x^{user} - L_x^{file}) \leq L_x^{R+}$$

and

$$L_y^{R-} \leq (L_y^{user} - L_y^{file}) \leq L_y^{R+}.$$

When an MLBS user who has registered in the MLBS provider wants to obtain mobile location-based service, the user needs to send D_{n+1}, D'_{n+1}, $[L_x^{R-}, L_x^{R+}]$, $[L_y^{R-}, L_y^{R+}]$ and part of attribute private key PK_L except A to the cloud service provider.

Cloud-Side Matching and Preliminary Decryption: After receiving the user's query, the cloud service provider does the following computation:

$$C_\Delta^{test} = \frac{\mathbf{e}(\hat{C}_0, \hat{D}_{\Delta,0} \cdot \prod_{i=1}^{n} D_{i,t,\Delta})}{\mathbf{e}(\prod_{i=1}^{n} C_{i,t,\Delta}, D_{\Delta,0})}$$

and for each value $R_x \in [L_x^{R-}, L_x^{R+}]$ and $R_y \in [L_y^{R-}, L_y^{R+}]$,

$$C'_{test} = \mathbf{e}(D_{n+1} \cdot H(0\|\beta_3)^{R_x}, C_{n+1})$$

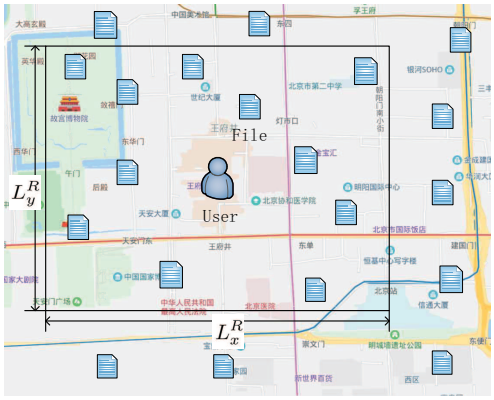


Figure 2: The MLBS geographic coordinate range

$$\cdot \mathbf{e}(D_{n+1}^{'} \cdot H(1\|\beta_4)^{R_y}, C_{n+1}^{'}).$$

The MLBS user can obtain corresponding MLBS data file if and only if $C_{\Delta}^{test} = C_{\Delta}$ and $C_{test}^{'} = C^{'}$. Otherwise, the user cannot obtain any MLBS.

If the matching phase is successful, then the CSP does the preliminary decryption phase. Finally, the CSP sends C^p, C and C_{file} of corresponding MLBS data file to the user.

Getting MLBS: The MLBS user can obtain the secret key pk after computing $\frac{C \cdot A}{C^p} \rightarrow py$. Then the MLBS user can obtain corresponding MLBS after using py to decrypt C_{file}.

7 Security Analysis

In this section, we describe the security of our proposed CMPD framework.

The confidentiality of MLBS data: In our proposed CMPD framework, we first encrypt the MLBS data file using a symmetric encryption algorithm. Then we encrypt the secret key py using our CP-ABE algorithm. Only the MLBS user knows the transformation key A. As analyzed in [21], the cloud service provider transforms our CP-ABE ciphertext into ElGamal type ciphertext according to the transformation key A. The cloud service provider can obtain no information about the access policy without knowing A, because of the security of ElGamal type ciphertext. Hence, no one except the user can obtain information about plaintext from the corresponding encrypted MLBS data file.

The privacy of access policy and user's attributes: We analyze the privacy of access policy in our proposed CMPD framework by comparing it with the basic anonymous CP-ABE scheme presented by Zhang *et al.* [27], which is proved the selective ciphertext-policy and chosen-plaintext secure (CPA-secure) under the condition of the DBDH assumption and the D-Linear assumption.

Our anonymous CP-ABE partial ciphertext components

$$< C_{\Delta}, \hat{C}_0, C_1, \hat{C}_1, \{\{C_{i,t,\Delta}, C_{i,t}, \hat{C}_{i,t}\}_{1 \le t \le n_i}\}_{1 \le i \le n} >$$

and partial attribute private key components

$$< D_{\Delta,0}, \hat{D}_{\Delta,0}, \{D_{i,t,\Delta}\}_{1 \le i \le n} >$$

are computed the same as the basic anonymous CP-ABE scheme presented by Zhang *et al.* [27]. As analyzed in [27], these above components don't reveal any information of the access policy. The matching and preliminary decryption phases of our framework perform at the cloud service provider, the corresponding components of ciphertext are chosen based on the indexes $[i,t]$ of the attribute private key. Since the index order of each multiple value in the $\mathbb{U} = \{u_{i,1}, u_{i,2}, \ldots, u_{i,n_i}\}$ is secret, people who only know the indexes $[i,t]$ don't know what the corresponding multiple value $u_{i,t}$ is. Since MLBS provider keeps β_1, β_2 secret, our proposed framework avoids the combination of CSP and users guessing access policies and user's attributes. Hence, our framework don't reveal any information of the access policy and user's attributes.

The confidentiality of geographic coordinate: The geographic coordinate L_x^{file}, L_y^{file} of the MLBS data file are in the formula

$$\mathbf{e}(\beta_3, H(0\|\beta_3))^{z^{'} L_x^{file}}, \mathbf{e}(\beta_4, H(1\|\beta_4))^{(z-z^{'}) L_y^{file}}.$$

Hence, the information of L_x^{file}, L_y^{file} cannot be revealed even in the matching phase, because of the complexity assumptions. Similarly, the geographic coordinate L_x^{user}, L_y^{user} of the MLBS user are in the formula $H(0\|\beta_3)^{L_x^{user}}, H(1\|\beta_4)^{L_y^{user}}$ and the information of L_x^{file}, L_y^{file} cannot be revealed even in the matching phase, because of the complexity assumptions. Hence, our framework cannot reveal any information of geographic coordinate whether it belongs to MLBS data files or MLBS users.

8 Performance Analysis

In this section, we compare our proposed framework with anther three LBS frameworks [21, 23, 32] in terms of efficiency.

Table 2 mainly introduces the efficiency comparison including PK size, pairings in decryption, whether there is a matching phase and the computation on the user-side. We denote the bit length of an element in $\mathbb{G}$ and $\mathbb{G}_T$ as $|\mathbb{G}|$ and $|\mathbb{G}_T|$ respectively. We denote by $\mathcal{M}$ a multiplication operation, by $\mathcal{D}$ a division operation, by $\mathcal{E}$ an exponentiation operation, by $\mathcal{H}$ a hash operation, by $\mathcal{P}$ a public key cryptographic algorithm operation and by

Table 2: The efficiency comparison of LBS schemes

Schemes	Fine-grained	Pairings in decryption	Decryption matching	User-side computation
[21]	✓	$8N+2$	×	$2\mathcal{M}+\mathcal{D}+3\mathcal{E}+2\mathcal{H}$
[32]	×	8	×	$2\mathcal{M}+7\mathcal{E}+3\mathcal{H}+2\mathcal{P}+2\mathcal{B}$
[23]	✓	8	×	$(\|S_x\|+1)\mathcal{M}+4\mathcal{D}+3\mathcal{B}$
Our	✓	8	✓	$\mathcal{M}+\mathcal{D}+2\mathcal{E}+2\mathcal{H}$

$\mathcal{B}$ a bilinear pairing operation. In [23], S_x is a set of child nodes and we denote the size of S_x as $|S_x|$. We assume $N=\sum_{i=1}^{n} n_i$.

