IJNS
International Journal of Network Security

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

Vol. 20, No. 1 (Jan. 2018)

INTERNATIONAL JOURNAL OF NETWORK SECURITY

PUBLISHING OFFICE

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

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

PUBLISHER: Candy C. H. Lin

International Journal of Network Security

Vol. 20, No. 1 (Jan. 1, 2018)

A Survey of Reversible Data Hiding for VQ-Compressed Images

Yu-Lun Wang[1], Jau-Ji Shen[1], Min-Shiang Hwang[2,3],
(Corresponding author: Min-Shiang Hwang)

Department of Management Information Systems, National Chung Hsing University[1]
Department of Computer Science and Information Engineering, Asia University[2]
No. 500, Lioufeng Raod, Wufeng Shiang, Taichung 41354, Taiwan
Department of Medical Research, China Medical University Hospital, China Medical University[3]
No.91, Hsueh-Shih Road, Taichung 40402, Taiwan
(Email: mshwang@asia.edu.tw)
(Invited Jan. 11, 2017)

Abstract

Data hiding has been popular over the past decades, and many different methods of the data hiding have been proposed. Nowadays, data hiding in compressed images becomes more and more popular because of the rise of the social media. Thus, we survey the previous research of data hiding of compressed images to provide the representative methods and analyze their capacities and bit rates. Finally, we propose the future work of data hiding in compressed images.

Keywords: Data Hiding; Reversible Data Hiding; VQ Compressed

1 Introduction

Data hiding has been popular due to the importance of the ownership. Through hiding the data in the image, users can prove the ownership by extracting the unique information [5, 17, 40]. Recently, it becomes more and more popular in social media such as Facebook or Twitter, so images or videos spread on the internet more quickly. While a user uploads the multimedia file on the social media, the social media always compress the image into a smaller size. Therefore, more and more researcher start to propose some new method of the compressed image data hiding [1, 16, 19, 25, 38, 39, 43]. Many compressed methods were proposed in the 1980's, such as VQ, BTC (block truncation coding) and LZW (Lemple-Ziv-Welch). In the data hiding in normal images, we pursue higher capacity and PSNR (peak signal-to-noise ratio) value [19, 28]. However, in the data hiding based on compressed images, higher capacity and lower bit rate are the goal that we pursued. Vector quantization is a popular image-compressed method proposed by Gray [12, 13, 14]. Through a machine learning method to get a codebook, the most famous algorithm was proposed by Linde et al.'s [22]. Figure 1 shows the VQ compression method. Each codebook has some index, and every index has some code words whose value are between 0 and 255.

In the compressed procedure, the image is divided into blocks, and each block has 4×4 pixels, total 16 pixels. The smallest difference between the pixels of the block and the codeword of each index is calculated in the codebook. The index value is used to represent each block.

2 Related Work

In image processing, data hiding has been proposed for many years. The target of the approach is to prevent a malicious user from duplicating or spreading the image on the internet. As mentioned earlier, the VQ compressed method [14] proposed in 1984 and the compressed procedure have been introduced.

In 2004, the first approach of VQ compressed image data hiding was proposed by Chang et al.'s [3] based on search-order coding which was proposed by Hsieh et al. [15]. Their methods have the problem of the expanded file size and acceptable capacity. In 2009, the method proposed by Chen et al.'s [4] has lower bit-rate, but the capacity is also much lower. In the same year, Lee et al.'s proposed [24] used a multiple codebook to enhance the capacity. In 2013, the methods proposed by Pan et al.'s [29] and Chang et al.'s [9] get lower bit-rate, but the capacity is lower than 5000. In 2013, Wang et al. [35] proposed the method joined with the state-codebook mapping. In 2015, Lin et l. [26] proposed the method combined with state-codebook and SOC, significantly decreasing the bit-rate. In 2016, Qin et al. [32] proposed the method added with a relative address type (RA type) into the SOC process to make the bit-rate lower.

In 2006, Chang et al. [10] proposed the method based

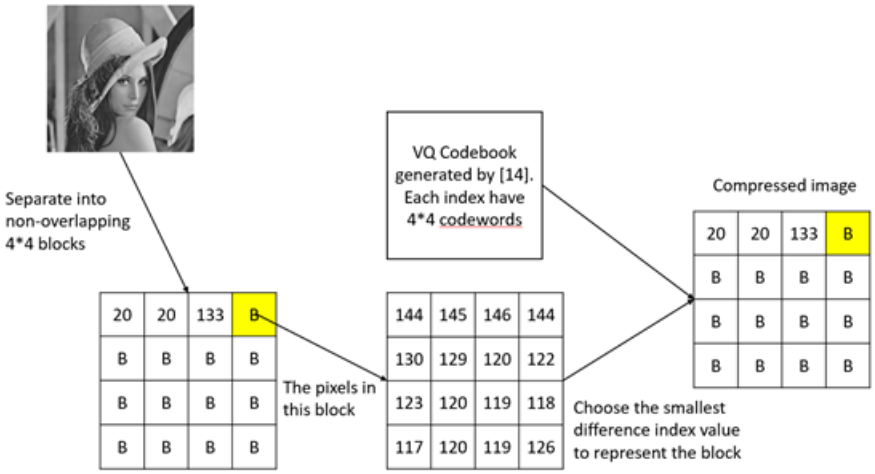


Figure 1: The VQ compression example

on side match VQ compression which was proposed by Kim [20]. In 2010, Lee et al. [23] proposed the method with higher capacity. In 2012, Shie et al. [33] proposed the method with higher capacity and average bit-rate. In 2013, Qin et al. [30] proposed the method based on an index mapping mechanism to have higher capacity. In 2014, Qin et al. [31] proposed the improved method, but the capacity is not good enough. In the same year, Wang et al. [41] proposed the method to increase the capacity, but the capacity is inefficient. In 2009, Chang et al. [7] proposed the method combined with the Joint Neighboring coding (JNC). In the same year, Wang et al. [36] proposed the method also based on JNC. Both methods have higher bit-rate.

In 2015, Kieu et al. [21] proposed the method to improve Chang et al.'s scheme [7]. In 2009, Chang et al. [6] proposed the method based on locally adaptive coding (LAS). In 2010, Yang proposed [42], reduce the bit-rate of Chang et al.'s scheme [6]. In 2011, Chang et al. [8] proposed the improved method to reduce the bit-rate further, but the capacity is still not enough. In 2015, Ma et al. [27] proposed the method based on LAS to improve LAS, decreasing the bit-rate obviously. In 2007, Chang et al. [11] proposed the method to sort the VQ index referring to the counts, which are separated into three clusters to embed data. In 2009, Yang et al. [41] proposed the method to increase the capacity and bit-rate. In 2015, Tu et al. [34] proposed the method with higher capacity. In the section 3, we will describe 5 representative method of the data hiding for VQ compression images. In 2012, Chang [2] proposed the method, which has high payload, but the bit-rate is also high too.

3 Representative Approaches

In the reversible data hiding in VQ compressed images, there are many approaches which have been proposed in the past decade. We will introduce five representative approaches in the following subsections.

3.1 Chang-Chen-Lin's Scheme

In 2004, Chang et al.'s proposed the first reversible data hiding scheme in a VQ compressed image [3]. Their methods are combined with the search-order coding (SOC) technique. Therefore, we will first introduce how SOC working. In 1996, Hsieh and Tsai proposed this method [15] which can compress the image more efficiently. After the VQ compressed process, SOC technique is used to compress images again. The process of the SOC is shown in the Figure 2.

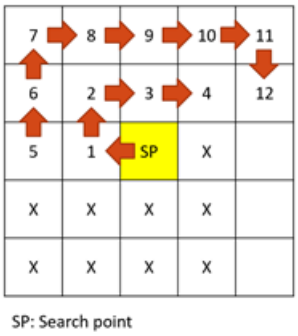


Figure 2: The process of the SOC scheme

The first column and first row will not be used in the SOC process. They choose a search point first, and search by order in the Figure 2. The first index in the search order is the representative bit 00; if the second index in the search order is same as the first index, we skip this index and do nothing. If the index is not the duplicate value compared with the previous index processed, we give the representative bits followed by 01, 10, 11. If the type of the index is SOC type, we add an indicator bit 1 in the head followed by the representative bits. Otherwise,

	OIV indicator = 1	SOC indicator = 0	
Situation 1:	block type = SOC	Situation 3:	block type = OIV
	secret bits = 0		secret bits = 1
	no change		no change
Situation 2:	block type = SOC	Situation 4:	block type = OIV
	secret bits = 1		secret bits = 0
	change the block type to OIV		change to the SOC type

Figure 3: The embedded rule of Chang et al.'s scheme

we add an indicator 0 in the head followed by the original VQ index value. In the Chang et al.'s scheme, they embed the secret bit according to the block type as shown in the Figure 3. In Situation 1, if the secret bit is 1, and the index type is SOC type, we do not change anything in this index. In Situation 2, if the secret bit is 0, and the index type is SOC type, we change the index into the original index value and transform the indicator to 0. In Situation 3, if the secret bit is 0, and the index type is original index value (OIV), we do not change anything in this index. In Situation 4, if the secret bit is 1, and the index type is OIV, we change the index type into SOC type. The change rule is shown in the Figure 4. We can judge whether the SOC type is fake because of the code stream length.

3.2 Chang-Wu-Hu's Scheme

In 2007, Chang et al.'s proposed a new scheme based on the frequency of the index count [11]. In the VQ compressed process, each index has a frequency value; and while the index is used, the frequency value will add 1. After the VQ compressed procedure is finished, they sort the codebook in a descending order according to the frequency value. Then we separate the code book into 3 clusters, and each cluster has $cbs/3$ indexes. In each cluster, cbs represents codebook size, and the remainder of the codebook will be abandoned.

The embedding process is shown in Figure 5. There are four situations in this scheme. First, if the index belongs to the cluster 1, we change the index to belonging to cluster 2 or cluster 3 according to the secret bit. And the length of the index value is always log2cbs. Second, if the index belongs to the cluster 2, we change the index belonging to cluster 1 and do not embed the secret bit in this index. Third, if the index belongs to the cluster 3, we change the index belonging to the cluster 3, and change the index belonging to cluster 1; and we change the length of this index to $\log_2 cbs || \log_2(cbs/3)$. For instance, the first index value is 2, and the secret bit is 1. We change it to the same position in the cluster 3, and change index value to 12. The second index value 1 with secret bit 0 will be changed to 6. The third index value 7 is in the cluster 2, and we change it to cluster 1 with the same position. If the value of the index is 7, we change it to 2. The number in the Figure 4 with underline means this index without secret bits embedding. If next index value 4 with secret bit 1, we change this index to 14. The fifth index value 13 is in the cluster 3, so we change it to an indicator of which length is log2cbs. The codebook size of this example is 15, and the indicator is 0000 followed by the $\log_2(cbs/3)$ bits. The answer of log25 is 3 bits. Therefore, followed by 110, it means 3 in decimal and does not embed any secret bit in this index.

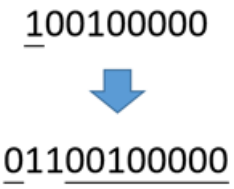


Figure 4: The change rule of Situation 4

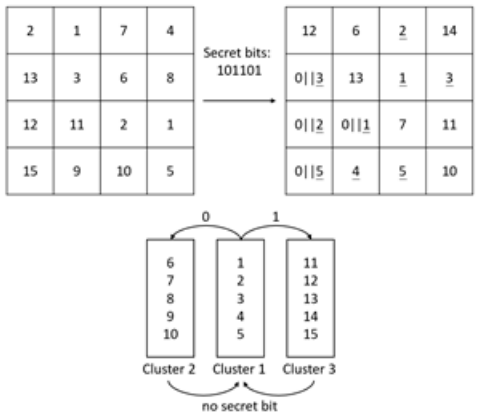


Figure 5: The example of the embedding process of Chang et al.'s method

In the extraction phase, they can accord to the index of the cluster it belongs to decide whether this index has

secret bits or not. If the index belongs to cluster 2 or cluster 3, we can change it to cluster 1 with the same position to recover the original index value. If not, we can determine whether this index belongs to cluster 2 or cluster 3 according to the indicator. For example, if first 4 bits are 0000, we can easily judge that this index belongs to original cluster 3, and the index value is correctly recovered. If the first secret bits are not 0000, we change it to cluster 2 in the same position to recover it.

3.3 Chang-Kieu-Chou's Scheme

In 2009, Chang et al.'s proposed a novel method based on locally adaptive coding (LAS) [6]. LAS method was proposed by Bentley et al. in 1986. We show the process of this method in Figure 6. In the LAS method, it is classified into two situations. In Situation 1, if the word has not been appeared, then we add this word in the list. In Situation 2, if the word has been added in the list, we output the position of the position of the word in the list and move the word to the head of the list.

String: THE CAR ON THE LEFT HIT THE CAR I LEFT		
Input	List	Output
THE	{THE}	1.THE
CAR	{CAR, THE}	2.CAR
ON	{ON, CAR, THE}	3.ON
THE	{THE, ON CAR}	3
LEFT	{LEFT, THE, ON, CAR}	4.LEFT
HIT	{HIT, LEFT, THE, ON, CAR}	5.HIT
THE	{THE, HIT, LEFT, ON, CAR}	3
CAR	{CAR, THE, HIT, LEFT, ON}	5
I	{I, CAR, THE, HIT, LEFT, ON}	6.I
LEFT	{LEFT, I, CAR, THE, HIT, ON}	5

Figure 6: The example of the LAS

In Chang et al.'s scheme, we use LAS method, and the embedding process is shown in Figure 7.

Index value: 45, 200, 150, 45, 150, 100, 120, 200, 120, 30				
Secret bits: 0100				
Input	List	Secret bits	Indicator	Output
45	{45}	X	0	$0\|\|(45)_2$
200	{200, 45}	X	0	$0\|\|(200)_2$
150	{150, 200, 45}	X	0	$0\|\|(150)_2$
45	{45, 150, 200}	0	0	$0\|\|(3)_2$
150	{150, 45, 200}	1	1	$1\|\|(150)_2$
100	{100, 150, 45, 200}	X	0	$0\|\|(100)_2$
120	{120, 100, 150, 45, 200}	X	0	$0\|\|(120)_2$
200	{200, 120, 100, 150, 45}	0	0	$0\|\|(5)_2$
120	{120, 200, 100, 150, 45}	0	0	$0\|\|(2)_2$
30	{30, 120, 200, 100, 150, 45}	X	0	$0\|\|(30)_2$

Figure 7: The example of Chang et al.'s scheme

There are 3 situations in the embedding process. In Situation 1, we scan the index value in scan order. If the current value is not in the list, we add this value in the list and output an indicator 0 and the index value without embedding any secret bit. In Situation 2, if the current value is in the list and the secret bit is 0, we move the index value to the head which is in the list and output the indicator 0 and the original position of the index value in the list. In Situation 3, if the current value is in the list, and the secret bits is 1, we move the index value to the head and output indicator 1 and the value of the index.

3.4 Tu and Wang's Scheme

In 2015, Tu and Wang's proposed an improved method based on the referred frequency [34]. Their scheme is similar to Chang et al.'s scheme [6], and the difference is that they use the index of the cluster 2 and cluster 3 to embed secret bits. The different part is shown on Figure 8. If the original index value belongs to cluster 2, we want to change it to cluster 1 and also embed some secret bits in this index. We change it to the cluster 1 in the same position followed by an indicator and secret bits. The indicator 0 means this index originally belongs to the cluster 2. And if the index originally belongs to cluster 3, the bit of the indicator will be 1. This method can decide how many secret bits we want to embed in each index, but the bit rate will change according to the secret bits we embed. More secret bits we embed mean the bit rate is less efficient.

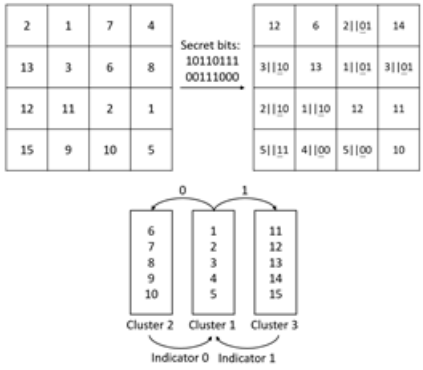


Figure 8: The example of the embedding

3.5 Qin and Hu's Scheme

In 2016, Qin and Hu's proposed a scheme based on an improved SOC method [32]. In ISOC, they make some difference after the SOC process is finished. If the index does not change to the SOC type, we compare it with the left or the upper index. Equation (1) shows the action of the index.

$$d = |X - X'|. \tag{1}$$

Users will define a threshold value T. While d is not greater than $T/2$, we change this index to the relative

addressing type (RA type). Table 1 shows the different length and indicator of each type.

Table 1: The summary of the ISOC

Index Type	Indictor	Encode Data	Coding Length
OIV	11	$(X)_{10}$	$2+\log_2 N$
SOC	0	$(m*)_{10}$	$1+\log_2 m$
RA	10	$(d)_{10}$	$2+\log_2 T$

X = search point value;
m = search order times;
d = difference calculate by Equation (1);
N = codebook size;
T = threshold defined by user;
m* = search order of the search point.

In Qin and Hu's scheme, after the ISOC process is finished, we embed each secret bit in each index. There are five situations in the embedding procedure.

First, if the index type is SOC type and the secret bit is 0, we do not change anything in this situation. Second, if the index type is SOC type and the secret bit is 1, we change it to the RA type; and this situation is classified to two cases. In case 1 we calculate the value d by Equation (1).

If d is less or equal than $T/2-1$, the encode bitstream will be $2+\log_2 T$. In the other hand, if d is greater than $T/2-1$, the encode bitstream will be $2+\log_2 T+\log_2 m$. In Situation 3, if the index type is RA type and the secret bit is 1, we do not change anything in this situation. In Situation 4, if the index type is RA type and the secret bit is 0, we change the index type to SOC type and the encode bitstream is $1+\log_2 m+\log_2 T$. In Situation 5, if the index type is OIV type, we do not do anything in this index. The summary of the embedding phase is shown as Table 2.

4 Comparisons

In the comparison part, we do the experiment of the representative scheme. The same codebook and the same secret bitstream are used in the experiment. The image size of the experiment is 512×512 and the codebook size is 256. In the VQ compression phase, the each non-overlapping block is 4×4, so each index has 8 codewords. Equation (2) shows the compression rate how to be calculated.

$$\text{bit rate}=\frac{\text{codestream}}{(H\times W)} \tag{2}$$

where the *codestream* means the code length after the embedding process is finished. The H and W mean the height and weight of the original cover image, respectively. We show the experiment result of these 5 representative methods in Tables 3 and 4 to display the capacity and compression rate, respectively. We can discover that method 1 index usually embeds 1 bit only; if we embed more bits in the index, the compression rate of the compression file will be significantly increased.

5 Conclusions and Future Work

Because of the rising of the social media, the compressed techniques are always used in the process of uploading images. This effect leads to the data hiding embed secret data in the compressed image becoming more popular. In this paper, we sort out the basic requirement in the data hiding for compressed images and compare five representative methods with different hiding techniques.

For the future development, according to the capacity requirement. How to embed more secret data with the similar bit-rate is the challenging issue. After separation, the number of blocks will be less than the number of original images pixels. We think that we can embed the secret data from the codeword in the codebook, which is the future work.

Acknowledgment

This research was partially supported by the Ministry Of Science and Technology, Taiwan (ROC), under contract no.: MOST 104-2221-E-468-004 and MOST 105-2410-H-468-009.

References

[1] K. Bharanitharan, C. C. Chang, H. R. Yang, and Z. H. Wang, "Efficient pixel prediction algorithm for reversible data hiding," *International Journal of Network Security*, vol. 18, no. 4, pp. 750–757, 2016.

[2] C. C. Chang, C. Y. Lin, Y. P. Hsieh, "Data hiding for vector quantization images using mixed-base notation and dissimilar patterns without loss of fidelity", *Information Sciences*, vol. 201, pp. 70-79, 2012.

[3] C. C. Chang, G. M. Chen, M. H. Lin, "Information hiding based on search-order coding for VQ indices", *Pattern Recognition Letters*, vol. 25, pp. 1253-1261, 2004.

[4] W. J. Chen, W. T. Huang, "VQ indices compression and information hiding using hybrid lossless index coding", *Digital Signal Processing*, vol. 19, pp. 433-443, 2009.

[5] C. C. Chang, K. F. Hwang, M. S. Hwang, "A digital watermarking scheme using human visual effects", *Informatics*, vol. 24, no. 4, pp. 505–511, Dec. 2000.

[6] C. C. Chang, T. D. Kieu, Y. C. Chou, "Reversible information hiding for VQ indices based on locally adaptive coding", *Journal of Visual Communication and Image Representation*, vol. 20, pp. 57-64, 2009.

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

Table 2: The summary of the Qin and Hu's scheme after the embedding phase finished

Original Index Type	Secret Bits	Indicator after Changed	Index Type after Changed	Coding Length after Changed
OIV	X	11	OIV	$2+\log_2 N$
SOC	0	0	SOC	$1+\log_2 m$
	1	10	RA	$2+\log_2 T$ (Case 1) $2+\log_2 T+\log_2 m$ (Case 2)
RA	0	0	SOC	$1+\log_2 T+\log_2 m$
	1	10	RA	$2+\log_2 T$

m = search order times;
N = codebook size;
T = threshold defined by user.

Table 3: The experiment result of the capacity comparison (Codebook size: 256; Block size: 4×4)

Schemes	**Trained Images**			**Non-trained Images**			
	Lena	Boat	Peppers	Mandrill	Barb	Goldhill	Zelda
Chang et al.'s Scheme [3]	16384	16384	16384	16384	16384	16384	16384
Chang et al.'s Scheme [11]	14262	12776	13792	11826	13199	14304	15799
Chang et al.'s Scheme [6]	16148	16137	16133	16139	16147	16153	16201
Tu et al.'s Scheme [34]	16384	16384	16384	16384	16384	16384	16384
Qin et al.'s Scheme [32]	9964	8776	9772	4577	8794	8804	10015

Table 4: The experiment result of the bit rate comparison (Codebook size: 256; Block size: 4×4)

Schemes	**Trained Images**			**Non-trained Images**			
	Lena	Boat	Peppers	Mandrill	Barb	Goldhill	Zelda
Chang et al.'s Scheme [3]	0.4941	0.5087	0.4941	0.5655	0.5102	0.5119	0.4922
Chang et al.'s Scheme [11]	0.5092	0.5246	0.5147	0.5199	0.5135	0.5071	0.5003
Chang et al.'s Scheme [6]	0.5552	0.5479	0.5606	0.5617	0.5617	0.5399	0.5536
Tu et al.'s Scheme [34]	0.5162	0.5275	0.5198	0.5348	0.5242	0.5159	0.5045
Qin et al.'s Scheme [32]	0.4641	0.4909	0.4721	0.5608	0.4914	0.4948	0.4695

[8] C. C. Chang, T. S. Nguyen, C. C. Lin, "A reversible data hiding scheme for VQ indices using locally adaptive coding", *Journal of Visual Communication and Image Representation*, vol. 22, pp. 664-672, 2011.