Our scheme has a matching phase before the preliminary decryption phase in the cloud, which reduces unnecessary computational consumption and improves the efficiency of decryption. In our scheme, the users don't need to compute bilinear pairings, which reduces the computational cost for resource-constraint mobile devices. It means that the user only needs to do some simple calculations during the query MLBS and the final decryption phase. A detailed comparison is shown in the Table 2. In summary, our scheme is more efficient than another.

9 Conclusion

In this paper, we propose a mobile location-based security service framework supporting Cloud-side Matching and Preliminary Decryption. The proposed framework takes advantage of cloud storage and cloud computing and ensures data confidentiality, access policy privacy, user's attribute privacy and geographic coordinate privacy. In particular, the computation cost of terminals is reduced significantly.

Acknowledgments

This research is supported by the National Key R&D Program of China under Grant 2017YFB0802000, the Innovation Capability Support Program of Shaanxi under Grant 2020KJXX-052, the Shaanxi Special Support Program Youth Top-notch Talent Program, the Key Research and Development Program of Shaanxi under Grants 2019KW-053 and 2020ZDLGY08-04, the Natural Science Basic Research Plan in Shaanxi Province of China under Grants 2019JQ-866 and 2018JZ6001. Yinghui Zhang is supported by New Star Team of Xi'an University of Posts and Telecommunications under Grant 2016-02.

References

[1] M. Ali, S. U. Khan, and A. V. Vasilakos, "Security in cloud computing: Opportunities and challenges," *Information Sciences*, vol. 305, pp. 357–383, 2015.

[2] J. Bethencourt, A. Sahai, and B. Waters, "Ciphertext-policy attribute-based encryption," in *IEEE Symposium on Security and Privacy (SP'07)*, pp. 321–334, 2007.

[3] D. Boneh, X. Boyen, and H. Shacham, "Short group signatures," in *Annual International Cryptology Conference*, pp. 41–55, 2004.

[4] D. Boneh and M. Franklin, "Identity-based encryption from the weil pairing," in *Annual International Cryptology Conference*, pp. 213–229, 2001.

[5] D. Boneh and B. Waters, "Conjunctive, subset, and range queries on encrypted data," in *Theory of Cryptography Conference*, pp. 535–554, 2007.

[6] Z. Cao, L. Liu, and Z. Guo, "Ruminations on attribute-based encryption," *International Journal of Electronics and Information Engineering*, vol. 8, no. 1, pp. 9–19, 2018.

[7] L. Cheung and C. Newport, "Provably secure ciphertext policy ABE," in *Proceedings of the 14th ACM Conference on Computer and Communications Security*, pp. 456–465, 2007.

[8] V. Goyal, O. Pandey, A. Sahai, and B. Waters, "Attribute-based encryption for fine-grained access control of encrypted data," in *Proceedings of the 13th ACM Conference on Computer and Communications Security*, pp. 89–98, 2006.

[9] H. Hu, Q. Chen, and J. Xu, "Verdict: Privacy-preserving authentication of range queries in location-based services," in *IEEE 29th International Conference on Data Engineering (ICDE'13)*, pp. 1312–1315, 2013.

[10] A. Kapadia, P. P. Tsang, and S. W. Smith, "Attribute-based publishing with hidden credentials and hidden policies," in *NDSS*, vol. 7, pp. 179–192, 2007.

[11] A. Khoshgozaran and C. Shahabi, "Blind evaluation of nearest neighbor queries using space transformation to preserve location privacy," in *International Symposium on Spatial and Temporal Databases*, pp. 239–257, 2007.

[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] C. T. Li and M. S. Hwang, "A lightweight anonymous routing protocol without public key en/decryptions for wireless ad hoc networks," *Information Sciences*, vol. 181, no. 23, pp. 5333–5347, 2011.

[14] X. Y. Li and T. Jung, "Search me if you can: Privacy-preserving location query service," in *Proceedings IEEE Infocom*, pp. 2760–2768, 2013.

[15] I. T. Lien, Y. H. Lin, J. R. Shieh, and J. L. Wu, "A novel privacy preserving location-based service protocol with secret circular shift for k-nn search," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 6, pp. 863–873, 2013.

[16] L. Liu, Z. Guo, Z. Cao, and Z. Chen, "An improvement of one anonymous identity-based encryption scheme," *International Journal of Electronics and Information Engineering*, vol. 9, no. 1, pp. 11–21, 2018.

[17] Y. Liu, Y. Zhang, J. Ling, and Z. Liu, "Secure and fine-grained access control on e-healthcare records in mobile cloud computing," *Future Generation Computer Systems*, vol. 78, pp. 1020–1026, 2018.

[18] R. Ostrovsky, A. Sahai, and B. Waters, "Attribute-based encryption with non-monotonic access structures," in *Proceedings of the 14th ACM Conference on Computer and Communications Security*, pp. 195–203, 2007.

[19] R. Paulet, M. G. Kaosar, X. Yi, and E. Bertino, "Privacy-preserving and content-protecting location based queries," *IEEE Transactions on Knowledge and Data Engineering*, vol. 26, no. 5, pp. 1200–1210, 2013.

[20] A. Sahai and B. Waters, "Fuzzy identity-based encryption," in *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pp. 457–473, 2005.

[21] J. Shao, R. Lu, and X. Lin, "Fine: A fine-grained privacy-preserving location-based service framework for mobile devices," in *IEEE Conference on Computer Communications*, pp. 244–252, 2014.

[22] C. Wang, K. Ren, S. Yu, and K. M. R. Urs, "Achieving usable and privacy-assured similarity search over outsourced cloud data," in *Proceedings IEEE Infocom*, pp. 451–459, 2012.

[23] Y. Xue, J. Hong, W. Li, K. Xue, and P. Hong, "Labac: A location-aware attribute-based access control scheme for cloud storage," in *IEEE Global Communications Conference (GLOBECOM'16)*, pp. 1–6, 2016.

[24] A. Ye, Y. Li, and L. Xu, "A novel location privacy-preserving scheme based on l-queries for continuous LBS," *Computer Communications*, vol. 98, pp. 1–10, 2017.

[25] S. Yu, C. Wang, K. Ren, and W. Lou, "Achieving secure, scalable, and fine-grained data access control in cloud computing," in *2010 Proceedings IEEE Infocom*, pp. 1–9, 2010.