[9] C. C. Chang, T. S. Nguyen, C. C. Lin, "A novel VQ-based reversible data hiding scheme by using hybrid encoding strategies", *The Journal of Systems and Software*, vol. 86, pp. 389-402, 2013.

[10] C. C. Chang, W. L. Tai, C. C. Lin, "A reversible data hiding scheme based on side match vector quantization", *IEEE Transactions on Circuits and System for Video Technology*, vol. 16, no. 10, pp. 1301-1308, 2006.

[11] C. C. Chang, W. C. Wu, Y. C. Hu, "Lossless recovery of a VQ index table with embedded secret data", *Journal of Visual Communication and Image Representation*, vol. 18, pp. 207-216, 2007.

[12] S. F. Chiou, I-En Liao, and M. S. Hwang, "A capacity-enhanced reversible data hiding scheme based on SMVQ", *Imaging Science Journal*, vol. 59, no. 1, pp. 17–24, Feb. 2011.

[13] S. F. Chiou, Y. C. Lu, I-En Liao, and M. S. Hwang, "An efficient reversible data hiding scheme based on SMVQ," *Imaging Science Journal*, vol. 61, no. 6, pp. 467–474, 2013.

[14] R. M. Gray, "Vector Quantization", *IEEE ASSP Magazine*, vol. 1, no. 2, pp. 4-29, 1984.

[15] C. H. Hsieh, J. C. Tsai, "Lossless Compression of VQ index with search-order coding", *IEEE Transactions on Image Processing*, vol. 5, no. 11, pp. 1579-1582, 1996.

[16] L. C. Huang, L. Y. Tseng, M. S. Hwang, "A reversible data hiding method by histogram shifting in high quality medical images", *Journal of Systems and Software*, vol. 86, no. 3, pp. 716–727, Mar. 2013.

[17] M. S. Hwang, C. C. Chang, K. F. Hwang, "Digital watermarking of images using neural networks", *Journal of Electronic Imaging*, vol. 9, no. 4, pp. 548–555, Jan. 2000.

[18] B. Jana, "Dual image based reversible data hiding scheme using weighted matrix," *International Journal of Electronics and Information Engineering*, vol. 5, no. 1, pp. 6–19, 2016.

[19] B. Jana, D. Giri and S. K. Mondal, "Dual-image based reversible data hiding scheme using pixel value difference expansion," *International Journal of Network Security*, vol. 18, no. 4, pp. 633–643, 2016.

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

[21] T. D. Kieu, S. Ramroach, "A reversible steganographic scheme for VQ indices based on joint neighboring coding", *Expert Systems with Applications*, vol. 42, pp. 713-722, 2015.

[22] Y. Linde, A. Buzo, R. M. Gray, "An algorithm for vector quantizer design", *IEEE Transaction Communication*, vol. 28, no. 1, pp. 84-95, 1980.

[23] J. D. Lee, Y. H. Chiou, J. M. Guo, "Reversible data hiding based on histogram modification of SMVQ indices", *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 4, pp. 638-648, 2010.

[24] C. C. Lee, W. H. Ku, S. Y. Huang, "A new steganographic scheme based on vector quantisation and search-order coding", *IET Image Processing*, vol. 3, no. 4, pp.243-248, 2009.

[25] F. Li, Q. Mao, and C. C. Chang, "A reversible data hiding scheme based on IWT and the sudoku method," *International Journal of Network Security*, vol. 18, no. 3, pp. 410–419, 2016.

[26] C. C. Lin, X. L. Liu, S. M. Yuan, "Reversible data hiding for VQ-compressed images based on search-order coding and state-codebook mapping", *Information Science*, vol. 293, pp. 314-326, 2015

hiding scheme for VQindices based on modified locally adaptive coding and double-layer embedding strategy", *Journal of Visual Communication and Image Representation*, vol. 28, pp. 60-70, 2015.

[27] X. Ma, Z. Pan, S. Hu, L. Wang, "New high-performance reversible data hiding method for VQ indices based on improved locally adaptive coding scheme", *Journal of Visual Communication and Image Representation*, vol. 30, pp. 191-200, 2015.

[28] S. Manoharan, D. RajKumar, "Pixel value differencing method based on CMYK colour model," *International Journal of Electronics and Information Engineering*, vol. 5, no. 1, pp. 37–46, 2016.

[29] Z. Pan, X. Ma, X. Deng, S. Hu, "Low bit-rate information hiding method based on search-order-coding technique", *The Journal of Systems and Software*, vol. 86, pp. 2863-2869, 2013.

[30] C. Qin, C. C. Chang, Y. C. Chen, "Efficient reversible data hiding for VQ-compressed images based on index mapping mechanism", *Signal Processing*, vol. 93, pp. 2687-2695, 2013.

[31] C. Qin, C. C. Chang, Y. P. Chiu, "A novel joint data-hiding and compression scheme based on SMVQ and image inpainting", *IEEE Transactions on Image Processing*, vol. 23, no. 3, pp. 969-978, 2014.

[32] C. Qin, 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.

[33] S. C. Shie, J. H. Jiang, "Reversible and high-payload image steganographic scheme based on side-match vector quantization", *Signal Processing*, vol. 92, pp. 2332-2338, 2012

[34] T. Y. Tu, 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.

[35] W. J. Wang, C. T. Huang, C. M. Liu, P. C. Su, S. J. Wang, "Data embedding for vector quantization image processing on the basis of adjoining state-codebook mapping", *Information Science*, vol. 246, pp. 69-82, 2013.

[36] J. X. Wang, Z. M. Lu, "A path optional lossless data hiding scheme based on VQ joint neighboring coding", *Information Science*, vol. 179, pp. 3332-3348, 2009.

[37] L. Wang, Z. Pan, X. Ma, S. Hu, "A novel high-performance reversible data hiding scheme using SMVQ and improved locally adaptive coding method", *Journal of Visual Communication and Image Representation*, vol. 25, no. 2, pp. 454-465, 2014.

[38] Y. L. Wang, J. J. Shen, M. S. Hwang, "An improved dual image-based reversible hiding technique using LSB matching", *International Journal of Network Security*, vol. 19, no. 5, pp. 858–862, 2017.

[39] Z. H. Wang, X. Zhuang, C. C. Chang, C. Qin, Y. Zhu, "Reversible data hiding based on geometric structure of pixel groups", *International Journal of Network Security*, vol. 18, no. 1, pp. 52–59, 2016.

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

[41] C. H. Yang, Y. C. Lin, "Reversible data hiding of a VQ index table based on referred counts", *Journal of Visual Communication and Image Representation*, vol. 20, pp. 399-407, 2009.

[42] C. H. Yang, 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, pp. 334-342, 2010.

[43] S. Zhang, T. Gao, L. Yang, "A reversible data hiding scheme based on histogram modification in integer DWT domain for BTC compressed images," *International Journal of Network Security*, vol. 18, no. 4, pp. 718–727, 2016.

Biography

Yu-Lun Wang study in the Department of Management Information Systems, Chung Hsing University.

Jau-Ji Shen received his Ph.D. degree from National Taiwan University in 1988. His research interests include digital image, software engineering, information security, and data base technique. His work experiences include the Director of National Formosa University Library and the Associate Dean of Management School in Chaoyang University of Technology. Now, he is a professor in the Department of Management Information Systems, National Chung Hsing University.

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

Isogeny-based Quantum-resistant Undeniable Blind Signature Scheme

M. Seshadri Srinath, V. Chandrasekaran
(Corresponding author: M. Seshadri Srinath)

Department Mathematics and Computer Science, Sri Sathya Sai Institute of Higher Learning
Prasanthi Nilayam, Puttaparthi, Andhra Pradesh 515134, India
(Email: srinathms@sssihl.edu.in)
(Received Oct. 7, 2016; revised and accepted Feb. 20, 2017)

Abstract

In this paper, we propose an Undeniable Blind Signature scheme (UBSS) based on isogenies between supersingular elliptic curves. The proposed UBSS is an extension of the Jao-Soukharev undeniable signature scheme [16]. We formalize the notion of a UBSS by giving the formal definition. We then study its properties along with the pros and cons. Based on this, we provide a couple of its applications. We then state the isogeny problems in a more general form and discuss their computational hardnesses. Finally, we prove that the proposed scheme is secure in the presence of a quantum adversary under certain assumptions.

Keywords: Isogeny; Post-quantum Cryptography; Supersingular Elliptic Curve; Undeniable Blind Signature Scheme

1 Introduction

Blind signature scheme is a protocol in which the *requester* requests the *signer* to sign a document without disclosing the contents of the document. In 1982, Chaum [5] proposed the first blind signature scheme. It is based on the *RSA problem* [23]. Since then a host of blind signature schemes and their variations have been proposed based on different hardness assumptions such as the Discrete Logarithm Problem (DLP), pairing-based problems and lattice-based problems [4, 24, 29]. However, all the known blind signature schemes suffer from a common drawback that they are not secure in the presence of a quantum adversary. The blind signatures by Chaum [5], Camenisch et al. [4] and Zhang and Kim [29] are not quantum secure due to the polynomial time quantum algorithm by Shor for solving integer factorization and discrete logarithms. The lattice-based blind signature by Rückert [24] uses Fiat-Shamir paradigm [9] which is not secure in the quantum random oracle model as shown in [7].

Blind signature provides both anonymity and authentication [15, 20]. Hence it is used in the privacy-preserving protocols such as e-cash and e-voting [21, 22]. However, the signer has neither any control on the content of the document nor on the way the signature is used. Therefore, there is a crucial need to give a certain degree of control to the signer. One possible way is to let the signer and the requester agree on a part of the message (e.g., certain metadata about the specific message). This can be achieved through the technique introduced by Abe and Fujisaki [1].

Alternatively, one could let the signer decide who can verify the signature. This will keep unauthorized verifiers at bay and provide a certain control on the way the signature is used. The *Undeniable Signature* scheme introduced by Chaum and van Antwerpen [6] precisely has the said requirement. In an undeniable signature scheme the signer can decide who can verify the signature.

So, it seems desirable to have a scheme that would provide anonymity and controlled verification satisfying the properties of both blind signature and undeniable signature. Such a scheme can be devised but not obvious. In 1996, Sakurai and Yamane [25] have come up with an undeniable blind signature scheme based on the DLP. Their technique is also applicable for blinding the RSA based undeniable signature described in [6]. However, their scheme is not quantum secure either.

In this paper, we propose a new undeniable blind signature scheme based on the hardnesses of *isogeny problems* over supersingular elliptic curves. The isogeny problems for supersingular curves (details in Section 5) do not have any subexponential quantum algorithm. Hence, our scheme is quantum resistant.

Soukharev et al. [28] give a construction of quantum secure *designated verifier signature* scheme based on the hardness of isogeny problems. They also show a generic construction of asymmetric key authenticated encryption scheme. Jao and Soukharev [16] have proposed an isogeny-based undeniable signature. We extend Jao-Soukharev scheme into an Undeniable Blind Signa-

ture scheme.

To sum up, the main contributions of this paper are:

1) The concept of an UBSS seems to have been first mentioned in the work of Sakurai and Yamane [25]. However, to the best of our knowledge, it has never been formally defined in the literature till date. In this paper, we make such an attempt and give a formal definition of UBSS. We also study its properties including its strengths and weaknesses.

2) In [17], Jao and Venkatesan, speculate the use of hardness assumptions related to isogeny problems in constructing blind signature. We confirm this speculation by constructing an undeniable blind signature scheme.

3) The existing isogeny-based schemes [8, 16], including the current work, use primes of special forms that depend on a given set of small primes. Therefore, we state isogeny problems in their general form. These definitions can be used for the construction of any isogeny-based cryptographic scheme.

The rest of the paper is organized as follows. In Section 2, a formal definition of a UBSS is given and its properties as well as the possible attacks are studied. In Section 3 a brief and relevant mathematical background about isogenies between supersingular elliptic curves is provided. Section 4 describes the proposed UBSS in detail. In Section 5, we state the isogeny problems in their general form and discuss related hardness assumptions. The security of the proposed scheme is proved in Section 6. We conclude in Section 7.

2 Undeniable Blind Signature: Definition and Properties

2.1 Formal Definition

One would expect that a UBSS combines the properties of undeniable signature scheme and blind signature scheme. This means that UBSS would offer anonymity of the message origination and controlled verification of the signature. We have not found any definition that would fulfill both the requirements. Hence, we provide a definition for UBSS.

Definition 1 (Undeniable Blind Signature Scheme)**.** *An interactive signature scheme given by the tuple*

$$\mathcal{UBSS} = (\mathsf{KeyGen}, \mathsf{Blind}, \mathsf{Sign}, \mathsf{Unblind}, \mathsf{Check}, \mathcal{CON}, \mathcal{DIS})$$

is said to be undeniable blind signature scheme if it satisfies the following:

1) The randomized *key generation algorithm* KeyGen takes as input a security parameter 1^λ and outputs a pair of keys (vk, sk) which are called the *verification key* and the *secret key* respectively. This is written as $(vk, sk) \leftarrow \mathsf{KeyGen}(1^\lambda)$.

2) The randomized *blinding algorithm* Blind takes as input a message m and outputs a blinded message m', denoted as $m' \leftarrow {}_r\mathsf{Blind}(m)$ where r is the random choice made by the algorithm.

3) The randomized or deterministic *signing algorithm* Sign takes as input a secret key sk and a message m. It outputs a signature σ, denoted $\sigma \leftarrow \mathsf{Sign}_{sk}(m)$.

4) The deterministic *unblinding algorithm* $\mathsf{Unblind}$ takes as input a blinded signature σ' and a random choice r. It outputs an unblinded signature σ, to be denoted by $\sigma := \mathsf{Unblind}_r(\sigma')$.

5) The deterministic *checking algorithm* Check takes as input a message m, a signature σ and the key pair (vk, sk). It outputs a bit b with $b = 1$ meaning *valid* and $b = 0$ meaning *invalid*. This is written as $b := \mathsf{Check}_{(vk,sk)}(m, \sigma)$.

6) The *confirmation protocol*, π_{con} initiated by the signer, assures the verifier that the signature is indeed valid.

7) The *disavowal protocol*, π_{dis} also initiated by the signer, assures the verifier that the signature is not valid.

It is required that, for every key pair (vk, sk) output by $\mathsf{KeyGen}(1^\lambda)$, every m in the message space, and every random choice r made by Blind, the following holds:

$$\mathsf{Check}_{(vk,sk)}(m, \mathsf{Unblind}_r(\mathsf{Sign}_{sk}({}_r\mathsf{Blind}(m)))) = 1$$

Additionally, if the signature algorithm is deterministic, we may also assume that the effect of *blinding-signing-unblinding* on a message is same as directly signing the message. In the above notation, this means

$$\mathsf{Unblind}_r(\mathsf{Sign}_{sk}({}_r\mathsf{Blind}(m))) = \mathsf{Sign}_{sk}(m).$$

2.2 Working of UBSS

We will now run through the protocol to illustrate the role of the different algorithms in the definition. The illustration also makes it clear when these algorithms are run and by whom.

At first the signer chooses a security parameter λ and runs $\mathsf{KeyGen}(1^\lambda)$ to obtain the key pair (vk, sk). The signing key sk is kept secret and the verification key vk is published by the signer. Let m be the message which the requester wishes to communicate anonymously. The requester first creates a blinded version m' of m by running the algorithm $\mathsf{Blind}(m)$. Let r be the random choice made by the algorithm Blind. The requester then sends m' along with his identity Id_R. The signer verifies the requester's identity (see Remark 1) and runs Sign_{sk} on m' to obtain the blinded signature σ'. After receiving σ' from the signer, the requester unblinds it by using the algorithm $\mathsf{Unblind}$ and the same random choice r made by

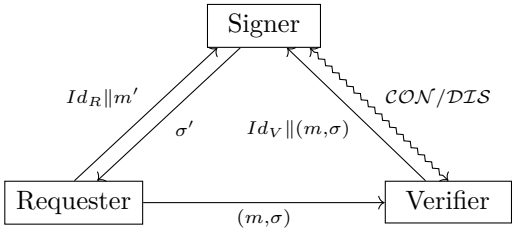


Figure 1: Illustration of the flow of information in an undeniable blind signature protocol

Blind. The requester then sends the original message m and the unblinded signature σ to the concerned party.

Any party who wishes to verify the signature sends the message-signature pair (m, σ) along with his identity Id_V to the signer. The signer verifies the identity of the verifier (see Remark 1). If Id_V is not the identity of an authorized verifier, then the signer simply ignores; otherwise, runs the algorithm Check. If Check returns *valid* then the signer initiates the confirmation protocol $\mathcal{CON}$; otherwise initiates the disavowal protocol $\mathcal{DIS}$. Figure 1 gives the flow of information in the UBSS.

Remark 1. *We intentionally do not specify how the signer verifies the identity of the requester and the verifier. It is the problem that can be best dealt with* mutual authentication *which can be done in one of the many ways [3, 11], all of which are* quantum secure.

2.3 Properties

The UBSS is desired to have the following three security properties viz., *unforgeability*, *blindness* and *invisibility*. The above properties are elaborated and their formal definitions are given below.

Unforgeability. As with any signature scheme, we require that the UBSS is unforgeable. The strongest notion of unforgeability is obtained when the adversary is allowed to corrupt both the requester and the verifier. The strongest notion of unforgeability for a UBSS is given here. The UBSS must be unforgeable against one-more forgery i.e., a requester who has received signatures for t messages (where t is polynomially bound by the security parameter), should not be able to output $t + 1$ distinct message-signature pairs even after collaborating with the verifier. This notion of unforgeability is formalized by the following security game:

1) The challenger runs $\mathsf{KeyGen}(1^\lambda)$ to obtain the key pair (vk, pk) and gives the verification key vk to $\mathcal{A}$.

2) $\mathcal{A}$ is allowed to make polynomially many queries to the signing oracle on chosen messages or any of their blinded versions adaptively and arbitrarily interleaved.

3) $\mathcal{A}$ is also allowed to submit message-signature pairs (m, σ) to the confirmation/disavowal oracle. If (m, σ) is valid (resp. invalid), then the oracle engages in confirmation (resp. disavowal) protocol with the adversary.

4) After making t queries to the signing oracle, $\mathcal{A}$ outputs t' distinct pairs (m_i, σ_i) such that

$$\mathsf{Check}_{(vk,sk)}(m_i, \sigma_i) = 1$$

Definition 2 (Unforgeability)**.** *Let $\mathcal{UBSS}$ be a given undeniable blind signature scheme as in Definition 1. We say that the $\mathcal{UBSS}$ is* unforgeable *if $Pr[t' > t]$ is negligible for any probabilistic polynomial-time (PPT) adversary $\mathcal{A}$ in the above game.*

Blindness. The blindness property is essential for preserving the anonymity of the message content originator. The signer should not be able to relate the message-signature pair and associated blinded versions. The strongest notion of blindness is obtained when the adversary is allowed to corrupt both the signer and verifier. Since the verification happens collaboratively with the signer, we allow the signer to view the signature after unblinding it. Incidentally, the existing definition of blindness for blind signature already accounts for this. Excepting notation, we consider the following security game as described by Schröder and Unruh in [26, Sec. 3 Defn. 4].

1) The adversary $\mathcal{A}$ runs $\mathsf{KeyGen}(1^\lambda)$ and generates a key pair (vk, sk).

2) $\mathcal{A}$ then chooses two messages m_0 and m_1 and gives them to the challenger.

3) The challenger chooses a random bit b hidden from $\mathcal{A}$ and reorders the messages as (m_b, m_{b-1}).

4) The challenger then blinds the two messages; $m'_b \leftarrow_{r_1} \mathsf{Blind}(m_b)$ and $m'_{b-1} \leftarrow_{r_2} \mathsf{Blind}(m_{b-1})$.

5) $\mathcal{A}$ engages in signing the blinded versions m'_b and m'_{b-1}. If signing requires multiple interactions, then $\mathcal{A}$ may engage parallely and arbitrarily interleaved.

6) The challenger receives the blinded signatures σ'_b and σ'_{b-1} and unblinds them; $\sigma_b := \mathsf{Unblind}_{r_1}(\sigma'_b)$ and $\sigma_{b-1} := \mathsf{Unblind}_{r_2}(\sigma'_{b-1})$.

7) The challenger then sends σ_b and σ_{b-1} to $\mathcal{A}$.

8) At the end of the attack game, $\mathcal{A}$ outputs a guess bit b'.

Definition 3 (Blindness)**.** *We say that the $\mathcal{UBSS}$ has* blindness *property if $|Pr[b' = b] - 1/2|$ is negligible for any PPT adversary $\mathcal{A}$ in the above game.*

Invisibility. A verifier should be able to accept (or reject) a signature only with the signer's cooperation via the confirmation (or disavowal) protocol and not otherwise. This notion is formalized by the following security game between a challenger $\mathcal{C}$ and an adversary $\mathcal{A}$.

1) The challenger runs $\mathsf{KeyGen}(1^\lambda)$ to obtain the key pair (vk, pk) and gives the verification key vk to $\mathcal{A}$.

2) $\mathcal{A}$ is permitted to issue a series of signing queries for messages m_i and their blinded versions to the signing oracle adaptively and receives signatures σ_i.

3) $\mathcal{A}$ is also allowed to submit message-signature pairs (m, σ) to the confirmation/disavowal oracle. If (m, σ) is valid (resp. invalid), then the oracle engages in confirmation (resp. disavowal) protocol with the adversary.

4) At some point, $\mathcal{A}$ chooses a message m^* and sends it to the challenger.

5) $\mathcal{C}$ chooses a random bit b. If $b = 1$, C runs $\sigma^* \leftarrow \mathsf{Sign}_{sk}(m^*)$, otherwise, C chooses a random value for σ^* from the signature space. $\mathcal{C}$ returns σ^* to $\mathcal{A}$.

6) $\mathcal{A}$ performs some signing queries again (see Remark 2).

7) $\mathcal{A}$ can also perform some queries to the confirmation/disavowal oracle but not allowed to query the challenge (m^*, σ^*).

8) At the end of the attack game, $\mathcal{A}$ outputs a guess bit b'.