[26] Y. Zhang, X. Chen, J. Li, D. S. Wong, and H. Li, "Anonymous attribute-based encryption supporting efficient decryption test," in *Proceedings of the 8th ACM SIGSAC Symposium on Information, Computer and Communications Security*, pp. 511–516, 2013.

[27] Y. Zhang, X. Chen, J. Li, D. S. Wong, H. Li, and I. You, "Ensuring attribute privacy protection and fast decryption for outsourced data security in mobile cloud computing," *Information Sciences*, vol. 379, pp. 42–61, 2017.

[28] Y. Zhang, J. Li, X. Chen, and H. Li, "Anonymous attribute-based proxy re-encryption for access control in cloud computing," *Security and Communication Networks*, vol. 9, no. 14, pp. 2397–2411, 2016.

[29] Y. Zhang, A. Wu, and D. Zheng, "Efficient and privacy-aware attribute-based data sharing in mobile cloud computing," *Journal of Ambient Intelligence and Humanized Computing*, vol. 9, no. 4, pp. 1039–1048, 2018.

[30] Y. Zhang, D. Zheng, and R. H. Deng, "Security and privacy in smart health: Efficient policy-hiding attribute-based access control," *IEEE Internet of Things Journal*, vol. 5, no. 3, pp. 2130–2145, 2018.

[31] X. Zhao, H. Gao, L. Li, H. Liu, and G. Xue, "An efficient privacy preserving location based service system," in *IEEE Global Communications Conference*, pp. 576–581, 2014.

[32] H. Zhu, R. Lu, C. Huang, L. Chen, and H. Li, "An efficient privacy-preserving location-based services query scheme in outsourced cloud," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 9, pp. 7729–7739, 2015.

Biography

Yinghui Zhang is a professor of National Engineering Laboratory for Wireless Security (NELWS), Xi'an University of Posts & Telecommunications since 2018. He has published over 80 research articles in ACM ASIACCS, IEEE Transactions on Dependable and Secure Computing, IEEE Transactions on Services Computing, Computer Networks, IEEE Internet of Things Journal, Computers & Security, IEEE Transactions on Industrial Informatics, etc. His research interests include public key cryptography, cloud security and wireless network security.

Xinwei Ma received the B.S. degree from the Northwestern Polytechnical University Mingde College in 2017. He is currently pursuing the M.Eng. degree with the Xi'an University of Post and Telecommunications, Xi'an, China. His research interests include attribute-based encryption.

Axin Wu received B.S. degree from Zhengzhou University of Light Industry in 2016, and M.Eng. degree from Xi'an University of Post and Telecommunications in 2019. Since then, he is currently in Ph.D program in Jinan University, Guangzhou, China. His research interests include cloud security and wireless network security.

Fangyuan Ren received the B.S. degree from the Xi'an University of Post and Telecommunications in 2017. She is currently pursuing the M.Eng. degree with the Xi'an University of Post and Telecommunications, Xi'an, China.

His research interests include authentication in 5G.

Dong Zheng received his Ph.D. degree in communication engineering from Xidian University, China, in 1999. He was a Professor at the School of Information Security Engineering, Shanghai Jiao Tong University. He is currently a Professor at National Engineering Laboratory for Wireless Security, Xi'an University of Posts & Telecommunications. He has published over 100 research articles including CT-RSA, IEEE Transactions on Industrial Electronics, Information Sciences, etc. His research interests include cloud computing security, public key cryptography.

Research on the Secure Financial Surveillance Blockchain Systems

Yi-Hui Chen[1,2], Li-Chin Huang[3], Iuon-Chang Lin[4], and Min-Shiang Hwang[5,6]
(Corresponding author: Min-Shiang Hwang)

Department of Information Management, Chang Gung University, Taoyuan 33302, Taiwan [1]
Kawasaki Disease Center, Kaohsiung Chang Gung Memorial Hospital, Kaohsiung 83301, Taiwan[2]
(Email: cyh@gap.cgu.edu.tw)
Department of Information Management, Executive Yuan, Taipei 10058, Taiwan[3]
Department of Management Information Systems, National Chung Hsing University, Taiwan[4]
Department of Computer Science & Information Engineering, Asia University, Taichung, Taiwan[5]
500, Lioufeng Rd., Wufeng, Taichung 41354, Taichung, Taiwan, R.O.C.
Department of Medical Research, China Medical University Hospital, China Medical University, Taiwan[6]
(Email: mshwang@asia.edu.tw)
(Received Apr. 13, 2020; Revised and Accepted June 21, 2020; First Online June 30, 2020)

Abstract

Supply chain finance refers to banks or loan institutions that can provide financial services to industries in which the core enterprise's supply chain industry is located. Blockchain with confidentiality, accountability, non-repudiation, and reliability can be effectively used in supply chain finance. Transactions and individuals can be stored in a distributed ledger, copying copies to each node of the blockchain to prevent data corruption and collapse. In addition, sensitive data stored in the blockchain is never afraid of being stolen, and always maintains its privacy. In this proposal, we intend to raise three research questions to design a mechanism that can not only help with credit checks but also be used for loan ledger management: 1) Research on the preservation and supervision of credit information on the blockchain; 2) Research on the blockchain in post-loan management; 3) Research on the financial supervision chain based on a time sequence.

Keywords: Blockchain; Distributed Ledger; Financial Surveillance Blockchain Systems; Peer-to-Peer Networks

1 Introduction

Supply chain means a unit consisting of a series of upstream suppliers, buyers, retailers, and distributors to the final downstream consumers [26]. From the initial purchaser to purchase raw materials from the supplier to the intermediate manufacturing, design, completion of the final product, and finally deliver the product to the consumer. The network formed by these series of actions and processes is a simple supply chain. In terms of manufacturing, the manufacturing supply chain is roughly composed of three parts: supply, manufacturing, and distribution. Supply is the process by which the purchaser must purchase raw materials from the supplier and strictly supervise its quality before manufacturing the goods; After the manufacturing makes up the raw materials for the purchaser, the raw materials are made or improved into the final product; Distribution is the use of retail, shipping, and orders to deliver the final product to the customer. However, in the real environment, the supply chain is more complicated, and each process of the upstream, middle, and downstream units can supply each other. The supply model is also quite complicated, including issues such as sales methods, efficiency assessment, and payment methods [2]. So not only the manufacturing supply chain that everyone is familiar with, this concept can be used in various fields. This paper mainly focuses on the financial supply chain.

In the traditional financing and lending model of financial institutions, the pre-operation is to collect credit from enterprises or individuals. The content of the credit report is that a third party department other than the bank sorts out its financial status, credit report, and the nature of the department's work, and then delivers it to a credit institution such as a bank as a reference condition for a loan or information. Therefore, traditional financial institutions are more willing to provide financial support than competitive core enterprises. However, smaller companies or upstream and downstream suppliers have higher operational risk and are more resistant to economic fluctuations than core companies. Coupled with the information asymmetry factors of banks and SMEs and the high cost of credit reporting, financial institutions are re-

luctant to provide loans to these suppliers. The result is an imbalance in the supply chain due to difficulties in capital turnover. Because of this, in recent years, various financial institutions have also actively responded to this phenomenon, and supply chain finance came into being.

In short, Supply Chain Finance (SCF) is a bank or lender that seeks and targets specific core companies, uses it as a starting point, and integrates and connects with upstream and downstream companies to provide a financing model of flexible financial products and services [2]. It is effectively investing funds in upstream and downstream of small and medium-sized enterprises (SMEs) can make capital flow flexibly in the supply chain and solve problems such as financing difficulties and supply chain imbalances for SMEs. Therefore, supply chain financing also has another name, called supply chain financing. The biggest difference between supply chain finance and the above-mentioned traditional financial credit industry is that through the evaluation of the credit and capabilities of specific core companies, financial institutions share information asymmetry between upstream and downstream SMEs [27]. So that upstream and downstream SMEs can obtain sufficient financial support to maintain the balanced development of the supply chain, as shown in Figure 1.

There are already international success stories, the most widely known being the Mexican National Financial Development Bank (NAFIN). The bank obtains information about large European and American channel providers and establishes contacts with domestic SME. The bank used the high credit of European and American core enterprises to reduce the default risk or other risks of SMEs. It successfully established a multinational supply chain financial platform between core enterprises and suppliers. However, due to the development of traditional supply chain finance, it only provides services for about 15% of suppliers. 85% of SMEs still suffer from cash flow.

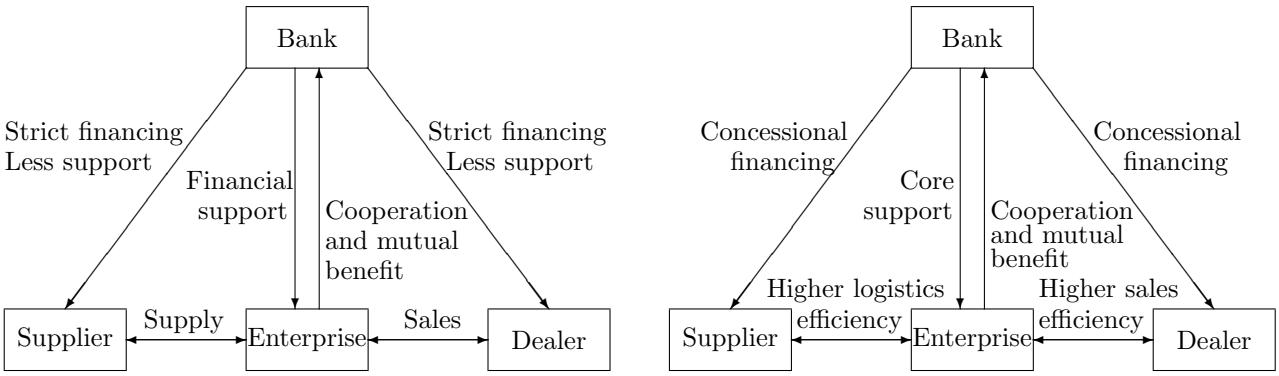
In order to solve this financial problem,blockchain is an opportunity for commercial financing. Blockchain is a decentralized database in the form of a peer-to-peer network, also known as a decentralized ledger [6, 14–16]. It has the characteristics of decentralized, permanent storage of data, and no tampering of data. Initially, it was mainly used for virtual encrypted electronic money. When it comes to blockchain cryptocurrency, one must mention the representative work of blockchain technology-Bitcoin. Satoshi Nakamoto released the Bitcoin white paper (Bitcoin: a peer-to-peer electronic cash system) in 2008, and officially released Bitcoin in 2009 [19, 24]. It is a virtual currency composed of point-to-point, proof of work, encryption and decryption technology, and electronic signature [4]. It is also the most famous cryptocurrency with the highest market value and unit value. According to official bitcoin information, as of 2017/12/05, the official exchange rate was 11726.03 US dollars. Compared with about 300 US dollars in 2016/12, it has increased by more than 30 times, and also let us know the infinite potential and possibilities of blockchain technology.

Satoshi Nakamoto issued the first Bitcoin block in 2009, which we call the Genius block. The field of the block contains the hash value function, timestamp, transaction number and content of the block, block index, difficulty calculation, etc. The value in the blockchain field is a hash function composed of multiple bits, which is a hexadecimal hash function generated after calculation by the encryption algorithm SHA-256 (secure hash algorithm) [7,8,10]. It is easy to verify, but difficult to reverse. This algorithm calculates the public wallet address and hash value function of each transaction. The reason why the blockchain is called "chain" is that the hash function of each new block is derived from the hash function of the previous block, and then calculated by an algorithm. With such infinite expansion, there will be a feeling of putting each block "chain" together. When a malicious attack changes the contents of a block, the block hash function will become completely different. It is why the blockchain has the characteristics of non-tampering. Because as long as the hash function can be easily compared, you can know whether the block is correct. Blockchain is not only used for cryptocurrencies but also widely used in IoT (Internet of Things) [1], DRM (Digital Rights Management), healthcare, and stock debt. The term smart contract also appeared, often referred to as blockchain 2.0, and then evolved into blockchain 3.0 again. Many companies in the world develop blockchain technology, such as IBM, R3 Alliance, and Ripple.

2 Related Works

In the past, experts and scholars have conducted research on financial lending and supervision, focusing on risk management and risk prediction [3, 11, 20] and its efficiency [9]. Through analysis to reduce risks [5, 21], improve financial stability [13]. Tao analyzed the relationship between companies through game theory [22]. Luo and Zhang used game theory analysis and provided recommendations to participants and regulators to solve financial risks, reduce regulatory costs, and improve the effectiveness of financial regulation [17]. Huang proposed data analysis to improve the effectiveness and safety of financial control [9]. Zhang proposed that the module structure is based on the financial supervision system, and it will immediately issue a warning signal when it encounters economic conditions [29].

In order to meet these requirements, many corresponding systems have also been proposed, and scholars such as Ye *et al.* proposed a web-based information system [28]. The system can analyze and predict financial risks. However, today's financial commodities are becoming more and more diverse, and trading systems are becoming more and more complex. To adapt to a more variable trading system, Tsai proposed to apply intelligent blockchain technology to financial supervision [23]. They use corpo-



(a) Traditional financial model (b) Supply chain financial model

Figure 1: The financial mode

rate bonds as an application area and design more complex and variable trading systems for commodities. Use the decentralized advantages of blockchain to eliminate the need for direct peer-to-peer transactions for intermediaries. Make transactions transparent, safe, and anonymous. However, this method does not consider loans to SMEs because the loan amount is not high, but the audit cost is too high.