Definition 4 (Invisibility)**.** *We say that the $\mathcal{UBSS}$ is* invisible *against full attack if $|Pr[b' = b] - 1/2|$ is negligible for any PPT adversary $\mathcal{A}$ in the above game.*

Remark 2. *If the signing algorithm is deterministic, we do not allow the adversary $\mathcal{A}$ to query m^* or any of its blinded versions to the signing oracle.*

2.4 Attacks: Blindness vs. Invisibility

A couple of attacks which exploit *blindness* property and *invisibility* property are demonstrated here. We show that all the existing schemes [13, 19] that combine these two requirements are vulnerable to the following attacks. At the end of the section, some suggestions to choose the appropriate model and suitable application are made in order that the system is secure.

The restriction in Remark 2 is a standard practice. However it seems rather forced. Suppose that the signing algorithm is deterministic and adversary $\mathcal{A}$ queries for a signature on a blinded version of m^*. If the UBSS is blind, then it is impossible for the signer to distinguish m^* from any of the previously signed messages. Hence, $\mathcal{A}$ can easily guess b and the signature is visible for the requester without actually engaging in the confirmation/disavowal protocol.

Suppose the signer does not conform to his inputs, say a different key pair (vk^*, sk^*) is used instead of (vk, sk) for signing all the messages form a particular requester. If the UBSS is invisible, it is impossible for the requester to know that the signer has used a different key pair. During the verification of a message-signature pair (m, σ), if $\mathsf{Check}_{(vk,sk)}(m, \sigma)$ returns *invalid*, and $\mathsf{Check}_{(vk^*,sk^*)}(m, \sigma)$ returns *valid*, then the signer can trace the origin of the message m. Thus, compromising the anonymity of the content originator. The signer seamlessly continues with the disavowal protocol. This anomaly could be seen as an advantage. Suppose the requester becomes aware that the signer has used a different key pair for signing. The requester may choose to give up the anonymity of the message to expose the signer. The signatures can be used as an evidence against the signer.

One way to circumvent the above attacks is to allow the requester to be a valid verifier. This makes the signatures visible to the requester and empowers the requester to check whether the signer has used the correct input.

The definition of UBSS is decoupled from the actual security model and the applications. While anonymity and invisibility appear to be conflicting goals, by choosing an appropriate security model, UBSS can be very useful in certain applications. For example, in the case of e-cash, one may consider the bank as a semi-honest signer. For security reasons, the bank could decide to verify signatures only for its customers. Then the bank should use UBSS instead of blind signatures.

Another example where the UBSS becomes a natural choice is *Anonymous Feedback System*. Suppose the chief organizer of an event wishes to take anonymous feedback from the participants. It should be done such that (i) only the participants should be able to give the feedback anonymously and (ii) only the organizing committee should be able to verify the authenticity of the feedback. The participants who give feedback request for a blind signature from the chief organizer. After obtaining the signature, the participants send the feedback along with the signature to the organizing committee. The committee members then verify the signature with the chief organizer. *E-voting* can be considered as a special case of anonymous feedback system. In ANONIZE [12], any adversary can verify the authenticity of the survey data and hence there is a possiblity of misuse of the data. UBSS averts any such misuse by allowing only the authorized verifiers to check the authenticity of the data.

This completes our discussion on the definition of the UBSS. In the next few sections, we give an example of a UBSS using the isogeny-based hardness assumptions.

3 Mathematical Background

This section briefly provides some necessary mathematical background. For further details, the reader is referred to [27] for mathematical, [10] for cryptographic, [8] for algorithmic aspects and the citations thereof.

Let $\mathbb{F}_q$ be the finite field (up to isomorphism) of characteristic p and cardinality q. It is a well known fact that two elliptic curves are isomorphic over an algebraic closure

of $\mathbb{F}_q$ if and only if they have the same j-invariant. Also, given two elliptic curves, the isomorphism between them can be efficiently computed. An elliptic curve $E/\mathbb{F}_q$ is said to be *supersingular* if $\#E(\mathbb{F}_q) \equiv 1 \mod p$. For equivalent definitions kindly refer [14, Ch. 13 Sec. 3 p. 259].

Isogenies. A homomorphism between two groups is a map that preserves the group structure. The *kernel* of a homomorphism is the subset of elements whose image is the identity. An *isogeny* is a group homomorphism between two elliptic curves with a *finite* kernel. Let $\phi : E_1 \to E_2$ be an isogeny between two elliptic curves E_1 and E_2. Thus $\phi(O_{E_1}) = O_{E_2}$ and ϕ can be written as

$$\phi(x, y) = \left(\frac{f_1(x, y)}{g_1(x, y)}, \frac{f_2(x, y)}{g_2(x, y)} \right),$$

where f_1, f_2, g_1, g_2 are polynomials in two variables x, y with co-efficients in $\mathbb{F}_q$. The *degree* of the isogeny ϕ, $\deg \phi = \max\{\deg f_1, \deg f_2\}$. An isogeny ϕ is said to be *separable* if $\deg \phi = \#\ker \phi$. An isogeny of degree ℓ is often referred to as an ℓ-isogeny. For any ℓ-isogeny $\phi : E_1 \to E_2$, there exists an ℓ-isogeny $\hat{\phi} : E_2 \to E_1$, called the *dual* of ϕ, such that $\phi \circ \hat{\phi} = \hat{\phi} \circ \phi = [\ell]$ where $[\ell]$ is a multiplication-by-ℓ map. Two elliptic curves E_1 and E_2 are said to be ℓ-isogenous if there exists an ℓ-isogeny ϕ between them. Tate's isogeny theorem says that E_1 and E_2 are isogenous over $\mathbb{F}_q$ if and only if $\#E_1(\mathbb{F}_q) = \#E_2(\mathbb{F}_q)$. An isogeny is uniquely identified (up to isomorphism) by its kernel. Any generator of the kernel will produce a unique isogeny up to isomorphism via Vélu's formulae. In our work we will be considering only supersingular elliptic curves and separable isogenies with cyclic kernels.

Isogeny Graph. An ℓ-isogeny graph is a graph in which the nodes are represented by isomorphism classes of elliptic curves. There is an edge from E_1 to E_2 in the ℓ-isogeny graph if there is an ℓ-isogeny form E_1 to E_2. The isogeny graph is undirected due to the existence of dual isogenies. The ℓ-isogeny graph of supersingular curves is connected. Given two random nodes in the isogeny graph finding a path of fixed length is hard. This hardness is used for constructing isogeny-based cryptosystems, explained in detail in Section 5.

4 A New Undeniable Blind Signature Scheme Based on Isogenies

In this section, we describe a new undeniable blind signature scheme based computing an isogeny between two supersingular elliptic curves over a finite field $\mathbb{F}_q$. We borrow the notation as in the paper of Jao and Soukharev [16].

4.1 Public Parameters

Choose a prime p of the form $p = \ell_A^{e_A} \ell_M^{e_M} \ell_C^{e_C} \ell_R^{e_R} \cdot f \pm 1$. Generate a random supersingular elliptic curve E_0 defined over the field $\mathbb{F}_{p^2}$. Choose base points $\{P_A, Q_A\}$, $\{P_M, Q_M\}$, $\{P_C, Q_C\}$ and $\{P_R, Q_R\}$ that generate $E_0[\ell_A^{e_A}]$, $E_0[\ell_M^{e_M}]$, $E_0[\ell_C^{e_C}]$ and $E_0[\ell_R^{e_R}]$ respectively. Choose a hash function $H : \{0,1\}^* \to \frac{\mathbb{Z}}{\ell_M^{e_M}\mathbb{Z}}$.

4.2 KeyGen

The signer generates two random numbers $m_A, n_A \in \mathbb{Z}/\ell_A^{e_A}\mathbb{Z}$. Computes the curve $E_A = E_0/\langle K_A \rangle$ where $K_A = [m_A]P_A + [n_A]Q_A$ is the generator of the kernel of the isogeny $\phi_A : E_0 \to E_A$. The signer also computes $\phi_A(P_C)$ and $\phi_A(Q_C)$.

Public Key: $E_A, \phi_A(P_C), \phi_A(Q_C)$

Private Key: m_A, n_A, K_A

$$E_0 \xrightarrow{\phi_A} E_A$$

Figure 2: The isogeny ϕ_A computed during the key generation phase

4.3 Blind

Let M be the message for which the signature is required. Let $h = H(M)$. Compute the isogeny ϕ_M and the curve

$$E_M = \frac{E_0}{\langle P_M + [h]Q_M \rangle}$$

The image points $\phi_M(P_A)$, $\phi_M(Q_A)$, $\phi_M(P_C)$, $\phi_M(Q_C)$, $\phi_M(P_R)$ and $\phi_M(Q_R)$ are also computed. Now this message curve E_M has to be blinded. Choose a random $r \in \frac{\mathbb{Z}}{\ell_R^{e_R}\mathbb{Z}}$ which is hidden from the signer. Compute the isogeny $\phi_{M,RM}$ and the curve

$$E_{RM} = \frac{E_M}{\langle \phi_M(P_R) + [r]\phi_M(Q_R) \rangle}$$

E_{RM} is the blinded curve on which the signer will sign. The blinding process is illustrated in Figure 3.

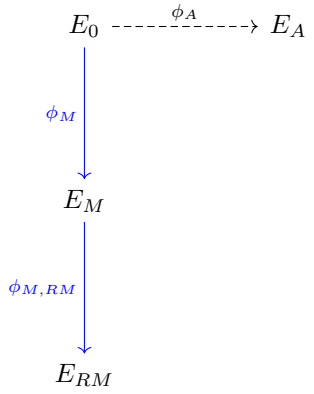


Figure 3: The isogenies ϕ_M and $\phi_{M,RM}$ computed while blinding the message. The dashed arrow is the isogeny unknown to the requester

Before sending the curve E_{RM} for signing, one has to compute the dual isogeny $\hat{\phi}_{M,RM}$, so that unblinding is possible. To do that, first we need to find a point $K \in E_M[\ell_R^{e_R}]$ of order $\ell_R^{e_R}$ such that $K \notin Ker\phi_{M,RM}$, say $K = \phi_M(Q_R)$. Compute the image point $\phi_{M,RM}(K) \in E_{RM}$. The isogeny with kernel $\langle \phi_{M,RM}(K) \rangle$ is the dual isogeny $\hat{\phi}_{M,RM}$.

Remark 3. *Strictly speaking, this will not be the dual of $\phi_{M,RM}$ because this isogeny will lead to a curve which is isomorphic to E_M. Since isomorphic curves represent the same node in the isogeny graph, this isogeny maps back to the same node. By the abuse of notation, we denote it as $\hat{\phi}_{M,RM}$.*

Now, choose basis $\{P'_R, Q'_R\} \in E_{RM}$ that generate $E_{RM}[\ell_R^{e_R}]$. Compute $m, n \in \frac{\mathbb{Z}}{\ell_R^{e_R}}$ such that

$$\phi_{M,RM}(K) = [m]P'_R + [n]Q'_R$$

This amounts to solving extended discrete logarithm problem on E_{RM}. Since E_{RM} is isogenous to E_0, by Tate's theorem, we have

$$\#E_{RM}(\mathbb{F}_{p^2}) = \#E_0(\mathbb{F}_{p^2})$$

Hence E_{RM} is a curve of smooth order. Therefore, m, n can be found efficiently using generalized Pohlig-Hellman algorithm. The masked curve E_{RM} along with the points

$$P'_A = \phi_{M,RM}(\phi_M(P_A))$$
$$Q'_A = \phi_{M,RM}(\phi_M(Q_A))$$
$$P'_C = \phi_{M,RM}(\phi_M(P_C))$$
$$Q'_C = \phi_{M,RM}(\phi_M(Q_C))$$

P'_R and Q'_R (all belonging to E_{RM}) is sent to the signer.

4.4 Sign

The signer computes the curve

$$E_{ARM} = \frac{E_{RM}}{\langle [m_A]P'_A + [n_A]Q'_A \rangle}$$

The signer also computes the image points $\phi_{RM,ARM}(P'_C)$, $\phi_{RM,ARM}(Q'_C)$, $\phi_{RM,ARM}(P'_R)$ and $\phi_{RM,ARM}(Q'_R)$, and sends all the computed values to the user.

4.5 Unblind

The requester computes the isogeny $\hat{\phi}_{AM,ARM}$ and the curve

$$E_{AM} = \frac{E_{ARM}}{\langle [m]\phi_{RM,ARM}(P'_R) + [n]\phi_{RM,ARM}(Q'_R) \rangle}$$

The requester also computes the points

$$P_S = \hat{\phi}_{AM,ARM}(\phi_{RM,ARM}(P'_C))$$

$$Q_S = \hat{\phi}_{AM,ARM}(\phi_{RM,ARM}(Q'_C))$$

The signature $\sigma = \{E_{AM}, P_S, Q_S\}$.

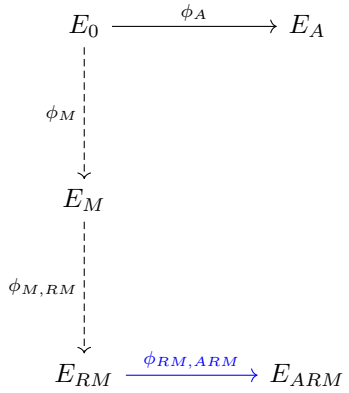


Figure 4: The isogeny $\phi_{RM,ARM}$ computed for signing the blinded message. The dashed arrows are the isogenies unknown to the signer.

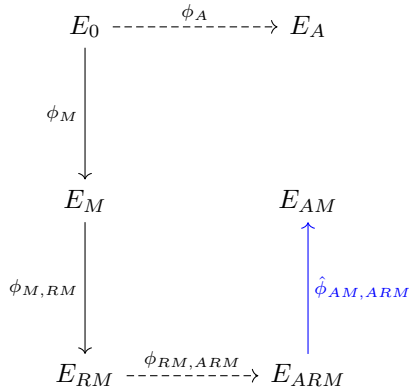


Figure 5: The isogeny $\hat{\phi}_{AM,ARM}$ computed while unblinding the signature. The dashed arrows are the isogenies unknown to the requester.

4.6 Check

At the end of Unblind algorithm, the signature curve generated by our scheme is isomorphic to Jao-Soukharev signature curve. Hence the signature verification can be done in the same way as in Jao-Soukharev signature. When a message M and signature σ is submitted for verification, the signer first checks whether the square (E_0, E_A, E_{AM}, E_M) in Figure 6 commutes. If it does, then the signer initiates the confirmation protocol $\mathcal{CON}$, initiates the disavowal protocol $\mathcal{DIS}$. The confirmation and disavowal protocols are same as in [16].

Remark 4. *Strictly speaking, the effect of* blinding-signing-unblinding *is not the same as directly signing the message. The action of an isogeny followed by the action of its dual is equivalent to multiplication-by-degree map [27, III.6.2a p. 83]. Hence, the points P_S and Q_S will have a factor of $\ell_R^{e_R}$ multiplied to them when compared to the Jao-Soukharev signature. But then, this factor is relatively prime to their order $\ell_C^{e_C}$. It would not affect the signature verification since both the pairs generate the*

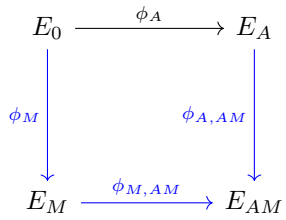


Figure 6: The isogenies ϕ_M, $\phi_{M,AM}$, $\phi_{A,AM}$ are computed to check whether the given signature E_{AM} is valid.

same kernel.

The prime used in our work is different form the primes already used in the literature [8, 16] for constructing isogeny-based cryptographic primitives. This motivates us to give generalized statements and hardness assumptions for isogeny-problems. We review them in the next section.

5 Isogeny Problems Revisited

The current work uses the prime p of the form $p = \ell_A^{e_A}\ell_M^{e_M}\ell_C^{e_C}\ell_R^{e_R} \cdot f \pm 1$ which has not been used so far in the literature. The security of the isogeny-based schemes depend on the size of the corresponding torsion subgroup. Hence, such a choice for the prime does not have any security implications so long as the torsion groups are large enough.

Let p be a prime of the form $p = f \cdot \prod_{i=1}^{N} \ell_i^{e_i} \pm 1$ where ℓ_i are distinct small primes, e_i are positive integers and $f \geq 1$ is a small cofactor. Let E_0 be a supersingular elliptic curve defined over $\mathbb{F}_{p^2}$ and having order $(p \mp 1)^2$. For each $1 \leq i \leq N$, let $\{P_i, Q_i\}$ be an arbitrarily chosen basis of $E_0[\ell_i^{e_i}]$. The above information forms the global parameters.

Problem 1 (Decisional Supersingular Isogeny (DSSI) problem)**.** *Given the global parameters and another curve E' defined over $\mathbb{F}_{p^2}$ such that $\#E_0(\mathbb{F}_{p^2}) = \#E'(\mathbb{F}_{p^2})$, decide whether E' is $\ell_i^{e_i}$-isogenous to E_0 for a specified $1 \leq i \leq N$.*

For a fixed but arbitrary $1 \leq i \leq N$, let $\phi_i : E_0 \to E_i$ be an isogeny whose kernel is $\langle [m_i]P_i + [n_i]Q_i \rangle$ where $m_i, n_i \in \mathbb{Z}/\ell_i^{e_i}\mathbb{Z}$ are chosen randomly and not both divisible by ℓ_i.

Problem 2 (Computational Supersingular Isogeny (CSSI) problem)**.** *Given the global parameters, the curve E_i and the points $\phi_i(P_j), \phi_i(Q_j)$ for all $j = 1, 2, \cdots, N$, $j \neq i$, find a generator of $\langle [m_i]P_i + [n_i]Q_i \rangle$.*

5.1 DSSI and CSSI Assumptions

The DSSI and CSSI assumptions are the assumptions that DSSI and CSSI problems are hard to solve for any $1 \leq i \leq N$. This notion is formalized in this section.

DSSI Assumption. The DSSI assumption says that the following two probability distributions are *computationally indistinguishable* for all i:

- $(E, E/\langle R \rangle)$ where $R \in E$ is a random point of order $\ell_i^{e_i}$.
- (E, E') where $E'/\mathbb{F}_{p^2}$ is a random curve such that $\#E(\mathbb{F}_{p^2}) = \#E'(\mathbb{F}_{p^2})$.

Let λ be the security parameter. Let $\mathcal{G}$ be a (possibly randomized) polynomial-time algorithm that, on input 1^λ, outputs the global parameters described above. Let us denote the set of all the global parameters by $\mathbb{G}$.

Definition 5. *We say that the DSSI problem is* hard *relative to $\mathcal{G}$ if $\forall\ 1 \leq i \leq N$ and for all bounded quantum polynomial-time algorithms $\mathcal{A}$, the quantity*

$$|Pr[\mathcal{A}(\mathbb{G}, E, E/\langle R \rangle) = 1] - Pr[\mathcal{A}(\mathbb{G}, E, E') = 1]|$$

is negligible and the probabilities in each case is taken over the experiment in which $\mathcal{G}(1^\lambda)$ outputs $\mathbb{G}$, $R \in E$ is a random point of order $\ell_i^{e_i}$ and E' is a random curve such that $\#E(\mathbb{F}_{p^2}) = \#E'(\mathbb{F}_{p^2})$.

CSSI Assumption. Consider the following experiment for a given parameter-generating algorithm $\mathcal{G}$, algorithm $\mathcal{A}$, and parameter λ.

The computational supersingular isogeny experiment $\mathsf{CSSIso}_{\mathcal{A},\mathcal{G}}(\lambda)$:

1) Run $\mathcal{G}(1^\lambda)$ to obtain the global parameters $\mathbb{G} = (p, E_0, \ell_i, e_i, P_i, Q_i)$.
2) For a fixed $1 \leq i \leq N$, choose $m, n \leftarrow \mathbb{Z}/\ell_i^{e_i}$ not both divisible by ℓ_i and compute
$$E' \equiv \frac{E_0}{\langle [m]P_i + [n]Q_i \rangle}$$
3) $\mathcal{A}$ is given $\mathbb{G}, i, E'$ and outputs a point $R \in E_0$.
4) The output of the experiment is defined to be 1 if $E' \equiv \frac{E_0}{\langle R \rangle}$ and 0 otherwise.

Definition 6. *We say that the CSSI problem is* hard *relative to $\mathcal{G}$ if $\forall\ 1 \leq i \leq N$ and for all bounded quantum polynomial-time algorithms $\mathcal{A}$ there exists a negligible function* negl *such that*

$$Pr[\mathsf{CSSIso}_{\mathcal{A},\mathcal{G}}(\lambda) = 1] \leq \mathsf{negl}(\lambda).$$

5.2 Hardness of CSSI and DSSI Assumptions

Since the DSSI and CSSI problems need to be hard for all values of i, it is expected that the parameter generating algorithm $\mathcal{G}$ outputs the prime p such that the values $\ell_i^{e_i}$ are roughly of the same size for all i. Hence, we assume $\ell_i^{e_i} \approx \sqrt[N]{p}$. The generic attack for solving DSSI and CSSI

problems that improve on exhaustive search involve solving the *claw problem* for the domain size $\ell_i^{e_i/2}$. The optimal complexity for the above black-box claw attack using a quantum computer is $O(\ell_i^{e_i/3}) = O(\sqrt[3N]{p})$. Suppose $\lambda = \log p$, then the complexity of the attack is $O(2^{\lambda/3N})$ which is clearly exponential in λ. Kohel et al. [18] have given a probabilistic algorithm for solving the quaternion analog of CSSI problem. However, translating it to CSSI problem is not known to be efficient. The quantum algorithm by Biasse et al. [2] yields a subexponential attack if the base curve is defined over $\mathbb{F}_p$. There is no known subexponential attack if the base curve is not defined over $\mathbb{F}_p$.

5.3 Other Isogeny Problems

There have been several other variants of DSSI and CSSI problems whose hardness have been assumed to build the cryptographic primitives. We present only those that are relevant to the current work. For a complete list, we refer the reader to [16, Sec. 5]. Henceforth in the rest of the paper, for the sake of simplicity, we follow the notation as in Section 4.

Problem 3 (Decisional Supersingular Product (DSSP) problem)**.** *Given an isogeny $\phi : E_0 \to E_3$ of degree $\ell_i^{e_i}$ and a tuple sampled with probability $1/2$ from one of the following two distributions:*

- *(E_1, E_2, ϕ') where the product $E_1 \times E_2$ is chosen at random among those $\ell_j^{e_j}$-isogenous ($i \neq j$) to $E_0 \times E_3$, and where $\phi' : E_1 \to E_2$ is an isogeny of degree $\ell_i^{e_i}$, and*
- *(E_1, E_2, ϕ') where E_1 is chosen at random among the curves having the same cardinality as E_0, and $\phi' : E_1 \to E_2$ is a random isogeny of degree $\ell_i^{e_i}$,*

determine from which distribution the tuple is sampled.