Mainelli & Smith proposed a qualitative multi-attribute model to identify the situation where the price of a single stock is affected by fraudsters who will actively promote stocks to support decisions in the field of financial market monitoring [18]. And provide data based on a large number of abnormal conditions. Lee *et al.* proposed the forward and backward analysis methods of the FDB Miner (FDBM) Information Extraction system [12]. This method aims to detect potential illegal Pump and Dump comments on the FDB by integrating the stock price per minute during the detection process to reduce false positives during the detection process. Wang *et al.* proposed a blockchain loan (LoC) in 2019 [25]. It is an intelligent financial loan management system based on smart contracts. They designed a digital account model for transferring assets between centralized and decentralized ledgers and introduced digital signatures to protect data privacy.

This article will propose three research directions:

1) Use blockchain traceability and transparency to integrate supply chain finance. In addition to making each transaction in the supply chain more comfortable to manage, it also allows financial institutions to easily view past import, export, or transaction data for SMEs. And under the protection of the security mechanism that cannot be tampered with or forged by the blockchain, the cost of bank credit and audit operations is significantly reduced. It makes the short-term capital turnover of SMEs easier and maintains the stability of the supply chain.

2) Tracking problems after loan. Post-loan tracking of traditional financial institutions is an important issue. In the past, few studies have proposed methods to solve the problem of severe difficulty and high cost of post-loan tracking.

3) Although personal privacy information is protected, it has resulted in a large number of financial institutions (such as banks, securities, insurance, trusts, and funds) signing a large number of personal data consent applications, resulting in reduced financial supervision efficiency.

Therefore, the financial system needs more effective and safer post-loan management and economic supervision methods.

3 Research Issues

In the blockchain environment architecture proposed in this article, the blockchain functions and architecture can be used in the industrial supply chain to improve user safety and performance. We apply various record keeping and encryption protection technologies to the blockchain environment. In this section, we propose three research topics: (1) Research on the preservation and supervision of credit information on the blockchain; (2) Research on the blockchain in post-loan management; (3) Research on the financial supervision chain based on a time sequence. Figure 2 shows the research architecture and describe it as follows.

3.1 Research on the Preservation and Supervision of Credit Information on the Blockchain

The traditional supply chain system is called the B2B (Business to Business) model. Because any node in the supply chain is only connected to its upstream or upstream, it does not fully understand the upward or downward supply chain relationship. However, in the era of

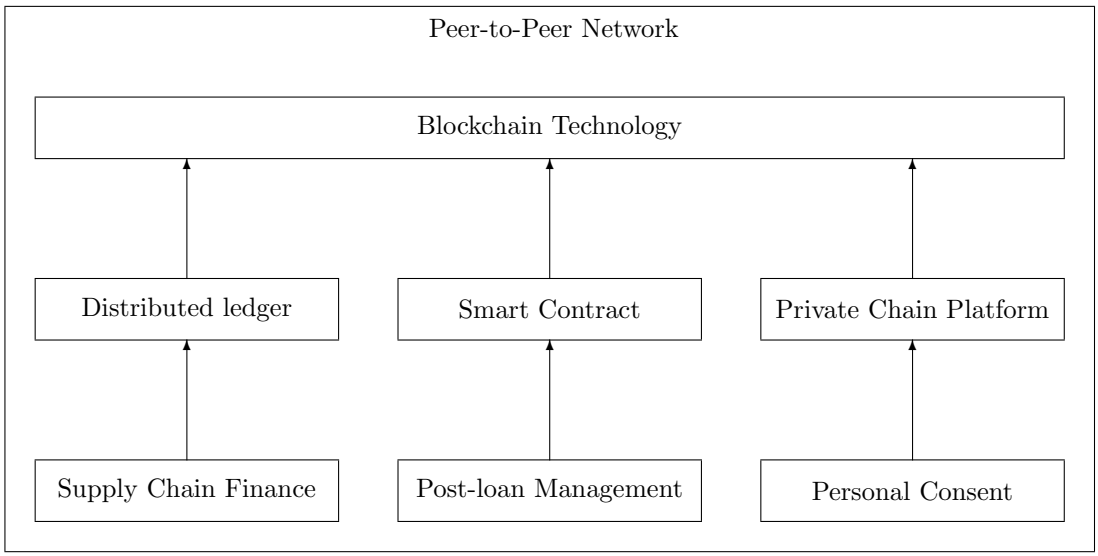


Figure 2: The research architecture

economic globalization, the product design from the initial raw material purchase to the final product model sales can be regarded as a long-chain single-chain system. However, the information currently understood by most nodes is limited to the units directly related to it, which is also one of the factors that cause asymmetric information in the supply chain.

In the absence of information, it is difficult for SMEs in the supply chain to verify whether they are one of the core business partners. Besides, the small size of the enterprise will also lead to a small amount of borrowing. When the audit costs of traditional financial institutions are too high, they will naturally resist the financing activities of SMEs. If we integrate the traceability and transparency of the blockchain into the supply chain finance, in addition to making each transaction in the supply chain more comfortable to manage, it also allows financial institutions to quickly check the past imports and exports of SMEs And transaction data. And under the protection of the security mechanism that cannot be tampered with or forged by the blockchain, the cost of bank credit and audit operations is significantly reduced. It makes the short-term capital turnover of SMEs easier and maintains the stability of the supply chain.

In the research theme, the use of blockchain transparency and unforgeable features will solve the problem of supply chain information asymmetry. The research topic will focus on necessary blockchain technology, including consensus algorithms, cryptographic hash functions, etc. And design a set of decentralized ledger platform for blockchain. For traditional records and supervision, we finally introduced the process from the supply chain to the downstream.

Due to the non-tampering, traceability, and high privacy functions of the private blockchain, a blockchain platform was developed. It can enable banks or other financial institutions to conduct credit investigations on enterprises or individual units faster and save more costs. First, a decentralized ledger must be created to allow units that need to borrow from the bank to retain their supply chain transactions, expenditure income, and other accounting materials. It will enable financial institutions to collect credit information faster and more efficiently. Figure 3 and Figure 4 are the flow charts of this platform.

In Figure 3, as long as any unit initiates a transaction, whether upstream of the supplier or upstream of the manufacturer's supplier, as long as both parties agree to initiate the transaction, the transaction will be stored in the temporary storage pool for confirmation. Then through Figure 4, all pending transaction items are broadcast to each unit in the supply chain. After designing the consensus mechanism of the algorithm, it is finally broadcast to the database system of each unit. Each unit in the supply chain can have the same ledger, which is the last decentralized ledger in Figure 4. In the security mechanism, to avoid tampering, it will lead to the situation of the bank's credit information is untrue, and eventually, cause unnecessary misunderstandings, and even cause credit bankruptcy. We introduced a multi-node authentication mechanism on the public chain. Even if a unit wants to tamper with the content arbitrarily, as long as it is not more than half, each node can compare the old and new data in real-time. According to the principles of our design, the data will never tamper successfully, so in addition to permanent storage, the possibility of tampering is also very low.

The research topic needs to overcome the following potential problems: Although units in the supply chain can use blockchain technology to record at any time, however, data leakage between groups is undesirable, which leads to leakage of business opportunities. When financial institutions ask for credit, will they have too much

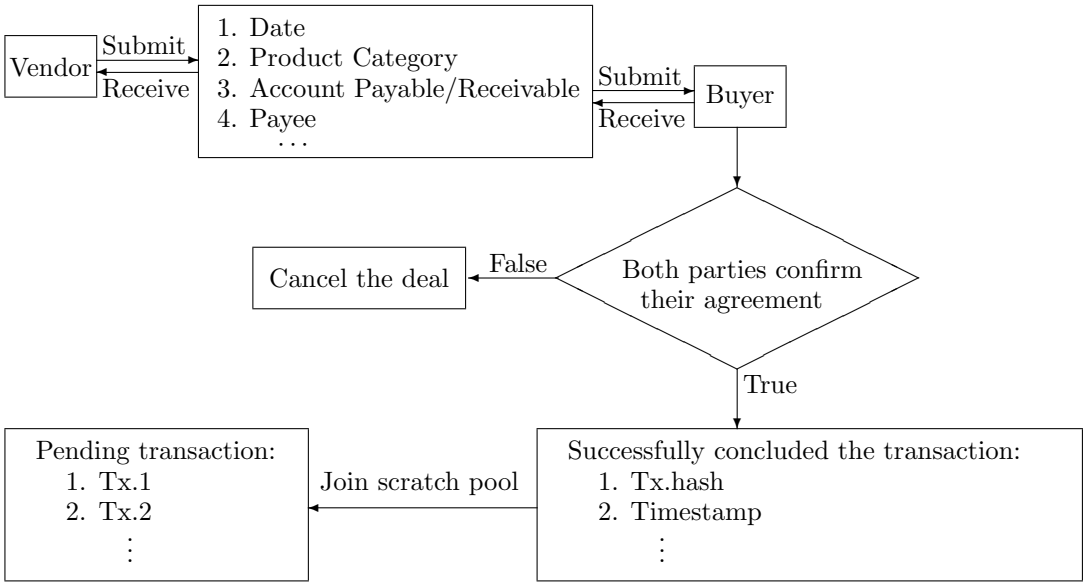


Figure 3: Supplier and buyer forming transaction phase

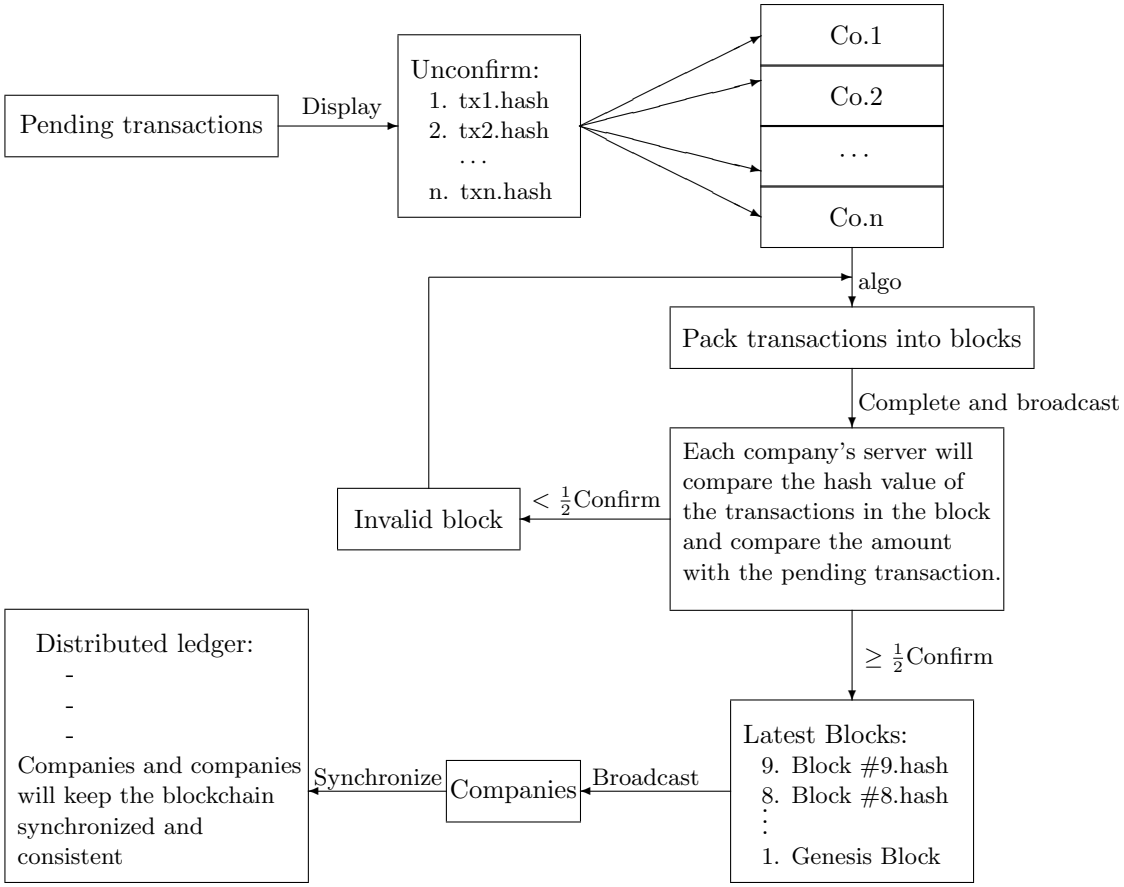


Figure 4: The synchronization to decentralized ledger process

essential data themselves? The specific situation must be considered according to the actual situation, and a more comprehensive overview and analysis are required.

3.2 Research on the Blockchain in Post-loan Management

Post-loan management is the process from financial institution loan issuance to loan recovery. This intermediate link or other aspects of the management process is also the final link after completing the credit work. Contents include loan tracking and review, credit risk management, and daily credit management. The primary purpose is to confirm the customer's repayment willingness and repayment ability and other factors so that financial institutions can ensure loan security and case prevention and control issues, prevent loan overruns, and more effectively control loan risks. Generally speaking, the post-loan management part of financial institutions' credit management is always weak and unreliable.

There are many problems with post-loan management. The client's business and financial situation is usually constantly changing. The client's financial status may be good at the initial credit. Still, later due to environmental prices, economic downturn, investment failure, commercial interference, poor management of the company, or bad policies may be weak. The customer's operating financial situation has an adverse effect. Therefore, post-loan management not only supervises the pure financial operation of customers, but also oversees upstream and downstream manufacturers starting from the customer industry, and even partners and business credits of partners. It is also the information that financial institutions must pay attention to and track. We can only find problems that are not conducive to loan repayment in time and propose appropriate solutions. The blockchain ledger technology is transparent and traceable and cannot be tampered. In addition to checking and sorting the ledger records, it also saves time and effort and also ensures the accuracy of customer data. It is very suitable for recording various matters of post-loan management, effectively reducing costs and improving operational efficiency.