Problem 4 (Modified Supersingular Computational Diffie-Hellman (MSSCDH) problem)**.** *Given E_A, E_M and $\ker(\phi_M)$, determine E_{AM}.*

Problem 5 (One-sided Modified Supersingular Computational Diffie-Hellman (q-OMSSCDH) problem)**.** *For a fixed E_A and given oracle access of at most q times to MSSCDH for any set of inputs E_A, E_{M_i}, $\ker(\phi_{M_i})$, $(1 \leq i \leq q)$. Solve MSSCDH for E_A, E_M and $\ker(\phi_M)$ where $E_M \not\equiv E_{M_i}$ $\forall i$.*

Problem 6 (Modified Supersingular Decisional Diffie-Hellman (MSSDDH) problem)**.** *Given E_A, E_M, E_C and $\ker(\phi_M)$, decide whether $E_C \equiv E_{AM}$.*

Problem 7 (One-sided Modified Supersingular Decisional Diffie-Hellman (q-OMSSDDH) problem)**.** *For a fixed E_A and given oracle access of at most q times to MSSCDH oracle for any set of inputs E_A, E_{M_i}, $\ker(\phi_{M_i})$, $(1 \leq i \leq q)$. Solve MSSDDH for E_A, E_M, E_C and $\ker(\phi_M)$ where $E_M \not\equiv E_{M_i}$ $\forall i$.*

Signing Oracle. Given any supersingular elliptic curve $\mathcal{E}/\mathbb{F}_{p^2}$ of order $(\ell_A^{e_A}\ell_M^{e_M}\ell_C^{e_C}\ell_R^{e_R})^2$ and points $P, Q \in \mathcal{E}$ both of order $\ell_A^{e_A}$, the signing oracle outputs the curve $\mathcal{E}_A$ such that

$$\mathcal{E}_A \equiv \frac{\mathcal{E}}{[m_A]P + [n_A]Q}$$

where $m_A, n_A \in \mathbb{Z}/\ell_A^{e_A}\mathbb{Z}$ form the private key.

Problem 8 (One-More Supersingular Computational Diffie-Hellman (1MSSCDH) problem)**.** *After making q queries to the signing oracle, output at least $q+1$ distinct pairs of curves $\{E_{M_i}, E_{AM_i}\}$ where E_{M_i} are $\ell_M^{e_M}$-isogenous to E_0 and $\{E_A, E_{M_i}, E_{AM_i}\}$ is a Diffie-Hellman tuple for each $1 \leq i \leq t$.*

6 Security of the Proposed Construction

In this section, we prove that our UBSS has unforgeability, blindness and invisibility.

6.1 Unforgeability

The challenger chooses a security parameter and generates the secret key m_A, n_A. The corresponding public key $E_A, \phi_A(P_C), \phi_A(Q_C)$ is given to the adversary $\mathcal{A}$. $\mathcal{A}$ then issues a series of at most q signing queries to the challenger for the messages m_i $(1 \leq i \leq q)$. Let E_{M_i} and E_{AM_i} be the corresponding message curves and signatures respectively. $\mathcal{A}$ is allowed to submit the message-signature pairs (m, E_{AM}) to the signer for verification. If the signature is correct then the signer engages in confirmation protocol otherwise initiates disavowal protocol. At some point adversary then outputs q' message-signature pairs (m_j, E_{AM_j}). The adversary wins the game if $q' > q$.

Theorem 1 (Unforgeability)**.** *If the DSSP and 1MSSCDH assumptions hold, then the proposed UBSS is* unforgeable.

Proof. Suppose there exists an adversary $\mathcal{A}$ that forges the proposed UBSS. Without any loss of generality we may assume that $\mathcal{A}$ issued exactly q signing queries and output exactly $q+1$ valid message-signature pairs. The confirmation and disavowal protocols are shown to be zero-knowledge in [16, Sec. 7] provided DSSP is hard to solve. Hence we may further assume that $\mathcal{A}$ does not have access to the confirmation/disavowal oracle at all. But then $\mathcal{A}$ in turn solves 1MSSCDH problem. □

Remark 5. *Since the signature for a message m obtained at the end of the proposed UBSS protocol is the Jao-Soukharev signature for m, we also need to assume that solving q-OMSSCDH problem is hard. This is omitted in the statement of Theorem 1 as 1MSSCDH assumption is stronger than q-OMSSCDH assumption.*

6.2 Blindness

To prove that the proposed signature scheme has blindness property, the following security game is used. The adversary $\mathcal{A}$ is given the security parameter. $\mathcal{A}$ generates the secret key m_A, n_A and the corresponding public key $E_A, \phi_A(P_C), \phi_A(Q_C)$. The adversary outputs two messages $\{m_0, m_1\}$. The same two messages are ordered as $\{m_b, m_{1-b}\}$ according to a random bit b which is hidden from $\mathcal{A}$. Then $\mathcal{A}$ engages in two parallel interactive protocols, possibly with two different users. If the users output the corresponding signatures, then $\mathcal{A}$ is also given E_{AM_0} and E_{AM_1}. $\mathcal{A}$'s goal is to guess the value of the bit b and the blindness property requires that such a guess is negligibly close to $\frac{1}{2}$.

Theorem 2 (Blindness)**.** *If the DSSP is hard to solve, then the proposed UBSS has the* blindness *property.*

Proof. Given E_{M_0}, E_{M_1}, E_{RM_b}, $E_{RM_{1-b}}$, E_{AM_0}, E_{AM_1} the goal of the adversary $\mathcal{A}$ is to figure out the value of the bit b. Note that $\mathcal{A}$ also has the knowledge of the isogenies $\phi_{s_0} : E_{M_0} \to E_{AM_0}$, $\phi_{s_1} : E_{M_1} \to E_{AM_1}$, $\phi'_{s_b} : E_{RM_b} \to E_{ARM_b}$ and $\phi'_{s_{1-b}} : E_{RM_{1-b}} \to E_{ARM_{1-b}}$. To decide whether $b = 0$ or $b = 1$ is equivalent to deciding whether, $E_{RM_b} \times E_{ARM_b}$ is $\ell_R^{e_R}$-isogenous to $E_{M_0} \times E_{AM_0}$ or not. Further, this essentially amounts to solving DSSP on the inputs $(E_{M_0}, E_{AM_0}, \phi_{s_0})$ and $(E_{RM_b}, E_{ARM_b}, \phi'_{s_b})$. □

6.3 Invisibility

The challenger chooses a security parameter and generates the secret key m_A, n_A. The corresponding public key $E_A, \phi_A(P_C), \phi_A(Q_C)$ is given to the adversary $\mathcal{A}$. $\mathcal{A}$ then issues a series of at most q signing queries to the challenger for the messages m_i. Let E_{M_i} and E_{AM_i} be the corresponding message curves and signatures respectively. $\mathcal{A}$ is allowed to query E_{M_i} and any of its blinded versions to the signing oracle. $\mathcal{A}$ is also allowed to submit the message-signature pairs (m_j, E_{AM_j}) to the confirmation/disavowal protocols. At some point $\mathcal{A}$ outputs a message m^*. The challenger chooses a random bit b. If $b = 0$, the challenger replies with the correct signature E_{AM^*} otherwise chooses a random curve E_R with $\#E_R(\mathbb{F}_{p^2}) = \#E_0(\mathbb{F}_{p^2})$. According to the definition of invisibility, the message curve E_M and none of its blinded versions are allowed to query the signing oracle.

Theorem 3 (Invisibility)**.** *If the DSSP and q-OMSSDDH assumptions hold, then the proposed UBSS is* invisible*.*

Proof. If the DSSP assumption holds, then the confirmation and disavowal protocols are shown to be zero-knowledge [16, Sec. 7] in the presence of a quantum adversary. Hence we may assume that the adversary $\mathcal{A}$ does not have access to confirmation/disavowal oracle. Instead, the access is given to an oracle which on querying (m, E) outputs valid or invalid depending on whether E is a valid signature for m or not. Further, $\mathcal{A}$ is not allowed to query the signing oracle for the curve E_M or any of its blinded versions. Hence showing the invisibility of our signature scheme is equivalent to showing that the Jao-Soukharev signature is invisible. The reader may refer [16, Sec. 6] for the proof of invisibility. □

7 Conclusion

We give a formal definition of UBSS as well as modified definitions of blindness, invisibility and unforgeability; concepts that are key in defining UBSS. As we mentioned earlier, though the concept of UBSS is not new and has been mentioned in Sakurai and Yamane [25], this is the first time a formal definition has been given. We also show that blindness and invisibility play against each other. This affects the specifics of how UBSS can be used for the application at hand. We then described a new UBSS based on the isogeny problem for supersingular elliptic curves. We also give the generalized statements of isogeny problems. This makes it convenient for constructions of isogeny-based cryptographic primitives. We finally prove that our UBSS has the desired properties under the assumptions that DSSP, OMSSDDH and 1MSS-CDH are hard to solve.

Acknowledgments

The first author thanks Vijay M. Patankar for guidance, helpful discussions and a careful first reading of the manuscript. He also thanks David Jao for suggestions (i) to formalize the definition of UBSS (ii) to give a careful thought to its applications (iii) to state the isogeny problems in a general form and (iv) to improve the exposition. This work is supported by Indian Space Research Organization through Sponsored Research programme. The authors dedicate this work to the Founder Chancellor, Sri Sathya Sai Institute of Higher Learning.

References

[1] M. Abe and E. Fujisaki, "How to date blind signatures," in *International Conference on the Theory and Applications of Cryptology and Information Security: Advances in Cryptology*, pp. 244–251, Springer-Verlag, 1996.

[2] J. F. Biasse, D. Jao, and A. Sankar, "A quantum algorithm for computing isogenies between supersingular elliptic curves," in *Progress in Cryptology (INDOCRYPT'14)*, LNCS 8885, pp. 428–442, Springer, 2014.

[3] D. Boneh and M. Zhandry, "Secure signatures and chosen ciphertext security in a quantum computing world," in *Advances in Cryptology (CRYPTO'13)*, LNCS 8043, pp. 361–379, Springer Berlin Heidelberg, 2013.

[4] J. L. Camenisch, J. M. Piveteau, and M. A. Stadler, "Blind signatures based on the discrete logarithm problem," in *Advances in Cryptology (EUROCRYPT'94)*, LNCS 950, pp. 428–432, Springer Berlin Heidelberg, 1995.

[5] D. Chaum, "Blind signatures for untraceable payments," in *Advances in Cryptology*, pp. 199–203, Springer US, 1983.

[6] D. Chaum and H. van Antwerpen, "Undeniable signatures," in *Advances in Cryptology (CRYPTO'89)*, LNCS 435, pp. 212–216, Springer New York, 1990.

[7] Ö. Dagdelen, M. Fischlin, and T. Gagliardoni, "The Fiat-Shamir transformation in a quantum world," in *Advances in Cryptology (ASIACRYPT'13)*, LNCS 8270, pp. 62–81, Springer Berlin Heidelberg, 2013.

[8] L. De Feo, D. Jao, and J. Plut, "Towards quantum-resistant cryptosystems from supersingular elliptic curve isogenies," *Journal of Mathematical Cryptology*, vol. 8, pp. 209–247, June 2014.

[9] A. Fiat and A. Shamir, "How to prove yourself: Practical solutions to identification and signature problems," in *Proceedings on Advances in cryptology (CRYPTO'86)*, pp. 186–194, Springer-Verlag, 1987.

[10] S. Galbraith, *Mathematics of Public Key Cryptography*, Cambridge University Press, Apr. 2012.

[11] S. A. Goorden, M. Horstmann, A. P. Mosk, B. Škorić, and P. W. H. Pinkse, "Quantum-secure authentication of a physical unclonable key," *Optica*, vol. 1, pp. 421–424, Dec. 2014.

[12] S. Hohenberger, S. Myers, R. Pass, and A. Shelat, "Anonize: A large-scale anonymous survey system," in *Proceedings of the 2014 IEEE Symposium on Security and Privacy*, pp. 375–389, 2014.

[13] Z. Huang, Z. Chen, and Y. Wang, "Convertible undeniable partially blind signatures," in *19th International Conference on Advanced Information Networking and Applications (AINA'05)*, vol. 1, pp. 609–614, Mar. 2005.

[14] D. Husemoller, *Elliptic Curves*, Springer, 2^{nd} edition, 2004.

[15] M. S. Hwang, C. C. Lee, Y. C. Lai, "An untraceable blind signature scheme", *IEICE Transactions on Foundations*, vol. E86-A, no. 7, pp. 1902–1906, July 2003.

[16] D. Jao and V. Soukharev, "Isogeny-based quantum-resistant undeniable signatures," in *Post-Quantum Cryptography*, LNCS 8772, pp. 160–179, Springer International Publishing, 2014.

[17] D. Y. Jao and R. Venkatesan, *Use of Isogenies for Design of Cryptosystems*, CA Patent 2,483,486, Dec. 24, 2013.

[18] D. Kohel, K. Lauter, C. Petit, and J. P. Tignol, "On the quaternion ℓ-isogeny path problem," *LMS Journal of Computation and Mathematics*, vol. 17, pp. 418–432, 2014.

[19] A. Koide, R. Tso, and E. Okamoto, "Convertible undeniable partially blind signature from bilinear pairings," in *IEEE/IFIP International Conference on Embedded and Ubiquitous Computing (EUC'08)*, vol. 2, pp. 77–82, Dec. 2008.

[20] C. C. Lee, M. S. Hwang, and W. P. Yang, "A new blind signature based on the discrete logarithm problem for untraceability", *Applied Mathematics and Computation*, vol. 164, no. 3, pp. 837–841, May 2005.

[21] C. T. Li, M. S. Hwang, C. Y. Liu, "An electronic voting protocol with deniable authentication for mobile Ad hoc networks," *Computer Communications*, vol. 31, no. 10, pp. 2534–2540, June 2008.

[22] I. C. Lin, M. S. Hwang, C. C. Chang, "Security enhancement for anonymous secure E-voting over a network", *Computer Standards & Interfaces*, vol. 25, no. 2, pp. 131–139, May 2003.

[23] R. L. Rivest and B. Jr. Kaliski, "RSA problem," in *Encyclopedia of Cryptography and Security*, pp. 1065–1069, Springer US, 2011.

[24] M. Rückert, "Lattice-based blind signatures," in *Advances in Cryptology (ASIACRYPT'10)*, LNCS 6477, pp. 413–430, Springer Berlin Heidelberg, 2010.

[25] K. Sakurai and Y. Yamane, "Blind decoding, blind undeniable signatures, and their applications to privacy protection," in *Information Hiding*, LNCS 1174, pp. 257–264, Springer Berlin Heidelberg, 1996.

[26] D. Schröder and D. Unruh, "Security of blind signatures revisited," in *Public Key Cryptography (PKC'12)*, LNCS 7293, pp. 662–679, Springer Berlin Heidelberg, 2012.

[27] J. H. Silverman, *The Arithmetic of Elliptic Curves*, Springer, 2^{nd} edition, June 2009.

[28] V. Soukharev, D. Jao, and S. Seshadri, "Post-quantum security models for authenticated encryption," in *7th International Workshop on Post-Quantum Cryptography (PQCRYPTO'16)*, pp. 64–78, Springer, 2016.

[29] F. Zhang and K. Kim, "ID-based blind signature and ring signature from pairings," in *Advances in Cryptology (ASIACRYPT'02)*, LNCS 2501, pp. 533–547, Springer Berlin Heidelberg, 2002.

Biography

M. S. Srinath has two masters degrees; one in mathematics and one in computer science. Currently he is pursuing doctoral research at the Department of Mathematics and Computer Science, Sri Sathya Sai Institute of Higher Learning. His areas of interest include isogeny-based cryptography, post-quantum cryptography, and mathematical problems in cryptography.

V. Chandrasekaran has a masters degree in engineering from Indian Institute of Science and a Ph.D. from University of Melbourne. He is a senior member of IEEE. Currently he is an Honorary Professor at the Department of Mathematics and Computer Science, Sri Sathya Sai Institute of Higher Learning. His areas of interest include cryptography, neural networks, image processing, video processing, and computer vision.

New Protocol E-DNSSEC to Enhance DNSSEC Security

Kaouthar Chetioui, Ghizlane Orhanou, and Said El Hajji
(Corresponding author: Kaouthar Chetioui)

Laboratory of Mathematics, Computing and Applications, Faculty of Science, Mohammed V University in Rabat
BP. 1014 RP, Rabat, Morocco
(Email: kaoutharchetioui@gmail.com)
(Received Sep. 28, 2016; revised and accepted Jan. 15, 2017)

Abstract

The Domain Name System (DNS) is an essential component of the internet infrastructure. Due to its importance, securing DNS becomes a necessity for current and future networks. DNSSEC, the extended version of DNS has been developed in order to provide security services. Unfortunately, DNSSEC doesn't offer query privacy; we can see all queries sent to resolver in clear. In this paper, we evaluate the security of DNS and DNSSEC protocols, and we would see clearly that DNSSEC is insufficient to secure DNS protocol; it doesn't ensure confidentiality to data transiting over the network. That's why, we propose a new method named 'E-DNSSEC' which aims to add, in addition to DNSSEC security features, queries confidentiality, by encrypting them between DNSSEC servers. After that, an implementation of E-DNSSEC protocol will be given. Finally, we conclude by an analysis to prove the positive impact of this method to enhance DNSSEC security.

Keywords: Confidentiality; DNSSEC; E-DNSSEC

1 Introduction

DNS is a distributed database globally accessible using a request/response architecture. The DNS protocol resolves domain names readable by humans to (IP) Internet Protocol addresses. So, the DNS resolution is the first step in any network communication. It is therefore essential that the DNS infrastructure be robust and secured [5]. That's why, we need to enhance DNS protocol security to be able to ensure at least authentication, integrity and confidentiality.

TSIG (Transaction Signatures) defined in RFC 2845 [10] is a solution used in order to ensure the integrity of channels; it allows two machines talking DNS to check the identity of the caller. Unfortunately, this mechanism does not authenticate source data, only it secures transmission data between two parties who share the same secret key. The original source data can come from a compromised zone master or can be corrupted during transit from an authentic zone master to some "caching forwarder" [10]. These signature mechanisms are reserved only to protect zone transfers and dynamic update messages. So, TSIG is mostly used between master and slave DNS servers to secure zone transfers and today almost all transfers between authoritative severs are protected by TSIG.

DNSSEC (DNS Security Extension) defined in [RFC4033-4035], is proposed and standardized in 1997 [3], it solves some security issues related to DNS protocol. DNSSEC secures data sent by DNS servers; it ensures two security objectives namely authentication and integrity of source of data. These extensions use cryptography to sign DNS records and put the signature in DNS. Thus, a suspicious DNS client can retrieve the signature and using the key of the server, it can check if data is correct. DNSSEC allows delegation of signatures and the register of a TLD (Top-Level Domain) can announce that this subdomain is signed. By using DNSSEC, we can also build a chain of trust from the root server.

Despite of services provided by DNSSEC protocol, it has some gaps which make its deployment slowed:

- The compatibility with the existent equipment and software;
- The deployment of DNSSEC in a wide range of DNS servers and DNS resolvers (clients);
- The protection of data transiting in the network that ensures confidentiality service;

When communications requires private channels, SSH or IPsec are used to interact with DNS. These technologies are considered as there is no DNS solutions proposed for this case. But, all of them suffer from different security problems [4].

In this paper, we propose a new method E-DNSSEC which uses cryptography to encrypt DNSSEC query transiting across the network. This method aims to add

a new strong security service to DNSSEC protocol and consequently enhance security in DNS service and Internet communications. We describe first the structure of DNSSEC message especially DNSSEC query which is subject of different type of attacks. We'll give some types of these attacks which justify the necessity of thinking about new method to secure data in transit. Then, we explain the new E-DNSSEC protocol by giving different steps of the query encryption. The result of the E-DNSSEC implementation will be presented. After that, we give security analysis of the results obtained in order to prove the efficiency of our method to enhance DNSSEC security. Finally, we conclude by comparing this solution with other existing methods and we give some perspectives.

2 DNSSEC Structure and Security Issues

In this section, we focus on the structure of DNSSEC query and DNSSEC resolution process and finally, we describe the limitations of DNSSEC protocol.

As DNSSEC is the extended version of DNS, it has the same tree structure as DNS [9], but it adds some improvements; it includes new records, services and techniques to secure DNS protocol. DNSSEC uses cryptography to secure zone files. So, each zone has at least a pair of key. The public key of the child zone (like "example.ma") is signed by the private key of the parent zone (in this example ".ma") with the exception of the root which is signed by itself. This process forms the trust chain (Chain of Trust).

DNSSEC uses cryptographic keys ZSK (Zone Signing Key) and KSK (Key Signing Key) and adds new resource records (RR KEY, SIG, NSEC and DS) to DNS messages in addition to the original DNS service records [1, 6].

1) DNSKEY RR record: The DNSKEY resource record stores all public key pairs that are necessary for signing the zone.

2) RRSIG record: The RRSIG record results from the signing of RRsets generated by the private key, it provides the digital signature to the provided data. So, it accompanies every RR and it's considered as the basic block of DNSSEC which is necessary to verify the authenticity of the returned data.

3) NSEC record: is used by the DNSSEC protocol when the requested name doesn't exist. It's called proof of nonexistence and occasionally denial of existence.

4) DS record: It allows a parent zone to validate the KEY record of its child zone.

In addition to that, DNSSEC slightly modify the header of the classic DNS packet with the use of AD and CD bits:

- AD (Authenticated Data): As specified by RFC 2535 [7] indicates in a DNS response that all information included in the "Answer" and "Authority" have been authenticated by the server according to its security policy. However, in practice this does not seem very useful since a properly configured DNS server should not respond to a request with data that have not been authenticated.

- CD (Checking Disable): This bit specifies whether the resolver accepts unverified responses, when set to 1. Otherwise (value 0), the principle of verification is active.

When a DNSSEC resolver (client) sends a query to DNSSEC (server), the query is transformed by the system on DNS request readable by the resolver. This transformation depends on the local OS and data structure on the system. We keep the standard format of query as presented in Figure 1.

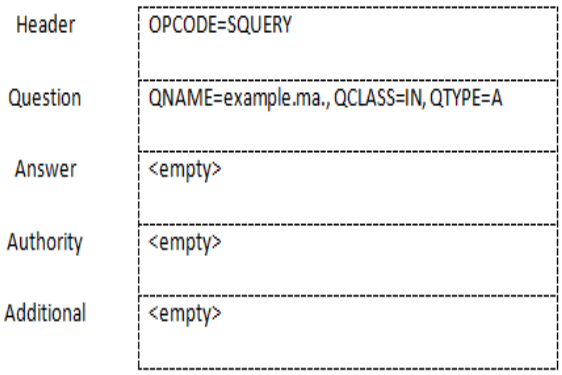


Figure 1: Format of standard query in DNS message [10]

As mentioned in Figure 1, the standard query contains other fields in addition to the domain named entered by the client (ex: "example.ma"). The communication in DNS follows the client/server model. So, when the server receives the query, it looks for the response in its database to have the IP address associated for the desired domain name, especially, it makes search in zone files which contain all information about domains names.