In the research topic, the transparency of using blockchain will improve the efficiency of post-loan management. This topic requires an understanding of the loan process between domestic and foreign financial institutions and companies. For example, a letter of credit guaranteed by a third party, an ERP system (Enterprise Resource Planning) for enterprise resource planning, a continuous loan after the initial investment is completed, and a post-loan management part before full repayment. And research and integrate smart contract deployment on the Ethereum platform to save a lot of time and labor costs.

From the initial loan to the final repayment, the traditional financial post-loan management is the weakest and most inefficient part. The main reason is that tracking is more difficult and costly. Imagine that if a financial institution lends tens of thousands of individual units or enterprises, it will be an extensive project to track the financial status after a loan. To build a blockchain platform, use the blockchain to store all data, and even use the blockchain 3.0 smart contract for many applications. Financial institutions create their smart contracts for each unit or enterprise, allowing lenders to perform operations such as automatic repayment or transaction deferral. Banks will save the time and cost of collecting payments regularly, and can even keep banks from requesting money from units that cannot repay loans.

The following are the blockchain post-loan management process of the research topics:

1) Overseas companies sign contracts and place orders with domestic manufacturers.

2) Overseas companies open individual accounts in the international bank.

3) Overseas companies remit 30% of the first part to the individual account of the international bank.

4) The international bank writes letters of credit in the form of smart contracts and places the deals on the blockchain.

5) The local bank remit 30% of the first part of funds to domestic manufacturers by smart contracts.

6) The domestic manufacturer's ERP system records the warehousing, logistics, and capital usage, and connects the blockchain system to update smart contracts in real-time.

7) The local bank performs post-loan management by updating smart contracts.

8) The local bank gradually issues loans based on the contract performance on smart contracts.

For smart contracts, the most extensive and standards-compliant platform is the Ethereum platform.

The research topic needs to overcome the following potential problems: The use of smart contracts to supervise and control the loan process, replacing traditional letters of credit, etc., are all built on the Ethereum platform. But Ethereum is currently only a transaction between numbers. If you need other functions, you need to develop different platforms and add a smart contract compiler. More research is needed in this area.

3.3 Research on the Financial Supervision Chain Based on a Time Sequence

In financial institutions, no matter what services are used, including foreign currency economic commodity financial management and insurance departments. The first task must be to confirm the identity of the customer and comply with the Personal Data Protection Law. The

purpose of the law is to regulate the collection and application of personal data by relevant units to avoid infringement of personality rights. To promote the cognitive processing and utilization of personal data, the objects of supervision of the Personal Data Protection Law include various institutions and individual units as well as financial banking institutions. But financial institutions include many different departments, and usually, the system of each department is independent. When customers sign the use of the law, different departments of financial institutions also need to repeat the operation once. Traditionally using paper records, in addition to being time-consuming, the process is inefficient. Reduce customer service perception and service willingness. If we use the blockchain to record customer's consent in a financial institution's proprietary account, not only can the information exchange of various departments be integrated, but the customer can also change the personal asset law for the first time when driving the power of the financial institution and solve the time cost of traditional paper records.

In the research topic, the time sequence of using blockchain can help financial institutions to track and monitor customers' wishes. This topic will focus on personal data protection law issues. Under the financial institution system, a private chain combined with a decentralized ledger platform is established. The consumer's privacy law application is stored on the blockchain, and the chronological order is used as the final basis. It does not affect the independent database between the various departments of the bank so that the information can be circulated among the different departments, and it also replaces the traditional paper, causing problems of waste of resources and difficult to save.

Because of this, the subject of this study is to build a proprietary private chain system platform for financial institutions. The system is specifically designed to store consumers' use of personal data, including the current signature time and whether they agree with the use of financial institutions. After using blockchain technology to form a record book, this information will be broadcast to each department's independent database system under the same financial institution. The data content of each unit is the same, and financial institutions can also set up other database systems or cloud systems to store these data. When more nodes have this account, both security and accuracy can be guaranteed.

This topic needs to consider many details, such as how to verify that the consumer's current signature is a real person, and whether to give consumers the right to query their signature status. There is also the speed and difficulty of each node verification and so on. It also involves the design of functions and algorithms in the system.

4 Conclusions

In this article, we have proposed three research topics: 1) Research on the preservation and supervision of credit information on the blockchain; 2) Research on the blockchain in post-loan management; 3) Research on the financial supervision chain based on a time sequence. The purposes of these topics are 1) Credit check and management using blockchain can help banks find it easier Industry transaction data, thereby reducing cost audits and investigations; 2) Loan management using blockchain to track loan repayment ability; 3) Financial monitoring period helps protect user's data.

The research topic needs to overcome the following potential problems: When a customer signs, the traditional method of verifying identity is double certificates, signatures, stamps, or handwriting. However, using blockchain technology, how to provide user-oriented proof when reading data in the future is the biggest challenge currently facing the system.

Acknowledgments

This research was partially supported by the Ministry of Science and Technology, Taiwan (ROC), under contract no.: MOST 108-2410-H-468-023 and MOST 108-2622-8-468-001-TM1.

References

[1] D. S. AbdElminaam, "Smart Kitchen: Automated cooker technique using IoT," *International Journal of Electronics and Information Engineering*, vol. 9, no. 1, pp. 1-10, 2018.

[2] P. Y. Chang, M. S. Hwang, C. C. Yang, "A blockchain-based traceable certification system," in *SICBS 2017: Security with Intelligent Computing and Big-data Services*, Advances in Intelligent Systems and Computing, vol. 733, Springer, pp 363-369, 2018.

[3] S. Chen, Q. Wang, S. Liu, "Credit risk prediction in peer-to-peer lending with ensemble learning framework," in *Chinese Control and Decision Conference (CCDC'19)*, pp. 4373-4377, 2019.

[4] S. Davidson, P. De Filippi, J. Potts, *Economics of Blockchain*, Mar. 8, 2016. (`http://dx.doi.org/10.2139/ssrn.2744751`)

[5] M. Degong, C. Yancun, Y. Wei, "Evaluation of the financial derivative risk from the probability of default angle," in *IEEE 3rd UKSim European Symposium on Computer Modeling and Simulation*, pp. 293-298, 2009.

[6] P. Fan, Y. Liu, J. Zhu, X. Fan, L. Wen, "Identity management security authentication based on blockchain technologies," *International Journal of Network Security*, vol. 21, no. 6, pp. 912-917, 2019.

[7] W. R. Ghanem, M. Shokir, and M. Dessoky, "Defense against selfish PUEA in cognitive radio networks based on hash message authentication code," *International Journal of Electronics and Information Engineering*, vol. 4, no. 1, pp. 12-21, 2016.