The process of DNSSEC resolution is described in Figure 2.

As described in Figure 2, we observe that the query (1) sent by recursive server to the authoritative server on a subdomain is neither encrypted nor signed.

Referring to RFC4033 [3], many security services are not provided by DNSSEC:

- DNSSEC doesn't provide confidentiality;

- DNSSEC doesn't provide access list control;

- DNSSEC doesn't protect from denial of service attacks.

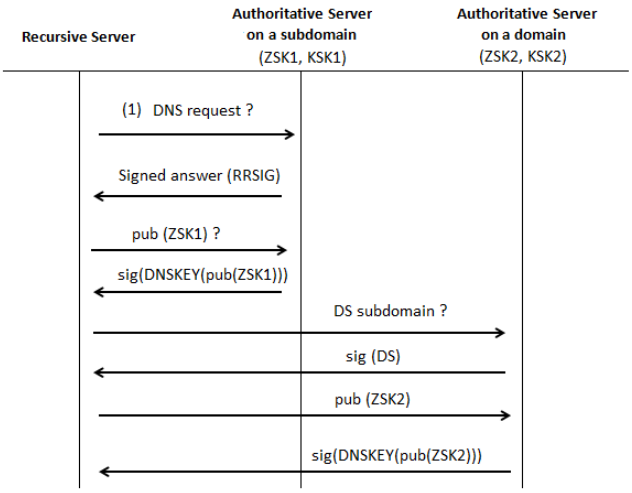


Figure 2: Diagram of DNSSEC resolution process

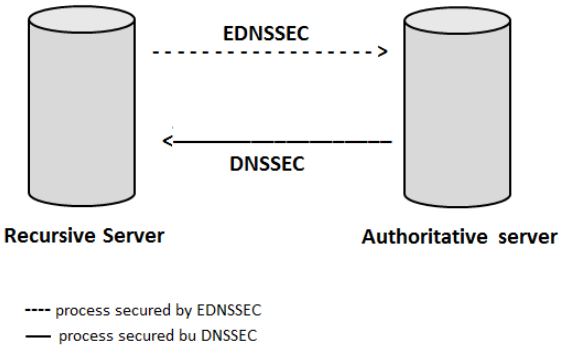


Figure 3: Processes secured by DNSSEC and E-DNSSEC protocols

In addition, DNS security extensions use public key cryptography to sign and authenticate DNS resource records, RFC 4035 [3]. An active attacker who can obtain the CD bit in a DNS query message or the AD bit in a DNS response message can use these bits to decrease the protection that DNSSEC attempts to provide to resolvers in recursive mode.

For this reason, the use of these control bits by a DNSSEC resolver in recursive mode requires a secure channel.

In the next section, we present the new method E-DNSSEC which encrypts data (DNSSEC query) transmitted across the network. This solution aims to secure the channel between DNS servers.

3 The New E-DNSSEC Protocol

In this section, we will describe our proposal which consists on E-DNSSEC protocol. We begin by a description of DNSSEC resolution process and after that, we present the new E-DNSSEC process in order to enhance DNSSEC security level. As we aim to encrypt DNSSEC query between DNSSEC client and DNSSEC server, we will refer to our proposal as E-DNSSEC, (Encrypted DNSSEC query).

The idea of E-DNSSEC protocol is to take the query from recursive server and encrypt it before sending it to the authoritative server; and in the reception, the authoritative server decrypts it before starting resolution and finally, it sends the response secured using DNSSEC protocol. So, the principal objective of this method is to combine DNSSEC properties with E-DNSSEC protocol to secure DNS message from the beginning of resolution to the end consequently ensuring authenticity, confidentiality and integrity of data transiting in the network.

In our demonstration as mentioned in Figure 3, we use two entities: a resolver (handling outgoing requests and incoming responses) and a DNS server (handling incoming requests and outgoing responses). These entities communicate on same DNS server because we suppose that the request received by the server needs recursion to be resolved. In recursive resolution, a stub resolver is created to be a DNS client; the stub resolver after consulting /etc/hosts, sends a recursive DNS query to a DNS server. In this case, the DNS server asks the resolver for the resolution of a request and sends the response to the stub resolver or another that queried it.

In the following, we will describe what happens during a resolution process inside a DNS server during the lifetime of a DNS query. Figure 4 below explains the steps of DNSSEC resolution.

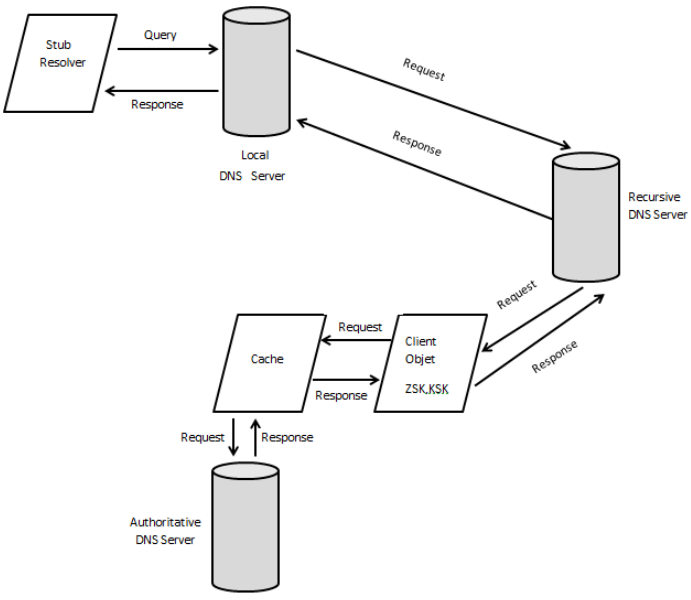


Figure 4: DNSSEC resolution process

Referring to Figure 4, the process of DNSSEC resolution starts when the stub resolver sends a query to its local DNS server. The local DNS server dispatches the re-

quest to the client manager, which is responsible for principal DNS communication, it's attached to each network interface and it manipulates all DNS messages arriving to the DNS server. After receiving the request by client manager, a client object is created in order to resolve the request. First, it verifies the public key signature, if the DNSSEC verification is successful; it looks for the response in the cache and in the authoritative DNS server. Once the response is found, it returned it to the DNS server, which sends it back to the Stub Resolver [8].

In our study, we focus on encrypting the query before sending it to the DNS server. So, E-DNSSEC keeps the structure of DNSSEC protocol and add the process of encrypting the query in a DNS message. Figure 5 shows when the process of encryption starts and finishes.

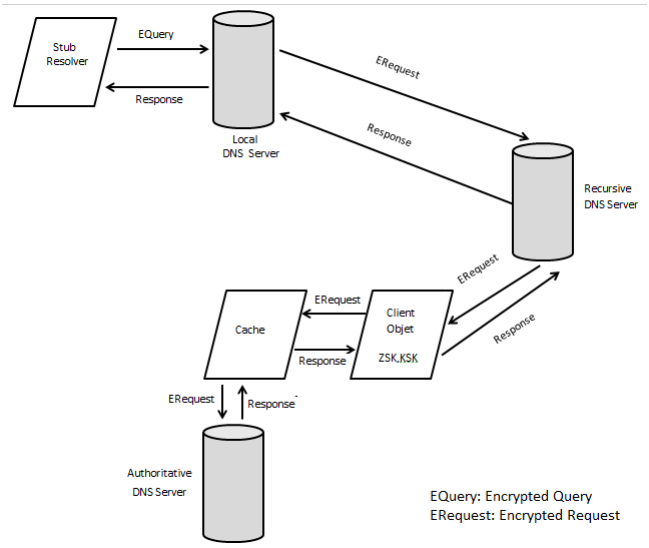


Figure 5: E-DNSSEC resolution process

To resolve a query by E-DNSSEC protocol, the local DNS server encrypts the query and inserts it in the DNS request. When the recursive DNS server receives the DNS request, it recuperates the E-query and decrypts it. The client object verifies public key signatures as usually done by DNSSEC protocol and looks for the response in its cache, if the response is found in the cache it returns it to the DNS server, else, the process of resolution continues by encrypting the query and sending the request to the authoritative DNS server. Once the response is found, the process of encrypting query is stopped. After that, the response is sent to the DNS server and the Stub Resolver. During this step, the response is secured using signed resource records of DNSSEC protocol.

To implement E-DNSSEC protocol, we use the original version BIND9. BIND is a complex program with its own tasks, threads, scheduler and memory management. We select version 9 of BIND, its source code is freely available. So, in this implementation, we look after the IP address of the domain name 'example.ma' and we use RSA algorithm for encryption. The different steps of our implementation are described below:

DNS ——> DNSSEC ——> E-DNSSEC

1) Firstly, we implement a simple DNS server with the domain name 'example.ma' and we interrogate it with 'dig' command:

```
; <<>> DiG 9.9.5 <<>> server.example.ma
;; global options: +cmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 19267
;; flags: qr aa rd ra; QUERY: 1, ANSWER: 1, AUTHORITY: 2, ADDITIONAL: 1

;; OPT PSEUDOSECTION:
; EDNS: version: 0, flags:; udp: 4096
;; QUESTION SECTION:
;server.example.ma.              IN      A

;; ANSWER SECTION:
server.example.ma.      86400   IN      A       192.168.1.1

;; AUTHORITY SECTION:
example.ma.             86400   IN      NS      192.168.1.1.example.ma.
example.ma.             86400   IN      NS      192.168.1.10.example.ma.
```

Figure 6: Result of Dig command using DNS

Figure 6 represents the result of this interrogation of a simple DNS server by 'dig' command. The response reflects a typical positive response to a dig command and includes the following items: Question Section, Answer Section, Authority Section and Additional Section. In this paper, we are interested only to the Question Section. Figure 7 shows that this section transit in clear across the network and could be intercepted by a malicious person [6].

2) We implement DNSSEC protocol. This implementation is done following several steps [6]: creating keys (ZSK, KSK), including the keys in the zone files, signature of the zone and reconfiguration of the 'named.conf' using DNSSEC. The results of 'dig' command are illustrated in Figure 8.

   As shown in Figure 8, after implementing DNSSEC protocol, the output of 'dig' command is changed especially in the Answer Section; the response is signed but the Question Section is still unchanged.

3) Based on DNSSEC architecture, we enhanced BIND source files by implementing the query encryption. We interrogate the server by 'dig' command as shown in Figure 9.

Figure 9 indicates that the Question Section is different compared to the previous result; the query 'server.example.ma' is encrypted.

4 Analysis Of E-DNSSEC Protocol

In this section, we aim to compare the new E-DNSSEC protocol with existing protocols. Furthermore, having good security results by using E-DNSSEC protocol is very

```
[root@localhost ~]# tcpdump -nn host 192.168.1.10 and host 192.168.1.1
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on eth0, link-type EN10MB (Ethernet), capture size 96 bytes
15:41:54.340664 IP 192.168.1.10.51965 > 192.168.1.1.53: 49733+ A? server.example.ma. (35)
15:41:54.340776 IP 192.168.1.1 > 192.168.1.10: ICMP host 192.168.1.1 unreachable - admin prohibited, length 71
```

Figure 7: DNS query captured by Tcpdump

```
[root@localhost ~]# dig +cd +multi example.ma dnskey
;; Truncated, retrying in TCP mode.

; <<>> DiG 9.6.1b1-RedHat-9.6.1-0.3.b1.fc11 <<>> +cd +multi example.ma dnskey
;; global options: +cmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 41792
;; flags: qr aa rd ra cd; QUERY: 1, ANSWER: 2, AUTHORITY: 0, ADDITIONAL: 0

;; QUESTION SECTION:
;example.ma.            IN DNSKEY

;; ANSWER SECTION:
example.ma.             86400 IN DNSKEY 257 3 5 (
                                AwEAAcQK6XP43phN3h4x7eMbpkEidvPjWmeynsZAfECq
                                quRMdplwgkuMftFX3HsjBpo5u7IhGYc1MkyeTV8vTajq
                                rIcm4TerWq3PkjL+0iKUTPo3mca5vjxRyw5CNORLYbwA
                                ZhIP/RAOZ1eKW0tX/46z5FI8crFm1ycX1Dp2eJYQmuuA
                                ODOKorEX2biALW1oSmxshvq66yqbew3MVwMFFVbu4i/J
                                LzfgpXosgGKcTHaFIYNiR15JJm9g0x/QBQM58LJUyyOr
                                5OzmCDny5hwYjbfmbyOkW82/KSQZ9mR4QIDexf6UH5ug
                                c7OOJBnVjCH+yvKC92NuTProcGiWlCs=
                                ) ; key id = 25924
```

Figure 8: Result of Dig command using DNSSEC

```
[root@localhost ~]# dig server.example.ma +dnssec

; <<>> DiG 9.9.5 <<>> server.example.ma +dnssec
;; global options: +cmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: SERVFAIL, id: 16884
;; flags: qr rd ra; QUERY: 1, ANSWER: 0, AUTHORITY: 0, ADDITIONAL: 1

;; OPT PSEUDOSECTION:
; EDNS: version: 0, flags: do; udp: 4096
;; QUESTION SECTION:
;A33301A32D126CA38.             IN      A
```

Figure 9: Result of Dig command using E-DNSSEC

useful to prevent our systems from different types of attacks, but having good performance is also very important. So, we present also the impact of this method in the DNSSEC performance. The principal objective of E-DNSSEC protocol is to enhance the DNS protocol security. For that, we propose a new method that encrypts DNS query and we aim to ensure confidentiality without using a tunnelling protocol like IP security (IPsec). This new method prevents DNS system from cache poisoning and Man in The Middle attacks. Authentication: In E-DNSSEC, authentication can be ensured by different ways:

Access control list: It's the first method deployed by BIND system in order to prevent DNS from IP spoofing attack. Access control lists give a list of IP addresses that are authorized to query the resolvers. This method is still inadequate to prevent DNS from attacks [7].

MAC function: When a server receives a query from another server, it calculates its MAC function and compare it with the hash, if they are the same, it considers that the request comes from an authorized name server and it treats it, else it rejects the request.

DNSSEC: By using DNSSEC, resolvers can verify the authentication of source data. Also, it uses signatures to authenticate parent with child zone (chain of trust) [2].

Confidentiality: By using E-DNSSEC, queries are transmitted securely across the network; every query is encrypted before being sent to the destination. This encryption is used in order to ensure a high level of security to data transmitted between servers.

Integrity: Deploying E-DNSSEC has a major advantage is that it keeps DNSSEC properties. As known, by using DNSSEC, DNS servers are required to sign the Resource Records in zone for which they are authoritative and answers the queries by returning the corresponding SIG RRs. DNSSEC uses new resource records to ensure integrity of data. So, due to this signed resources records, the destination is sure that the request is not modified or changed in transmission. Finally, the higher level of authentication and confidentiality in the encapsulation process ensure a high level in integrity of data.

Our comparison is based essentially on the frequent DNS attacks like IP spoofing, file corruption and cache poisoning which are the famous attacks that make DNS system very weak. These attacks affect all DNS communication, whether it was between two DNS servers or a DNS server and a client.

Table 1: Types of attacks at each point and the possible solutions

Number	Area	Types of Attacks	Solutions			
			System Adm.	TSIG	DNSSEC	E-DNSSEC
1	Zone file (local)	File corruption	X			X
2	Dynamic update (server-server)	IP spoofing				X
3	Cache (server-client)	Cache poisoning			X	X
4	Resolver (Remote cache-client)	Data interception, IP spoofing		X	X	X

Table 1 below defines the position of different attacks in each area of a DNS system and the possible solutions.

According to Table 1, we can see that no ancient solution ensures data confidentiality between servers or between server and client, and the only solution that exist now is using IPsec to secure the canal, but this solution is still insufficient as we have explained in Section 2.

5 Conclusion

In this paper, we have proposed a new method E-DNSSEC. E-DNSSEC keeps all security properties of DNSSEC and encrypts the query in order to secure DNS across the network from different type of attacks. This solution adds confidentiality service to DNS system in addition to the authentication and integrity which are ensured by DNSSEC protocol. Furthermore, we have presented a functional implementation of E-DNSSEC protocol, we compare this method by other existing solutions and finally we discuss the impact of this solution on the security performance of DNS protocol.

References

[1] R. Aitchison, *Pro DNS and BIND 10*, Apress, 2011.

[2] P. Albitz, and C. Liu, *DNS & BIND*, O'Reilly, 2006.

[3] R. Arends, R. Austein, M. Larson, D. Massey and S. Rose, *DNS Security Introduction and Requirements*, RFC 4033, Mar. 2005.

[4] D. Atkins, R. Austein, *Threat Analysis of the Domain Name System (DNS)*, RFC 3833, Aug. 2004.

[5] P. Barnes, "Using DNS to protect networks from threats within," *Network Security*, vol. 2014, no. 3, pp. 9–11, Mar. 2014

[6] K. Chetioui, G. Orhanou, S. El Hajji and A. Lakbabi, "Security of the DNS protocol - Implementation and weaknesses analyses of DNSSEC," *International Journal of Computer Science Issues*, vol. 9, no. 2, pp. 340–345, 2012.

[7] D. Estlake, *Domain Names System Security Extensions*, RFC 2535, Mar. 1999.

[8] D. Herrmann, M. Maa and H. Federrath, "Evaluating the security of a DNS query obfuscation scheme for private Web surfing," in *IFIP International Information Security Conference SEC 2014: ICT Systems Security and Privacy Protection*, pp. 205–219, Marrakech, Morocco, June 2014.

[9] A. Velagapalli and M. Ramkumar, "Trustworthy TCB for DNS servers," *International Journal of Network Security*, vol. 14, no. 4, pp. 187-205, July 2012.

[10] P. Vixie, O. Gudmundsson, D. Eastlake and B. Wellington, *Secret Key Transaction Authentication for DNS (TSIG)*, RFC 2845, May 2000.

Biography

Kaouthar Chetioui, she has received her Ph.D degree in 2017 in computer science from Mohammed V University in Rabat, Morocco. She received in 2011 a Master degree in Cryptography and Security of Information from the same university and in 2009 a licence in Network Technologies from Sidi Mohammed Ben Abdellah University, Fez, Morocco. Her main research domains include networks and Internet protocols security.

Ghizlane Orhanou, Professor in Faculty of Sciences and member of the Laboratory of Mathematics, Computing and Applications, Mohammed V University in Rabat, Morocco since 2013. She received Ph.D degree in Computer sciences from the Mohammed V University in Rabat in 2011 and the Habiltation to direct theses in 2016 from the same University. She received in 2001 a Telecommunication Engineer diploma from Telecommunication Engineering Institute (INPT - Morocco), and worked for about 3 years as GPRS and Intelligent Network Engineer, and for 9 years as System and Network Security Engineer. Her main research interests include network and information systems security.

Said El Hajji, Professor in the Mathematics Department since 1991 at Mathematical and Computer Sciences, Faculty of Sciences, University of Mohammed V-Rabat. Responsible of the Mathematics, Computing and Applications Laboratory. He received Ph.D degree from Laval University in Canada. His main research interests include modeling and numerical simulations, security in networks and Information systems.

New Intrusion Detection System Based on Support Vector Domain Description with Information Gain Metric

Mohamed El Boujnouni and Mohamed Jedra
(Corresponding author: Mohamed El Boujnouni)
Laboratory of Conception and Systems, Mohammed V University
Avenue Ibn Battouta B.P. 1014, Rabat, Morocco
(Email: med_elbouj@yahoo.fr)
(Received Nov. 20, 2016; revised and accepted Feb. 21 & Mar. 4, 2017)

Abstract

With the vulgarization of Internet, the easy access to its resources and the rapid growth in the number of computers and networks, the security of information systems has become a crucial topic of research and development especially in the field of intrusion detection. Techniques such as machine learning and data mining are widely used in anomaly-detection schemes to decide whether or not a malicious activity is taking place on a network. This paper presents a new intrusion detection system (IDS) based on information gain criterion to select relevant features from network traffic records and a new version of support vector domain description to classify the extracted features and to detect new intrusions. Experimental evaluation on NSL-KDD, a filtered version of the original KDD99 has shown that the proposed IDS can achieve good performance in terms of intrusions detection and recognition.

Keywords: Information Gain Metric; Network Intrusion Detection System; NSL-KDD; Support Vector Domain Description

1 Introduction

In computer science domain, an intrusion can be defined as the attempts to compromise the confidentiality, integrity, or availability of a computer or network. Thus, Network Intrusion Detection Systems (NIDS) are a critical defense layer of any network security architecture. The main task of NIDS is to monitor network traffic for suspicious contents, and to alert system administrators when a malicious activity is taken place. The detection of intrusions can be performed basing on analyzing the events which occur in the monitored network. Two primary approaches are used: misuse or signature detection and anomaly detection. The first technique, existing in the majority of commercial NIDSs, aims to detect known attacks by using predefined attack patterns and signatures so it looks for a specific event that has already been recognized and registered. The second technique detects attacks by comparing the deviation from a model describing the normal behavior of the monitored resource. There are advantages and disadvantages associated with each approach: Misuse detection methods can detect malicious network traffics without generating high false alarms but they are basically limited to known attacks. This leads to the necessity for frequent updates of the intrusions database. On the contrast anomaly detection methods based on heuristics or rules are able to detect known and unknown attacks. This propriety is very important since new kinds of vulnerabilities and intrusions are constantly appearing. However, new legitimate behavior can be falsely identified as malicious, resulting in a false positive. Recently new hybrid intrusion detection systems that exploit benefits of both misuse and anomaly detection techniques are developed and showed great success [21, 28].

Anomaly detection approach is based on techniques such as: Threshold detection, rule-based measures, statistical measures, machine learning and data mining methods. The first technique expresses some attributes of user and system behavior in terms of counts. Then it compares the latter with a tolerance level. The second approach tries to define a set of rules that can be used to decide whether a given behavior is normal or not. Statistical measures analyze the distribution of the network traffic attributes and can be parametric or non-parametric, the first one is assumed to fit a particular pattern while the second is learned from a set of historical values. The last technique based on machine learning and data mining learns from a set of training data and constructs a model able to classify new network traffic as legitimate or malicious.

In this paper we aim to design a new intrusion detec-

tion system based on the last technique described above. The proposed NIDS works in three steps: At first, a data encoding and normalization operations are performed on the network traffic records. Then, information gain (IG) method is applied to extract relevant features from the preprocessed data. Finally, a new version of SVDD called SVDD with small sphere and parametric volume (SSPV-SVDD) [3] will be trained with the extracted features and used as a novelty detection model able to detect unknown attacks. Experimental evaluation of our approach will be performed using NSL-KDD a benchmark dataset widely used to evaluate the performance of NIDSs.

This paper is organized as follows: Section 2 presents an overview of some previous applications of machine learning and soft computing methods to detect network intrusions. Section 3 describes in details our new network intrusion detection system. This section is divided into three parts: The first one describes the architecture of the proposed NIDS, the second presents the techniques of data encoding, data normalization and relevant feature extraction and finally the third presents the application of SSPV-SVDD to detect network intrusions. The last section investigates empirically the performance of the proposed NIDS using NSL-KDD. This section is divided into two parts: The first one describes NSL-KDD dataset and the second presents the experimental setting and the results of applying the proposed NIDS on NSL-KDD. A conclusion is provided in the final section.

2 Related Work

There are numerous important research papers regarding the use of machine learning and soft computing techniques to detect network intrusion. For example Liu et al. [18] proposed a genetic clustering method for intrusion detection. Their method is able to establish clusters automatically and to detect attacks by labeling normal and abnormal groups. Javadzadeh and Azmi [13] proposed a hybrid approach to design NIDSs. Their method is able to generate fuzzy rules based on a fuzzy genetic machine learning algorithm and to detect multiple attacks. Aghdam and Kabiri [1] focused on feature selection of network traffic for intrusion detection purpose. They proposed a new method based on ant colony optimization (ACO) algorithm and nearest neighbor classifier to eliminate irrelevant and redundant features from network traffic records. Wang et al. [29] presented an application of artificial neural networks (ANN) and fuzzy clustering on intrusion detection. Their approach works sequentially: Firstly fuzzy clustering technique was applied to create different training subsets. Then, based on the latter, different ANN models are trained to formulate different base models. Finally, a fuzzy aggregation module is employed to aggregate these results. Li et al. [35] introduced an application of multiple kernel support vector machine (SVM) for intrusion detection. This new version of SVM improves the standard one by calculating the weights of kernel functions and Lagrange multipliers simultaneously and automatically without user intervention. Mukherjee and Sharma [20] presented an intrusion detection method based on naive Bayes classifier with a new feature reduction method. In order to select the most relevant features the authors investigate the performance of three standard features selection methods, namely correlation-based features selection, information gain and gain ratio. Then they proposed a new features reduction method named feature vitality. The reduced data sets are further classified using Native Bayes classifier. Li et al. [17] proposed the use of K-means clustering and particle swarm optimization (PSO) algorithm to deal with network intrusions. The key idea behind using PSO is to reach a good overall convergence and to overcome falling into local minima. Wankhade et al. [30] discussed the development of a secured information system by applying various data mining techniques on intrusion detection systems for the effective identification of both known and unknown attacks. Tao et al. [24] presented one-class classification approach to detect network intrusions based on SVDD. They used genetic algorithm to determine the optimal parameter of the kernel function. Then they analyzed the behavior of the classifier basing on the selected parameters. Yu Zhang et al. [36] proposed an optimized method of SVDD based on particle swarm optimization algorithm (PSO). Their method adopts PSO to eliminate the superfluous parameters in SVDD and carries out dimension reduction to data. GhasemiGol et al. [9]. presented a novel approach to describe the normal behavior of computer networks using minimal hyper-ellipse instead of hypersphere used by SVDD. The hyper-ellipse creates tighter boundary around the positive examples. The boundary was used to detect new attacks. Zhou et al. [38] presented an improved intrusion detection method based on kernel learning. They used Kernel principal component analysis (KPCA) as preprocessor of the dataset. Then they applied SVDD on the preprocessed data. Kenaza et al. [16] introduced an adaptive SVDD-based learning for false alarm reduction in intrusion detection. In their work they aimed to take into consideration the dynamic aspect of a monitored environment, and they proposed an adaptive SVDD-based learning approach that aims at continuously enhancing the performances of the SVDD classifier by refining the training dataset. Yang et al. [32] proposed a new method for anomaly intrusion detection based on SVDD. In their work they considered intrusion detection problem as one-class classification and then they built SVDD model for normal data. This model was used to detect known and unknown attacks. Yang et al. [33] introduced a new framework for adaptive anomaly detection based on SVDD classifier and change detection algorithm. The proposed framework consists of four main components: preprocessor, change detector, model generator and anomaly detector.

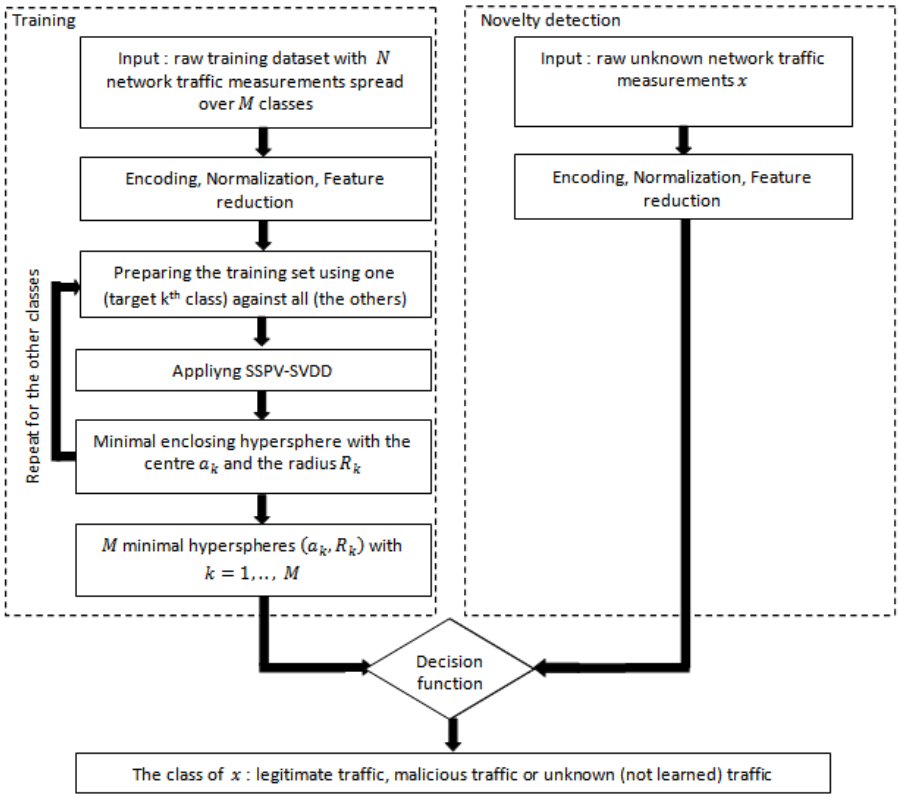


Figure 1: Architecture of the proposed network intrusion detection system

3 The Proposed Network Intrusion Detection System

3.1 Architecture of the Proposed NIDS

Figure 1 shows the architecture of the proposed network intrusion detection system. The proposed NIDS works in two steps:

- Training: After preprocessing (encoding, normalization and feature reduction) a set of network traffic measurements having M classes describing normal and attacks behaviors, SSPV-SVDD with Gaussian kernel will be applied on each class. The result is a set of M minimal hyperspheres each of which has a center a_k and a radius R_k with $k = 1, \cdots, M$ and encloses the samples of a specified class.

- Novelty detection: After preprocessing an unknown network traffic measurement x with the same technique used in the training step. The decision function expressed by Equation (5) will be evaluated. The result is either the class label of x or no one of the learned classes which signify that x is a new type of attacks.

3.2 Encoding, Normalization, and Feature Reduction

Network traffic contains different forms of data (continuous, discrete and symbolic) with significantly varying resolution and ranges, in order to handel this dataset with SSPV-SVDD a preprocessing is required. The latter is based on 3 steps:

Step 1: Convert symbolic attributes to numeric values. The conversion is performed using the encoding tables shown in Tables 1, 2, 3, 4.

Table 1: Encoding of symbols in the 2^{nd} field of NSL-KDD dataset

Symbol	tcp	udp	icmp
Code	1	2	3

Step 2: Normalize numeric values [37]. The data attributes are scaled to fall within the interval $[x_{min}^{new}, x_{max}^{new}]$ that can be [-1, 1] or [0,1].The scaling is performed using Equation (1). Likewise, before testing, the same way is applied to scale testing data. The main advantage is to avoid attribute in greater numeric ranges dominate those in smaller numeric

Table 2: Encoding of symbols in the 3^{rd} field of NSL-KDD dataset

Symbol	ftp_data	other	private	http	remote_job	name	netbios_ns	eco_i
code	1	2	3	4	5	6	7	8
Symbol	mtp	telnet	finger	domain_u	supdup	uucp_path	Z39_50	smtp
code	9	10	11	12	13	14	15	16
Symbol	csnet_ns	uucp	netbios_dgm	urp_i	auth	domain	ftp	bgp
code	17	18	19	20	21	22	23	24
Symbol	ldap	ecr_i	gopher	vmnet	systat	http_443	efs	whois
code	25	26	27	28	29	30	31	32
Symbol	imap4	iso_tsap	echo	klogin	link	sunrpc	login	kshell
code	33	34	35	36	37	38	39	40
Symbol	sql_net	time	hostnames	exec	ntp_u	discard	nntp	courier
code	41	42	43	44	45	46	47	48
Symbol	ctf	ssh	daytime	shell	netstat	pop_3	nnsp	IRC
code	49	50	51	52	53	54	55	56
Symbol	pop_2	printer	tim_i	pm_dump	red_i	netbios_ssn	rje	X11
code	57	58	59	60	61	62	63	64
Symbol	urh_i	http_8001	aol	http_2784	tftp_u	harvest		
code	65	66	67	68	69	70		

Table 3: Encoding of symbols in the 4^{rd} field of NSL-KDD dataset

Symbols	SF	S0	REJ	RSTR	SH	RSTO	S1	RSTOS0	S3	S2	OTH
Codes	1	2	3	4	5	6	7	8	9	10	11

Table 4: Encoding of symbols in the 41^{rd} field of NSL-KDD dataset

Class	**Sampling Rate**	**Length**	**Code**
Non-attack	1	Normal	0
DOS	10	Back, land, neptune, pod, smurf, teardrop, apache2, processtable, worm, udpstorm, mailbomb	1
Probe	6	Ipsweep, portsweep, nmap, satan, saint, mscan	2
R2L	16	Warezclient, guess_passwd, ftp_write, multihop, imap, warezmaster, phf, spy, snmpgetattack, httptunnel, snmpguess, named, sendmail, xlock, xsnoop	3
U2R	7	Rootkit, buffer_overflow, loadmodule, perl, ps, xterm, sqlattack	4

ranges.

$$x^{new} = x^{new}_{min} + \frac{x^{new}_{max} - x^{new}_{min}}{x^{old}_{max} - x^{old}_{min}}(x^{old} - x^{old}_{min}). \quad (1)$$

With x^{old}_{max} and x^{old}_{min} are respectively the maximum and the minimum values of the attribute that x belongs to, x^{old} is the value before normalization and x^{new} is the value after normalization that will belong to the interval $[x^{new}_{min}, x^{new}_{max}]$.

Step 3: Extract relevant features using information gain measure that can be expressed as follows: Let S be a set of M classes that contains s labeled training points where each class I includes s_i samples. Expected information needed to classify a given sample is evaluated using the following equation:

$$I(s_1, s_2, ..., s_M) = -\sum_{i=1}^{M} \frac{s_i}{s} log_2(\frac{s_i}{s}).$$

An attribute A with values $\{A_1, A_2, .., A_v\}$ can split the training set S into v subsets $\{S_1, S_2, .., S_v\}$ where S_j is the subset which has the value A_j for attribute A and contains s_{ij} points of class i. The entropy of the attribute A can be expressed as:

$$E(A) = \sum_{j=1}^{v} \frac{s_{1j} + ... + s_{Mj}}{s} I(s_{1j}, s_{2j}, ..., s_{Mj}). \quad (2)$$

Information gain for A is given by the equation:

$$Gain(A) = I(s_1, s_2, , s_M) - E(A). \quad (3)$$

3.3 Application of SSPV-SVDD to Detect Network Intrusion

Support Vector Domain Description is a relatively new classification method inspired by Support Vector Machine (SVMs). SVDD was originally developed by Tax and Duin [26, 27] and then improved by many researchers. This classifier aims to enclose the data of interest through the smallest hypersphere where its boundary serves to classify new unknown samples. Due to its high generalization capability, SVDD have been applied successfully to a wide range of problems, such as: Biometric authentication [10], novelty detection [7, 31], fault diagnosis [6, 34], credit ratings [8, 22], disease diagnosis [5, 15], digital investigations and computer security [4, 19], financial fraud detection [2, 14], etc. SVDD inherits many of the advantages of SVMs, including SVDD has a solid mathematical foundation based on the statistical learning theory. Also, it benefits from kernel functions that maps a linearly inseparable data points represented in the original space into a high dimensional feature space in which they become separable. In addition, training a given dataset with SVDD implies solving a constrained quadratic problem (QP) with a single minimum which avoids the risk of becoming trapped by local minimum solutions. Moreover, the classification of a new unknown sample requires checking the sign of a decision function basing only on a small subset of the training data known as support vectors (SVs) which reduces the time required to classify new unknown instances. Furthermore, training SVDD requires setting a small number of parameters which limits the intervention of users.

The proposed NIDS is designed with an improved version of SVDD called SSPV-SVDD [3]. The latter aims to improve SVDD by introducing a new regularization parameter that offers the following advantages: 1) It allows user to customize the hyperspherical boundary between different classes; 2) It plays a compromise between the acceptance of negative data and the rejection of target data; 3) It allows to distinguish between the set of samples existing on the boundaries.

SSPV-SVDD considers a dataset $S = \{(x_1, y_1), (x_2, y_2), \cdots, (x_N, y_N)\}$ with $i = 1, \cdots, N$ and $x_i \in \mathbb{R}^d$. The label y_i equals $+1$ for the target samples, and -1 for the negative ones. The objective of SSPV-SVDD is to find the smallest hypersphere, with a center a and a radius R that includes the maximum number of target samples and excludes the majority of negative ones following the value of a regularization parameter called p. This problem is formulated as follows:

Minimize:

$$R^2 + C\sum_{i=1}^{N} \varepsilon_i$$

Subject to:

$$\|x_i - a\|^2 \leq R^2 - p.y_i + \varepsilon_i, \forall i = 1, .., N, \quad \text{with } y_i = +1$$

$$\|x_i - a\|^2 \geq R^2 - p.y_i - \varepsilon_i, \forall i = 1, .., N, \quad \text{with } y_i = -1.$$

Where $\|.\|$ is the Euclidean norm. p is a strictly positive real number. ε_i are slack variables that measure the violation amount of the constraints. To allow the presence of outliers a positive parameter C was introduced, the latter gives the tradeoff between the volume of the sphere and the rejection of target samples.

It's an optimization problem with constraints that may be solved by Lagrange's method. The primal problem of SSPV-SVDD can be written as follows:

$$L(R, \varepsilon, a) = R^2 + C\sum_{i=1}^{N} \varepsilon_i - \sum_{i=1}^{N} \varepsilon_i y_i - \sum_{i=1}^{N} \alpha_i y_i (R^2 - \|x_i - a\|^2 - p y_i).$$

Where α_i and μ_i are Lagrange multipliers, Annulling the

partial derivatives of L with respect to R, a, ε_i yields:

$$\frac{\partial L}{\partial R} = 0 \quad \Rightarrow \quad \sum_{i=1}^{N} \alpha_i y_i = 1.$$
$$\frac{\partial L}{\partial a} = 0 \quad \Rightarrow \quad a = \sum_{i=1}^{N} \alpha_i x_i y_i$$
$$\frac{\partial L}{\partial \varepsilon_i} = 0 \quad \Rightarrow \quad \alpha_i = C - \mu_i.$$

The dual optimization problem can be written as:

Maximize:

$$L(\alpha) = -\sum_{i=1}^{N}\sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j x_i x_j + \sum_{i=1}^{N} \alpha_i (y_i x_i x_i + p)$$

Subject to:

$$0 \leq \alpha_i \leq C \; and \; \sum_{i=1}^{N} \alpha_i y_i = 1. \tag{4}$$

In multi-classes case, each class $k = 1, .., M$ is represented with a small sphere (a_k, R_k). To classify a new unknown sample z we have simply to evaluate the decision function given bellow:

$$class \; of \; z = argmax_{k=1,...,M}(1 - \sqrt{\frac{\|z - a_k\|^2}{R_k^2}})$$
$$with \; \|z - a_k\| \leq R_k.$$

The radius R_k that corresponds to the k^{th} class (minimal sphere), can be expressed as follows:

$$R_k^2 = \|x_s - a_k\|^2 + y_s p. \tag{5}$$

Where x_s is a training point that belongs to the set of Support Vectors of the k^{th} class having $0 < \alpha_s < C$.

4 Experimental Setup and Performance Evaluation

4.1 Description of KDD CUP 99 Dataset

In order to evaluate the performance of our new intrusion detection system, we propose to perform several experiments basing on KDD99 database. The latter is a version of the 1998 DARPA intrusion detection evaluation data set prepared and managed by MIT Lincoln Labs [12]. KDD99, widely used by NIDS researchers, includes a wide variety of intrusions simulated in a network environment. In this dataset, each record is composed by 41 fields and categorized into one of 5 classes that are normal and 4 types of attacks: DoS (denial of service), R2L (root to local), U2R (user to root) and Probing (surveillance).

As the number of records in the original KDD99 is very large and contains redundant data, we propose to work with the new filtered version of the original. This new dataset, called NSL-KDD, was created by Tavallaee et al. [25] after a statistical analysis of KDD99. NSL-KDD has the following major advantages over the original KDD99 [11]. It eliminates redundant records to which the classifier will be biased towards during the training. and it contains a reasonable number of records. Therefore, experiments can be achieved using the whole dataset without the need to randomly select a reduced subset.

Similar to KDD-99, NSL-KDD has the same number of attributes with a reduced number of samples. For KDD99 the training dataset contains 494021 records, and the testing dataset consists of 311029 records while NSL-KDD contains 125973 patterns for training and 22544 patterns for testing.

Figure 2 shows a random record taken from NSL-KDD. It can be seen that each attribute can take one of the following forms: continuous, discrete, or symbolic. The before last attribute in each record describes the class label that can take 40 values normal and 39 attacks that can be classified in 4 main intrusions DoS, R2L, U2R and Probing.

```
0,tcp,ftp_data,SF,491,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,2,2,0.00,
0.00,0.00,0.00,1.00,0.00,0.00,150,25,0.17,0.03,0.17,0.00,0.00,
0.00,0.05,0.00,normal,20
```

Figure 2: Random record taken from NSL-KDD dataset

4.2 Experimental Results

The intrusion dataset NSL-KDD is divided on two subsets training and testing. In the experiment the classifier SSPV-SVDD will be trained by the first subset, and then tested by the second. The parameter C is fixed at 100 and the kernel is Gaussian with $\sigma \in \{0.15, 0.30, ..., 1.5\}$. Gaussian kernel was chosen because of its wide uses in pattern recognition problems and its good generalization capability. Since we are using the whole NSL-KDD dataset having 125973 records the traditional QP solvers can 't be applied directly because they need to store Gram matrix of size N×N (125973×125973). To deal with this problem we propose to solve the QP of SSPV-SVDD with the algorithm called Sequential Minimal Optimization (SMO) [23], SMO is an iterative algorithm that decomposes the Quadratic Problem given by Equation (4) to the extreme in such a way that the working set only has two samples and their optimal Lagrange multipliers can be solved analytically. The recognition rates are calculated for the training and testing sets using the following

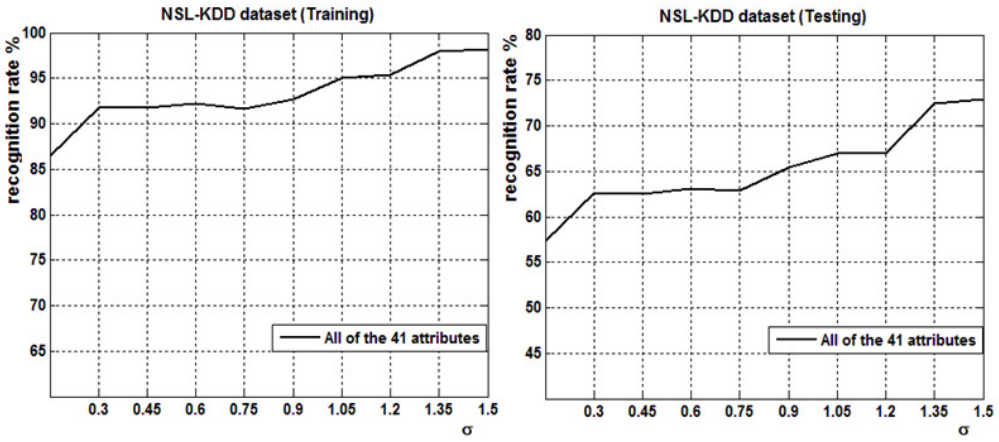


Figure 3: Classification accuracy of NSL-KDD using SSPV-SVDD with 41 attributes

equations:

$$\%(Normal) = \frac{\#of\ normal\ records\ well\ classified}{Total\ number\ of\ records} \times 100$$

$$\%(DoS) = \frac{\#of\ attacks\ DOS\ well\ classified}{Total\ number\ of\ records} \times 100$$

$$\%(R2L) = \frac{\#of\ attacks\ R2L\ well\ classified}{Total\ number\ of\ records} \times 100$$

$$\%(U2R) = \frac{\#of\ attacks\ U2R\ well\ classified}{Total\ number\ of\ records} \times 100$$

$$\%(Probing) = \frac{\#of\ attacks\ Probing\ well\ classified}{Total\ number\ of\ records} \times 100$$

$$Global\ recognition\ rate = \%(Normal) + \%(DoS) + \%(R2L) + \%(U2R) + \%(Probing).$$

Figure 3 shows the classification accuracy of NSL-KDD dataset using the entire 41 attributes. The figure is divided into two parts: The left side shows the global recognition rate of the training dataset with different values of σ. It can be seen that the recognition rate grows with σ until a maximum value that reaches 98%. This signifies that relatively 123454 training instances out of 125973 are well enclosed by the minimal five hyperspheres. The right side illustrates the generalization capability of our NIDS. It can be observed that the recognition rate increases with the kernel width until a maximum of 72.5% which means that 16344 of records out of 22544 are well classified. To increase further the recognition rate next experiment will be performed with the most significant attributes instead of the all ones.

Figure 4 shows the 41 attributes of NSL-KDD dataset sorted in descending order of their information gain measurement, the latter is evaluated using Equations (2) and (3). The main objective of this experiment is to reduce the number of NSL-KDD attributes by selecting the most relevant ones. This will decrease the space and time complexities required to solve the QP of SSPV-SVDD expressed by Equation (4) and could obtain a tight description of NSL-KDD dataset. By analyzing the histogram, we observe that the IG differs from an attribute to another and the last IGs are practically nulls. We will choose to eliminate the attributes with IG<0.001.