[8] H. Huang, X. Chen, Q. Wu, X. Huang, J. Shen, "Bitcoin-based fair payments for outsourcing computations of fog devices," *Future Generation Computer Systems*, vol. 78, pp. 850–858, 2018.

[9] Y. Huang, "Design and implementation of financial supervision and management information system," in *IEEE International Conference on Smart City and Systems Engineering (ICSCSE'16)*, pp. 214-217, 2016.

[10] M. S. Hwang and I. C. Lin, *Introduction to Information and Network Security (6ed, in Chinese)*, Taiwan: Mc Graw Hill, 2017.

[11] Y. Jin, Y. Zhu, "A data-driven approach to predict default risk of loan for online peer-to-peer (P2P) lending," in *15th International Conference on Communication Systems and Network Technologies*, Gwalior, MP, India, pp. 609-613, 2015.

[12] P. S. Lee, M. Owda, K. Crockett, "Novel methods for resolving false positives during the detection of fraudulent activities on stock market financial discussion boards," *International Journal of Advanced Computer Science and Applications*, vol. 9, no. 1, 2018

[13] L. Li and Q. Zhang, "Study of financial stability based on nonlinear dynamic theory," in *2010 International Conference on E-Product E-Service and E-Entertainment*, pp. 1-4, 2010.

[14] Z. C. Li, J. H. Huang, D. Q. Gao, Y. H. Jiang, L. Fan, "ISCP: An improved blockchain consensus protocol," *International Journal of Network Security*, vol. 21, no. 3, pp. 359-367, 2019.

[15] 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, 2017.

[16] Y. Liu, M. He, F. Pu, "Anonymous transaction of digital currency based on blockchain," *International Journal of Network Security*, vol. 22, no. 3, pp. 444-450, 2020.

[17] X. Luo and P. A. Zhang, "Game analysis of financial supervision in international financial crisis," in *Chinese Control and Decision Conference (CCDC'11)*, pp. 46-50, 2011.

[18] M. Mainelli, M. Smith, "Sharing ledgers for sharing economies: an exploration of mutual distributed ledgers (Aka Blockchain Technology)," *The Journal of Financial Perspectives: FinTech*, vol. 3, no. 3, 2015.

[19] S. Nakamoto, *Bitcoin: A Peer-to-peer Electronic Cash System*, 2008.

[20] A. Namvar, M. Naderpour, "Handling uncertainty in social lending credit risk prediction with a Choquet fuzzy integral model," in *IEEE International Conference on Fuzzy Systems (FUZZ-IEEE'18)*, Rio de Janeiro, Brazil, pp. 1-8, 2018.

[21] H. Qingchun, "The ponder on supervision of rural cooperative financial institution in the post-financial crisis era," in *The 2nd International Conference on Information Science and Engineering*, pp. 6305-6307, 2010.

[22] S. Tao, "Lending relationship analysis of micro-loan company on game theory," in *International Conference on Management Science and Industrial Engineering*, pp. 149-152, 2011.

[23] M. W. Tsai, *Applying Smart Blockchain Technology and Financial Supervision Technology to Design Complex Multi-Party Transaction Systems*, Institute of Information Management, Jiaotong University, Master Thesis, 2018.

[24] F. Tschorsch and B. Scheuermann, "Bitcoin and beyond: A technical survey on decentralized digital currencies," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 3, pp. 2084-2123, 2016.

[25] H. Wang, C. Guo, S. Cheng, "LoC - A new financial loan management system based on smart contracts," *Future Generation Computer Systems*, vol. 100, pp. 648–655, 2019.

[26] T. Xiao, T. M. Choi and T. C. E. Cheng, "Pricing and benefit of decentralization for competing supply chains with fixed costs," *IEEE Transactions on Engineering Management*, vol. PP, no. 99, pp. 1-14, 2017.

[27] H. Yan, "Credit model of supply chain finance based on big data of e-commerce," in *4th International Conference on Industrial Economics System and Industrial Security Engineering (IEIS'17)*, pp. 1-4, 2017.

[28] H. Ye, L. Zhu and Y. Gan, "An overall framework study of financial supervision information system based on web service and MAS," in *7th International Conference on Service Systems and Service Management*, pp. 1-5, 2010.

[29] J. Zhang, "Design and implementation of real time warning module for bank financial supervision system," in *International Conference on Computer Systems, Electronics and Control (ICCSEC'17)*, pp. 1174-1179, 2017.

Biography

Yi-Hui Chen received her Ph.D. degree in computer science and information engineering at the National Chung Cheng University. Later on, she worked at Academia Sinica as a post-doctoral fellow. Then, she worked at IBM's Taiwan Collaboratory Research Center as a Research Scientist, the Department of M-Commerce and Multimedia Applications, Asia University as an associate professor. She is now an associate professor at the Department of Information Management, Chang Gung University. Her research interests include multimedia security, semantic web, text mining, and multimedia security.

Li-Chin Huang received the B.S. in computer science

from Providence University, Taiwan, in 1993 and M.S. in information management from Chaoyang University of Technology (CYUT), Taichung, Taiwan, in 2001; and the Ph.D. degree in computer and information science from National Chung Hsing University (NCHU), Taiwan in 2001. Her current research interests include information security, cryptography, medical image, data hiding, network, security, big data, and mobile communications.

Iuon-Chang Lin received the B.S. in Computer and Information Sciences from Tung Hai University, Taichung, Taiwan, Republic of China, in 1998; the M.S. in Information Management from Chaoyang University of Technology, Taiwan, in 2000. He received his Ph.D. in Computer Science and Information Engineering in March 2004 from National Chung Cheng University, Chiayi, Taiwan. He is currently a professor of the Department of Management Information Systems, National Chung Hsing University,and Department of Photonics and Communication Engineering, Asia University, Taichung, Taiwan, ROC. His current research interests include electronic commerce, information security, cryptography, and mobile communications.

Min-Shiang Hwang received M.S. in industrial engineering from National Tsing Hua University, Taiwan in 1988, and a Ph.D. degree in computer and information science from National Chiao Tung University, Taiwan, in 1995. He was a professor and Chairman of the Department of Management Information Systems, NCHU, during 2003-2009. He was also a visiting professor at the University of California (UC), Riverside, and UC. Davis (USA) during 2009-2010. He was a distinguished professor of the Department of Management Information Systems, NCHU, during 2007-2011. He obtained the 1997, 1998, 1999, 2000, and 2001 Excellent Research Award of National Science Council (Taiwan). Dr. Hwang was a dean of the College of Computer Science, Asia University (AU), Taichung, Taiwan. He is currently a chair professor with the Department of Computer Science and Information Engineering, AU. His current research interests include information security, electronic commerce, database and data security, cryptography, image compression, and mobile computing. Dr. Hwang has published over 300+ articles on the above research fields in international journals.

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.