Figure 5 shows the classification accuracy of NSL-KDD dataset using the attributes having IG≥0.001. Also, the figure is divided into two parts: The left side describes the recognition rate of the training dataset. It can be seen that the recognition rate is better than the previous experiment and it reaches 99%. The right side represents the novelty detection capability of our NIDS. It can be observed that the recognition rate grows with σ until a maximum value which outperforms the previous experiment and reaches a maximum of 77.5%. This means that the selection of the most significant attributes of NSL-KDD using information gain metric was performed successfully and have improved the classification accuracy of our NIDS.

5 Conclusions

In this paper, we have presented a new network intrusion detection system based on anomaly detection approach. The proposed system includes: Data transformation where symbolic attributes of network traffics are converted to numeric, normalization operation where the numerical attributes are scaled in a small specified range, relevant attributes selection where information gain method was applied as a measure to estimate the quality of the attributes, and finally a novelty detection model based on SSPV-SVDD as classifier and SMO as solver to decide whether a network traffic is an attack or normal.In contrast to numerous IDSs researchers who use just a small random subset of NSL-KDD in the experimental evalu-

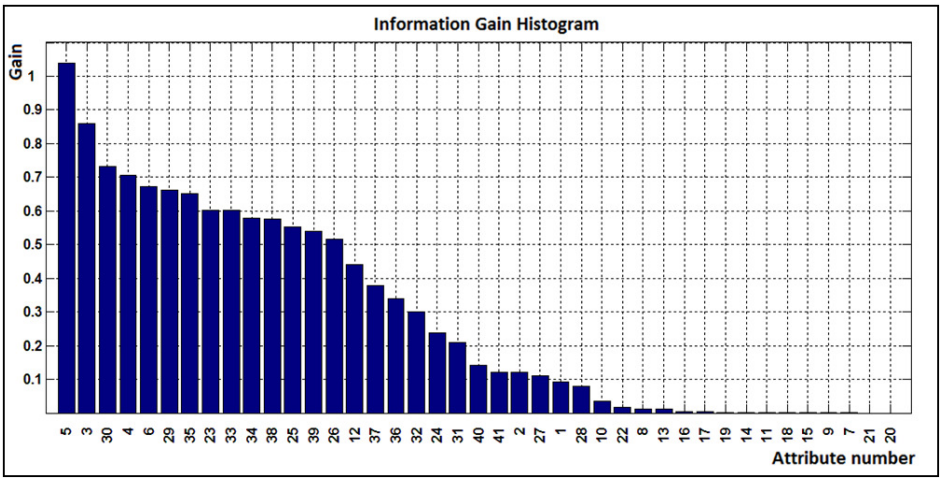


Figure 4: The 41 attributes of NSL-KDD shorted in descending order of IG

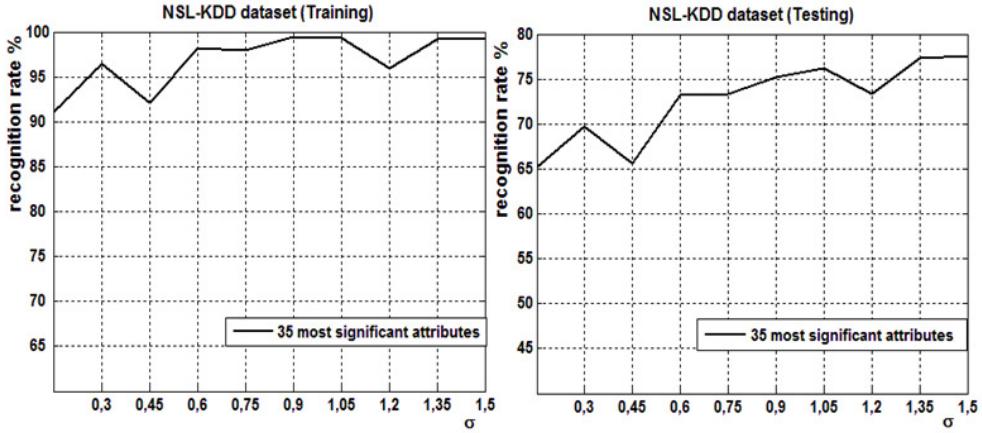


Figure 5: Classification accuracy of NSL-KDD using SSPV-SVDD with the attributes having IG$\geq$ 0.001

ation which gives good but inexact results, in this work we have tested our IDS with the whole NSL-KDD which contains 125973 of samples for training and 22544 samples for testing. The experimental results have shown that with the most significant attributes of NSL-KDD, the proposed IDS can learn 124713 network traffics and can classify successfully 17471 of unknown network behaviors which gives 77.5% of novelty detection rate. This proves that the proposed NIDS is efficient and accurate in detecting different kinds of attacks..

References

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

[2] S. H. An, K. Nam, M. K. Jeong, and Y. R. Choi, "User action-based financial fraud detection method by svdd," *International Journal of Security and Its Applications*, vol. 10, no. 2, pp. 247–254, 2016.

[3] M. El Boujnouni, M. Jedra, and N. Zahid, "A small sphere and parametric volume for support vector domain description," *Journal of Theoretical and Applied Information Technology*, vol. 46, no. 1, pp. 471–478, 2012.

[4] M. El Boujnouni, M. Jedra, and N. Zahid, "New malware detection framework based on n-grams and support vector domain description," in *Proceedings of the* 11^{th} *International Conference on Information Assurance and Security (IAS'15)*, pp. 123–128, Marrakesh, Morocco, Dec 2015.

[5] J. Cao, L. Zhang, B. Wang, F. Li, and J. Yang, "A fast gene selection method for multi-cancer classification using multiple support vector data description," *Journal of Biomedical Informatics*, vol. 53, pp. 381–389, 2015.

[6] L. Duan, M. Xie, T. Bai, and J. Wang, "A new support vector data description method for machinery fault diagnosis with unbalanced datasets," *Expert Systems with Applications*, vol. 64, no. 1, pp. 239–246, 2016.

[7] P. Duong, V. Nguyen, M. Dinh, T. Le, D. Tran, and W. Ma, "Graph-based semi-supervised support vector data description for novelty detection," in *Proceedings of the International Joint Conference on Neural Networks*, pp. 1–6, Killarney, Ireland, July 2015.

[8] C. Gangolf, R. Dochow, G. Schmidt, and T. Tamisier, "Svdd: A proposal for automated credit rating prediction," in *Proceedings of the International Conference on Control, Decision and Information Technologies (CoDIT'14)*, pp. 48 –53, Metz, France, Nov. 2014.

[9] M. GhasemiGol, R. Monsefi, and H. S. Yazdi, "Intrusion detection by ellipsoid boundary," *Journal of Network and Systems Management*, vol. 18, no. 3, pp. 265–282, 2010.

[10] Y. Guerbai, Y. Chibani, and B. Hadjadji, "The effective use of the one-class svm classifier for handwritten signature verification based on writer-independent parameters," *Pattern Recognition*, vol. 48, no. 1, pp. 103–113, 2015.

[11] L. M. Ibrahim, D. T. Basheer, and M. S. Mahmod, "A comparison study for intrusion database (kdd99, nsl-kdd) based on self organization map (som) artificial neural network," *Journal of Engineering Science and Technology*, vol. 8, no. 1, pp. 107–119, 2013.

[12] KDD, *UCI KDD Cup 1999 Data The UCI KDD Archive Information and Computer Science*, University of California Irvine, 1999. (`http://kdd.ics.uci.edu/databases`)

[13] G. Javadzadeh and R. Azmi, "Idufg: Introducing an intrusion detection using hybrid fuzzy genetic approach," *International Journal of Network Security*, vol. 17, no. 6, pp. 754–770, 2015.

[14] M. K. Jeong, S. H. An, and K. Nam, "Svdd-based financial fraud detection method through respective learnings of normal/abnormal behaviors," *International Journal of Security and Its Applications*, vol. 10, no. 3, pp. 429–438, 2016.

[15] R. Chandpa Kalpit and M. Jani Ashwini, "Comparative study between two-class svm and one-class svm classifiers for outlier detection for disease diagnosis," *International Journal of Data Mining And Emerging Technologies*, vol. 5, no. 1, pp. 42–48, 2015.

[16] T. Kenaza, A. Labed, Y. Boulahia, and M. Sebehi, "Adaptive svdd-based learning for false alarm reduction in intrusion detection," in *Proceedings of the* 12^{th} *International Conference on Security and Cryptography*, pp. 405–412, Colmar, Alsace,France, July 2015.

[17] Z. Li, Y. Li, and L. Xu, "Anomaly intrusion detection method based on k-means clustering algorithm with particle swarm optimization," in *Proceeding of the International Conference on Information Technology, Computer Engineering and Management Sciences*, pp. 157–161, Nanjing, Jiangsu, Sept 2011.

[18] Y. Liu, K. Chen, X. Liao, and W. Zhang, "A genetic clustering method for intrusion detection," *Pattern Recognition*, vol. 37, no. 5, pp. 927–924, 2004.

[19] Z. Liu, D. Lin, and F. Guo, "A method for locating digital evidences with outlier detection using support vector machine," *International Journal of Network Security*, vol. 6, no. 3, pp. 301–308, 2008.

[20] S. Mukherjee and N. Sharma, "Intrusion detection using naive bayes classifier with feature reduction," in *Proceeding of the* 2^{nd} *International Conference on Computer, Communication, Control and Information Technology*, pp. 119–128, Hooghly, West Bengal, India, Feb. 2012.

[21] E. U. Opara, O. A. Soluade, "Straddling the next cyber frontier: The empirical analysis on network security, exploits, and vulnerabilities," *International Journal of Electronics and Information Engineering*, vol. 3, no. 1, pp. 10–18, 2015.

[22] S. Pang, S. Li, and J. Xiao, "Application of the algorithm based on the pso and improved svdd for the personal credit rating," *Journal of Financial Engineering*, vol. 1, no. 4, 2014.

[23] J. C. Platt, *Advances in Kernels Methods: Support Vector Learning*. Cambridge, Mass: MIT Press, 1998.

[24] X. Tao, F. Liu, and T. Zhou, "A novel approach to intrusion detection based on support vector data description," in *Proceeding of the* 30^{th} *Annual Conference of IEEE Industrial Electronics Society*, pp. 2016–2021, Busan, South Korea, Nov. 2004.

[25] M. Tavallaee, E. Bagheri, W. Lu, and A. Ghorbani, "A detailed analysis of the kdd cup 99 data set," in *Proceeding of the* 2^{nd} *IEEE Symposium on Computational Intelligence for Security and Defence Applications*, pp. 1–6, Ottawa, Ontario, July 2009.

[26] D. M. J. Tax and R. P. W. Duin, "Support vector domain description," *Pattern Recognition Letters*, vol. 20, no. 11-13, pp. 1191–1199, 1999.

[27] D. M. J. Tax and R. P. W. Duin, "Support vector data description," *Machine Learning*, vol. 54, no. 1, pp. 45–66, 2004.

[28] A. Tayal, N. Mishra and S. Sharma, "Active monitoring & postmortem forensic analysis of network threats: A survey," *International Journal of Electronics and Information Engineering*, vol. 6, no. 1, pp. 49–59, 2017.

[29] G. Wang, J. Hao, J. Ma, and L. Huang, "A new approach to intrusion detection using artificial neural networks and fuzzy clustering," *Expert Systems with Applications*, vol. 37, no. 9, pp. 6225–6232, 2010.

[30] K. Wankhade, S. Patka, and R. Thool, "An overview of intrusion detection based on data mining techniques," in *Proceedings of the International Conference on Communication Systems and Network Technologies*, pp. 626–629, Gwalior, India, Apr. 2013.

[31] H. Wenjun, S. Wang, Y. Liu, F. Chung, and W. Ying, "Privacy preserving and fast decision for novelty detection using support vector data description," *Soft Computing*, vol. 19, no. 5, pp. 1171–1186, 2015.

[32] M. Yang, H. Zhang, J. Fu, and M. Luo, "Anomaly intrusion detection method based on svdd," *Computer Engineering*, vol. 31, no. 3, pp. 39–42, 2005.

[33] M. Yang, H. G. Zhang, J. M. Fu, and F. Yan, *A Framework for Adaptive Anomaly Detection Based on Support Vector Data Description*, LNCS 3222, Springer, 2004.

[34] G. Yin, Y. T. Zhang, Z. N. Li, G. Q. Ren, and H. B. Fan, "Online fault diagnosis method based on incremental support vector data description and extreme learning machine with incremental output structure," *Neurocomputing*, vol. 128, pp. 224–231, 2014.

[35] L. Yuping, L. Weidong, and W. Guoqiang, "An intrusion detection approach using svm and multiple kernel method," *International Journal of Advancements in Computing Technology*, vol. 4, no. 1, pp. 463–469, 2012.

[36] X. Y. Zhang, Z. W. Wei, and X. Lin, "Research on svdd network intrusion detection of the optimal feature selection for particle swarm," *Applied Mechanics and Materials*, vol. 716-717, no. 1, pp. 860–863, 2015.

[37] S. Zheng, H. Tang, Z. Han, and H. Zhang, "Solving large-scale multiclass learning problems via an efficient support vector classifier," *Journal of Systems Engineering and Electronics*, vol. 17, no. 4, pp. 910–915, 2006.

[38] Z. X. Zhou, Y. Jiang, L. T. Ming, M. F. Wang, G. C. Xie, and X. Li, "Improved intrusion detection method based on kernel learning," *Computer Engineering*, vol. 38, no. 14, pp. 21–25, 2012.

Biography

Mohamed El Boujnouni received his Ph.D. degree in Computer Science in July 2015, from Mohammed V University Faculty of Sciences, Rabat, Morocco. His research interests include machine learning, data mining, computational intelligence, and pattern recognition.

Mohamed Jedra holds Doctorat de Troisime Cycle and Doctorat d 'Etat degrees in Electronics Engineering and Informatics; all from Mohammed V University in Rabat, Morocco. From 1990 to 1999, he was the Network and Internet Center Manager in the Faculty of Sciences and Assistant Professor at the Department of Physics. In 1999, he became a Professor Habilit of Informatics in the same Department and in 2003 he was promoted to the position of Professor. From 1987 to the present, he was a member of the Conception and Systems Laboratory. He is the co-founder and the current Chair of the Architecture of Informatic Systems UFR /Formation and Research Unit in the Faculty of Sciences in Rabat since 2003. He is also the co-founder and the current Chair of the Security of Informatic Networks and Embedded Systems Master (ScuRISE) in the same faculty. His main research interests include computational intelligence, pattern recognition and biometric. Mohamed Jedra is a member of IEEE since 1995. He is also a member of IEEE Computer Society, IEEE Computational Intelligence Society and IEEE Signal Processing Society.

Establishing Systems Secure from Research with Implementation in Encryption Algorithms

Mikhail Styugin

Research Department, Siberian State Aerospace University
Department of Applied Mathematics and Computer Security, Siberian Federal University
660014 Krasnoyarsk, Russia
(Email: styugin@gmail.com)
(Received June 29, 2016; revised and accepted Nov. 12, 2016 & Jan. 15, 2017)

Abstract

Systems that have a complex technical implementation usually contain many vulnerabilities which cannot be found at the development stage. Security of complex systems can be improved by protecting them from external research. When the operating algorithm of a system remains concealed then it will be more difficult to compromise the system. The present paper reviews a method of modeling information systems, which allows formalizing the amount of information obtained by a researcher. Two methods of establishing systems protected from research are presented. One method is related to complicating the algorithms and the other one is related to their multiplication. Implementation in encryption systems proves fulfillment of cipher security conditions with their modification. Experimental study of the obtained encryptor demonstrated its effectiveness in protecting from many existing types of attacks aimed at block cipher algorithms.

Keywords: Block Ciphers; Cryptography; Indistinguishability; Researcher Model

1 Introduction

Recently methods and technologies for ensuring security of information systems more and more frequently address the objectives to complicate the system research process for an adversary. Evidently, the less information an attacker obtains the less opportunities it has to compromise the system as well as for unauthorized use of the system.

The current trend is also determined by constant increasing complexity of information systems. The complexity of modern information systems does not allow eliminating all potential vulnerabilities and errors at the design stage. The requirement to release a functionally complete application limits time for testing the completed systems. Whereas for an attacker the time for analyzing an application is nominally unlimited. This creates information asymmetry and requires new solutions to be found in information security, solutions to cover the undetected vulnerabilities and errors. Technologies for protecting information systems from external research are the solutions for those problems. One of the prominent trends in that area is, for example, Moving Target Defense [7] technologies. Recently over 150 different MTD techniques [14] related to LAN security [2], protection from program code injection [9], protection from XSS attacks [15], protection from DDoS attacks [10], etc.

The present paper considers protection of a system from research at the level of algorithms which is implemented in encryption algorithms. This area was chosen because currently requirements to encryption algorithms are the most formalized ones. The problem we have studied is about whether it is possible to formalize a system researcher and to conclude that system research security problem can be solved by using the model obtained.

Security of ciphers themselves is not a new problem. It seems obvious that with an unobservable encryption algorithm it will be more difficult for an adversary to accomplish a ciphertext attack [5, 11]. For instance, papers [6, 12, 19] consider modifications of symmetric encryption algorithms performed by changing rules of permutation and substitution. Paper [20] considers modification of the mode of operation so that the method for presentation of the next cipher block depends on the parameters obtained at the previous step. Cipher modification is also effective in security against side-channel attacks [17] which remain effective also for existing symmetry encryption standards such as AES.

The drawback of all the studied cipher modification solutions is that the cipher variation method is strictly defined. The knowledge of the method simplifies system analysis. A completely research secure system shall not disclose any information related to the methods of cipher text generation.

2 Formalization of a System Researcher

The researcher model is assumed as in paper [18]. The research target remaining to be a black box, the sufficient modifications shall be made as follows.

Take a tuple of three values (x, y, z). Value x stands for the number (share) of observable input values, value y is the observable number of functions of the black box and z is the number of observable output values. A completely observable box is $(1, 1, 1)$, a completely unobservable box is $(0, 0, 0)$. When the black box's output can be observed then it is denoted as $(0, 0, 1)$. Whereas the set of different input values and function values consists of N elements and the system researcher has a function for one of them with which it transforms into the required output value then the box is denoted as $(1/N, 1/N, 1)$.

An encryption system which does not disclose any information can be conceived as two boxes:

$$(0, 0, 1) \rightarrow (1, 0, 0).$$

The first block is denoted as A, hence the second block being reverse to the first one is denoted as A^{-1}. The notation for the obtained system is AA^{-1}.

In order to find the possibilities for researching the system the second level of variables can be introduced, which indicates that the researcher has the data presentation method. For example, one-time pads will have the following notation:

$$(0_1, 0_1, 1) \rightarrow (1, 0_1, 0_1).$$

That implies that even without having the final transformation function we know that it is denoted as $c = m \oplus k$. Hence, having the only one value and one function for one-time pad will enable to disclose all information about the system

$$((1/N)_1, (1/N)_1, 1) = (1, 1, 1).$$

For box $(0, 0, 1)$ in which the function's construction principle is unobservable (for example when the transformation implements a purely random function), then the attempt to obtain information on the only input value and the function will not enable to obtain any additional data:

$$(1/N, 1/N, 1) \neq (u, v, 1), \text{ where } u, v > 1/N.$$

Similarly when the system's function is an instance of a more general functionality then the following tuple can be defined:

$$(0_{0_1}, 0_{0_1}, 1).$$

By the analogy to the encryption in that notation we obtained the algorithm for algorithm generation.

3 Algorithm Research Security by Blurring

When the state of the researched system $(..., 1, ...)$ is unacceptable then the system can be transferred into state $(..., 0_1, ...)$. In case the latter state is also unacceptable then the system can be transferred into state $(..., 0_{0_1}, ...)$. The above procedure shall be called "blurring" of the box's properties.

Absolute ciphers in cryptography are not always a practical structure therefore information obtained by a researcher can be expressed as negligibly small values $\epsilon(n)$, which depend on some parameters as key length n, for example.

A symmetric cipher can be denoted as the following scheme:

$$(\epsilon(n)_1, \epsilon(n)_1, 1) \rightarrow (1, \epsilon(n)_1, \epsilon(n)_1).$$

Algorithm of the above cipher can be "blurred" by defining the algorithm for selecting the encryption algorithm. Then the following scheme is obtained:

$$(..., \epsilon(n)_{\epsilon(n)_1}, 1) \rightarrow (1, \epsilon(n)_{\epsilon(n)_1}, ...).$$

Let us consider an example. Given an AES algorithm, Algorithm 1 shall be used instead of the typical substitution table.

Algorithm 1 Substitution function

1: **Function** SubBytes (t: $0...2^8 - 1$, k:integer)
2: $a = t - 1 \bmod 28$;
3: $b_i = a_{(k+i) \bmod 8} \oplus a_{(k+i+4) \bmod 8} \oplus a_{(k+i+5) \bmod 8} \oplus a_{(k+i+6) \bmod 8} \oplus a_{(k+i+7) \bmod 8} \oplus (k_{(k+i) \bmod 8} \bmod 2^8)$;
4: $result = b$;

The above function substitutes an input value t for value b. It performs the substitution based on key k. Depending on the key value the *SubBytes* function generates 28! (factorial) substitution tables. That number is so high that we can effectively use keys of 128 bits or 512 bits as input with a negligibly small probability that substitution tables may repeat. Hence, now the system has two keys. One of the keys is used for generating substitution tables and the other one is a regular AES key. When the length of the key for selecting substitution tables equals m then upon analyzing the ciphertexts an adversary is not able to distinguish a substitution table with an accuracy greater than a negligibly small value $\epsilon(m)$.

Similarly the cipher can be blurred further by introducing the function for a function's modification:

$$(..., \epsilon(n)_{\epsilon(m)_{\epsilon(h)_1}}, 1) \rightarrow (1, \epsilon(n)_{\epsilon(m)_{\epsilon(h)_1}}, ...).$$

According to the Kerckhoffs' principle [8] a system's operation algorithm shall be open. In our case the principle is maintained, but the adversary's knowledge of the system is always moved to the last level of the scheme (Figure 1).

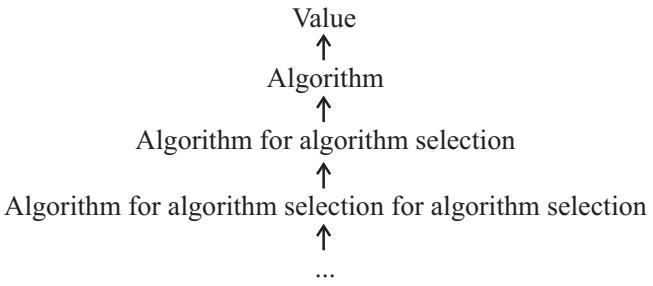


Figure 1: "Blurring" of the box's properties

Following the Kerckhoffs' principle requires that each tuple shall end with 1. Violation of the Kerckhoffs' principle, i.e. when the tuple ends with 0. Then the parameter is defined as absolutely unobservable.

Theorem 1. *When random variables k_1, k_2, ..., k_m with a bit length of $\|k_1\| = n_1$, $\|k_2\| = n_2$, ..., $\|k_m\| = n_m$ are used for creating an encryption algorithm and the cipher scheme is $(..., \epsilon(n_1)_{...\epsilon(n_m)_1}, 1) \rightarrow (1, \epsilon(n_1)_{...\epsilon(n_m)_1}, ...)$, then the probability for distinguishing the input state is less or equals to negligible value $\epsilon(n_1) \cdot \epsilon(n_2) \cdot ... \cdot \epsilon(n_m)$.*

Proof. For proving the above statement we should consider the fact that to find the required algorithm at step q, $\epsilon^{-1}(n_{q+1})$ operations at step $q+1$ must be performed for the algorithm for selecting an algorithm. As operations are performed consecutively, hence the total number of operations would be $\epsilon^{-1}(n_q)\epsilon^{-1}(n_{q+1})$. Thus, using the method of mathematical induction it is proved that the statement, which declares that the probability of distinguishing input in one operation would be less or equal to $\epsilon(n_1) \cdot \epsilon(n_2) \cdot ... \cdot \epsilon(n_m)$, is true. □

4 Research Security by Multiplication of Algorithms

Besides blurring a specific algorithm, multiplication of algorithms can also be implemented. The multiplication would be a sequence of direct and reverse boxes. The previous sections of the paper only simple schemes were considered, which are denoted as A for a separate algorithm and AA^{-1} for encryption systems.

Algorithm multiplication would be a consecutive recording of $ABCD...$, where every following algorithm has the output of the previous algorithm as its input as well as some set of properties. It is assumed that all algorithms are executed in polynomial time.

When one of the algorithms in the sequence has some indistinguishability property, then the whole sequence has the indistinguishability property provided that the parameters must not be reused. The statement can be formalized and proved by separate examples. Below are the examples as applied to cipher area.

Theorem 2. *When encryption scheme AA^{-1} has the indistinguishability parameter with respect to input data, then the parameter is included in any scheme where notation is in the form of $A_1...A_uAA_{u+1}...A_mA_m^{-1}...A_{u+1}^{-1}A^{-1}A_u^{-1}...A_1^{-1}$ provided that the parameters used in A are not used in the other algorithms of the scheme.*

Proof. The indistinguishability definition for encryption algorithms is applied as presented in [13]. An attacker provides a pair of messages m_0 and m_1 of equal length. The encryption algorithm gets a random number of message $b \leftarrow \{0,1\}$ and the message with the number is encrypted $c \leftarrow Enc_k(m_b)$. The obtained ciphertext is sent to the attacker (I) in it shall find the number of the encrypted message b' and in case $b' = b$ then the experiment is considered to be accomplished $PrivK_{I,\Pi}^{eav}(n) = 1$, otherwise $PrivK_{I,\Pi}^{eav}(n) = 0$. The encryption scheme is indistinguishable when there is such a negligibly small function $negl$ for all probabilistic polynomial time attackers I so that the following condition is fulfilled:

$$Pr[PrivK_{I,\Pi}^{eav}(n) = 1] \leq \frac{1}{2} + negl(n).$$

It is evident that operation of algorithms $A_1...A_u$ will have no impact on the indistinguishability condition as their output is input for algorithm A, for which the indistinguishability condition is fulfilled.

For algorithms $A_{u+1}...A_m$ it shall be noted that according to the theorem's conditions they have parameters that are different from the parameters of A. A proof by reduction shall be performed. Assume there are algorithms $A_{u+1}...A_m$ that are such that in conjunction with algorithm A, i.e. $withAA_{u+1}...A_m$ a distinguishable scheme will be obtained for which the following is fulfilled

$$Pr[PrivK_{I,\Pi}^{eav}(n) = 1] > \frac{1}{2} + negl(n).$$

As a result a probabilistic-polynomial time algorithm $A_{u+1}...A_m$ was obtained. The algorithm can distinguish output of algorithm A, which contradicts the indistinguishability condition of algorithm A. □

Similar conditions can be established for the other cipher indistinguishability requirements as well.

Theorem 3. *When encryption scheme AA^{-1} is CPA-secure, then every other scheme that has notation in the form of $A_1...A_uAA_{u+1}...A_mA_m^{-1}...A_{u+1}^{-1}A^{-1}A_u^{-1}...A_1^{-1}$ also has that property on condition that parameters used by A shall not be used by the other algorithms of the scheme.*

Proof. The proof is similar to Theorem 2. □

Theorem 4. *When encryption scheme AA^{-1} is CCA-secure then every other scheme that has notation in the form of $A_1...A_uAA_{u+1}...A_mA_m^{-1}...A_{u+1}^{-1}A^{-1}A_u^{-1}...A_1^{-1}$ also has that property on condition that parameters used by A shall not be used by the other algorithms of the scheme.*

Proof. The proof is similar to Theorem 2. □

Multiplication of cracking difficulty of ciphers in consequent implementation of their algorithms can be considered.

Theorem 5. *When encryption scheme AA^{-1} fulfils the indistinguishability condition which is expressed in the requirement $Pr[PrivK^{eav}_{I,A}(n) = 1] \leq \frac{1}{2} + negl(n)$ and encryption scheme BB^{-1} fulfils the indistinguishability condition which is expressed in the requirement $Pr[PrivK^{eav}_{I,B}(m) = 1] \leq \frac{1}{2} + negl(m)$ then encryption scheme $...A...B...B^{-1}...A^{-1}...$ shall fulfill the indistinguishability requirement expressed as follows $Pr[PrivK^{eav}_{I,...A...B...B^{-1}...A^{-1}...}(n,m) = 1] \leq \frac{1}{2} + negl(n)negl(m)$*

Proof. Proof similar to Theorem 1. □

The section below presents a practical implementation of algorithms with blurring and multiplication.

5 Establishing Research Secure Systems

Let us review at a practical example. Assume that a symmetric encryption scheme should be established for message exchange between two users. It is possible to generate a random sequence of 128 to 2048 bit for one session and provide the random sequence to both users. The classic symmetric encryption scheme requires a strictly set key size and a predefined algorithm. However, we want to establish an encryption system to generate a different algorithm every time by efficiently using the whole random sequence. Then the practical impossibility for an adversary to distinguish either the encrypted message or the algorithm implemented, does not allow performing any computational attacks or side-channel attacks.

Assume that the encoder program can be a $C \in \mathbf{C}$ algorithm. Then the cardinality of set $\mathbf{C}$ has to be large enough to use the provided random sequence. When it is possible to generate a 4096 bit sequence then the number of elements in set $\mathbf{C}$ shall be greater than 2^{4096}. It was demonstrated above that it is easy to accomplish by only correcting the encoder's substitution tables.

A basic requirement to a research secure algorithm. *An adversary shall not be able to compute the implemented $C \in \boldsymbol{C}$ algorithm with accuracy greater than the negligibly small function from the length of the random sequence, which is $negl(u)$.*

In order to create an encryptor that is guaranteed to fulfill the indistinguishability requirements, the cipher's research protection algorithm shall be divided in two operations. The first operation implies consequent multiplication of separate algorithms (boxes) with the predefined features indistinguishable (IND), CPA-secure and CCA-secure. At this stage a transformation sequence with the defined requirements and guaranteed multiplication of difficulty is established. Then each box is blurred with the operations of permutation and substitution, while the difficulty multiplication requirements are no longer applied to them. The encryptor scheme is shown in Figure 2.

The obtained encryption algorithm complies with the requirements of IND-CCA, IND-CPA and IND for adversaries that are able to perform up to $2^{\|k\|}$ operations.

Algorithm 2 is the encryptor mechanism at the pseudocode level. In Theorems 2, 3 and 4 above it was established that for fulfilling the requirements applied to the encryption algorithm the key cannot be reused in every "box". Therefore, function cut is introduced which cuts the random sequence in pieces of specified length as required.

Algorithm 2 Cipher

```
Function cipher (k, m)
Cut ← n bit
i = 0
while ∃k_cur = cut(i, k) do
    m = A_{hash(k_cur) mod m}(k_cur, m)
    i = i + 1
end while
result = m
```

In the above case, algorithm A_j is launched pseudorandomly (by function hash). Each of the algorithms $A_0, ..., A_{m-1}$ is a typical box with the confirmed requirements of IND, CPA and CCA. Thus, function cipher performs multiplication of algorithms as a sequence of boxes $A_1...A_uA_u^{-1}...A_1^{-1}$. Then algorithm blurring is performed in each box to establish research security as shown below

$$(..., \epsilon(n_1)_{...\epsilon(n_m)_1}, 1) \rightarrow (1, \epsilon(n_1)_{...\epsilon(n_m)_1}, ...).$$

The above blurring scheme is demonstrated in the example of the pseudocode in Algorithm 3.

Algorithm 3 An example of the pseudocode

```
Function A_j (k, m)
while hash(k) ≠ const do
    Gen func_sub, k
    Gen func_per, k
    m = Item(func_sub, func_per, k)
    k = func_sub(k)
    k = func_per(k)
end while
result = m
```

The above function runs until $hash(k)$ equals the predefined constant. The constant is defined by experiment, when the number of the algorithm blurring steps can be considered sufficient. The key is used for generating the unique substitution function $func\_sub(k)$ and the unique permutation function $func\_per(k)$. Then a block cipher with the obtained functions is launched. After that the function of substitutions and permutations is applied to the key itself and the cycle is repeated.

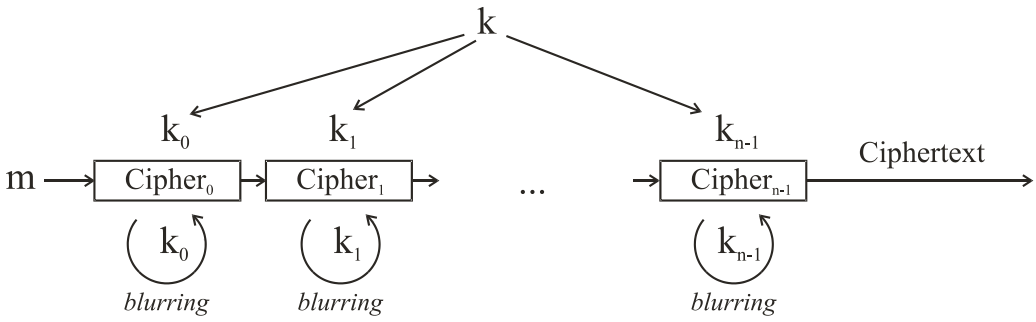


Figure 2: The research secure encryptor scheme

Practical implementation of the research security program was calculated hash functions only once and then it generated the code compliant to the algorithm's functionality to speed up the cipher execution.

6 Experimental Research

In order to perform experimental research of the research secure algorithm, the program mentioned in the previous paragraph has been developed to implement the multiplication algorithm and blurring algorithm. The researched algorithm was a block cipher. Advanced Encryption Standard (AES), International Data Encryption Algorithm (IDEA) and the GOST Russian encryption standard were used as boxes. It should be noted that while substitution and permutation tables are strictly defined in AES and IDEA, in GOST the algorithm blurring procedure can be carried out with these parameters without exceeding the limits of the standard.

All above algorithms comply with the indistinguishability requirements (IND), indistinguishability under chosen-plaintext attack (CPA-secure) and indistinguishability under chosen ciphertext attack (CCA-secure). In order to simplify the block assembly, key length of 128 bit was used as all the provided ciphers can operate with that key length.

A similar function was used as generator of substitution and permutation tables in Algorithm 1 with an input key as the parameter. The range of permutations and substitutions is defined by the type of the block cipher used. The generated cipher was tested for susceptibility to algebraic attacks that decrease the number of enumerating operations [1], algebraic attacks of side channels [13, 16] and differential attacks to decrease cipher cracking difficulty by many orders involving a large volume of memory [3, 4].

The obtained algorithm on a 128 bit random sequence demonstrated failure of all the above attack classes. However, multiplication of cracking difficulty with a longer sequence cannot be tested in practice.

7 Conclusions

The present paper provided a general approach to research security of program algorithms based on two methods. The first method is algorithm blurring by constant shifting the researcher's visibility point. Instead of getting the algorithm, a researcher can only get an algorithm for algorithm generation or an algorithm for generation of an algorithm for generation of an algorithm, etc. That shift enables making the system more complex by introducing additional parameters of randomness or pseudorandomness. The second method is based on algorithm multiplication. Its special feature is that it can be used for expanding any of the indistinguishability properties of its individual components to the whole algorithm.

Theoretic and experimental study of the methods as applied to encryption algorithms demonstrated their effectiveness against many existing attacks aimed at block ciphers and involve algebraic analysis and exploitation of side channels.

Acknowledgments

This study was funded by RFBR according to the research project No. 16-29-09456 ofi_m and Grant of RF President (MK-5025.2016.9).

References

[1] A. Bogdanov, D. Khovratovich, and C. Rechberger, "Biclique cryptanalysis of the full AES," in *Proceedings of The 17th International Conference on the Theory and Application of Cryptology and Information Security*, pp. 344–371, Seoul, Korea, Dec. 2011.

[2] M. Carvalho and R. Ford, "Moving-target defenses for computer networks," *IEEE Security and Privacy*, vol. 12, no. 2, pp. 73–76, 2014.

[3] N. T. Courtois, "An improved differential attack on full gost," Tech. Rep. IACR, 2012.

[4] N. T. Courtois and M. Misztal, "Differential cryptanalysis of gost," Tech. Rep. IACR, 2011.

[5] T. Gulom, "The encryption algorithm GOST28147-89-PES16-2 and GOST28147-89-RFWKPES16-2,"

International Journal of Electronics and Information Engineering, vol. 6, no. 1, pp. 1–11, 2017.

[6] M. I. Husain, K. Courtright, and R. Sridhar, "Lightweight reconfigurable encryption architecture for moving target defense," in *Proceedings of The IEEE Military Communications Conference (MILCOM'13)*, pp. 214–219, San Diego, USA, Nov. 2013.

[7] S. Jajodia, A. K. Ghosh, V. Swarup, C. Wang, and X. S. Wang, *Moving Target Defense. Creating Asymmetric Uncertainty for Cyber Threats*, Advances in Information Security, Springer, 2011.

[8] J. Katz and Y. Lindell, *Introduction to Modern Cryptography: Principles and Protocols*, Boca Raton: Chapman and Hall/CRC, 2007.

[9] P. Larsen, S. Brunthaler, and M. Franz, "Automatic software diversity," *IEEE Security and Privacy*, vol. 13, no. 2, pp. 30–37, 2015.

[10] D. Ma, Z. Xu, and D. Lin, "Defending blind ddos attack on sdn based on moving target defense," *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering*, vol. 152, no. 1, pp. 463–480, 2015.

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

[12] H. Mestiri, F. Kahri, B. Bouallegue, and M. Machhout, "A high-speed aes design resistant to fault injection attacks," *Microprocessors and Microsystems*, vol. 41, no. 1, pp. 47–55, 2016.

[13] Mohamed Saied Emam Mohamed, Stanislav Bulygin, Michael Zohner, Annelie Heuser, and Michael Walter, "Improved algebraic side-channel attack on AES," *Journal of Cryptographic Engineering*, vol. 3, no. 3, pp. 139–156, 2014.

[14] H. Okhravi, T. Hobson, D. Bigelow, and W. Streilein, "Finding focus in the blur of moving-target techniques," *IEEE Security and Privacy*, vol. 12, no. 2, pp. 16–26, 2014.

[15] J. Portner, J. Kerr, and B. Chu, "Moving target defense against cross-site scripting attacks," *Lecture Notes in Computer Science*, vol. 8930, pp. 85–91, 2015.

[16] M. Renauld, F. X. Standaert, and N. V. Charvillon, "Algebraic side-channel attacks on the AES: Why time also matters in DPA," *Lecture Notes in Computer Science*, vol. 5747, pp. 97–111, 2009.

[17] W. Shan, L. Shi, X. Fu, X. Zhang, C. Tian, Z. Xu, J. Yang, and J. J. Li, "A side-channel analysis resistant reconfigurable cryptographic coprocessor supporting multiple block cipher algorithms," in *Proceedings of The 51st Annual Design Automation Conference*, pp. 1–6, San Francisco, United States, June 2014.

[18] M. Styugin, "Protection against system research," *Cybernetics and Systems: An International Journal*, vol. 45, no. 4, pp. 362–372, 2014.

[19] Y. Wang, L. Wang, R. Yao, Z. Zhang, and C. Jiang, "Dynamically reconfigurable encryption system of the AES," *Cryptography. Wuhan University Journal of Natural Sciences*, vol. 11, no. 6, pp. 1569–1572, 2006.

[20] P. Zacek, R. Jasek, and D. Malanik, "Using the deterministic chaos in variable mode of operation of block ciphers," *Artificial Intelligence Perspectives and Applications. Series Advances in Intelligent Systems and Computing*, vol. 347, no. 1, pp. 347–354, 2015.

Biography

Mikhail Styugin is a senior lecturer at Siberian Federal University and a scientist at Siberian State Aerospace University (Krasnoyarsk, Russia). He holds a PhD degree in computer science. He conducts research in area of information security system and technologies of information warfare. He owns two companies that develop solutions in area of information security system in the Internet .

A Hybrid Intrusion Detection System: Integrating Hybrid Feature Selection Approach with Heterogeneous Ensemble of Intelligent Classifiers

Amrita, Kiran Kumar Ravulakollu
(Corresponding author: Amrita)

Department of Computer Science and Engineering, Sharda University
SET, Plot No. 32-34, Knowledge Park III, Greater Noida, Uttar Pradesh 201306, India
(Email: amrita.prasad@sharda.ac.in)
(Received Aug. 9, 2016; revised and accepted Nov. 21 & Dec. 23, 2016)

Abstract

This paper proposes Hybrid Feature Selection Approach – Heterogeneous Ensemble of Intelligent Classifiers (HyFSA-HEIC) for intelligent lightweight network intrusion detection system (NIDS). The purpose is to classify for anomaly from the incoming traffic. This system hierarchically integrates HyFSA and HEIC. The HyFSA will obtain the optimal number of features and then HEIC is built using these optimal features. HyFSA helps to decrease the computation time of the system and make it lightweight to work in real time. The aim of HEIC is to obtain accurate and robust classifier and enhance overall performance of the system. The results demonstrate that proposed system outperforms other ensemble and single classifier methods used in this paper. It has true positive rate (99.9%), accuracy (99.91%), precision (99.9%), receiver operating characteristics (99.9%), low false positive rate (0.1%) and lower root mean square error rate (3.06%) with a minimum number of selected 6 features. It also reduces time to build and time to test the model by 50.79% and 55.30% respectively on reduced features set. The results evince that detection rate, accuracy and precision of the system is increased by incorporating feature selection approach with heterogeneous ensemble of intelligent classifiers and significantly reduce the computation time.

Keywords: Classifier; Ensemble; Feature Selection; Network Intrusion Detection System

1 Introduction

Information security has become an essential key component in all areas with the increasing Internet connectivity and traffic volume. Also the security of networks plays a vital role in information security [27]. Therefore, Intrusion Detection Systems (IDS) for network have become an essential component for information security to protect it from continuous increase of network-based intrusions and attacks. The role of Network Intrusion Detection Systems (NIDS) is to actively monitor the function of network and detect malicious activities in real time and raise an alert. Network intrusion detection is a classification task that is capable of distinguishing between normal and attack or intrusion traffic connection i.e. two-class problem and further classification of the different attacks type. Also NIDS has to examine huge amount of data with high dimensional network traffic in real time and on-line processing. Therefore, it is necessary to build accurate, intelligent and lightweight NIDS to protect networks as well as information system.

In high dimensional feature set, some features may be redundant and some others may be irrelevant. These redundant and irrelevant features can increase the computation time and degrade the performance and accuracy of NIDS. For this reason feature selection method is used as a pre-processing step to obtain the subset of relevant features to construct NIDS. It selects the minimal cardinality feature subset that maintains the detection rate and accuracy as the original feature set. For classification, the main issue is to select the best classification method as every method has its own advantages and disadvantages [11]. Therefore heterogeneous ensemble of intelligent classifiers has been proposed to overcome the limitation of single classification method. Ensemble method exploits the strength of each classifier of the ensemble to acquire accurate and robust classifier. It combines different classification method to improve the performance and accuracy of the model.

The ultimate goal of NIDS is to achieve best possible detection accuracy and reliability. It can be achieved by

combining different decision-making model into one system. This leads to the design of hybrid system. The hybrid system integrates different decision-making models or learning techniques to boost the performance of the system than the individual decision making or learning technique. The primarily focus on the design of hybrid system is integration and interaction of different learning techniques covering computational phases from data preprocessing up to final decision making.

In an attempt to develop lightweight and efficient anomaly based NIDS for two-class classifications, i.e., intrusions or attack and normal, a novel hybrid system for network intrusion detection, HyFSA-HEIC (Hybrid Feature Selection Approach – Heterogeneous Ensemble of Intelligent Classifiers) is proposed. The HyFSA-HEIC integrates hybrid feature selection approach (HyFSA) and heterogeneous ensemble of intelligent classifiers (HEIC) to make it lightweight and accurate. HyFSA proposed in [3] has been used for the selection of optimal features set. To increase the performance of NIDS, HEIC has been developed. The performance of HyFSA-HEIC has been tested on non-redundant datasets of "10% KDD" and "Corrected Test". True Positive Rate (TPR), Precision (PRE), Accuracy (ACC), False Positive Rate (FPR), Receiver Operating Characteristics (ROC), Time-span to build model (TBM), Time-span to test model (TTM) and Root mean squared error (RMSE) have been used as performance evaluation metrics.

The paper is organized as follows. Review of related work is presented in Section 2. Section 3 introduces feature selection approaches, classification methods and ensembles employed in this work. Performance evaluation measures utilized in this paper are discussed in Section 4. The proposed system and experimental setup adopted in this paper is presented in Section 5, experimental results and analysis in Section 6, and conclusion and future direction in Section 7.

2 Related Works

Many hybrid based systems or approaches have been proposed and investigated in the literature to enhance the accuracy of the IDS in the recent years. Each technique has its own strength and weakness. As well as performance of each technique is varies in terms of Detection Rate Accuracy, Precision, False Positive Rate and error rate. Panda et al. [23] investigated novel hybrid intelligent technologies for making intelligent decision which combines the supervised or unsupervised with classifier using data filtering to detect network attacks. The approach was evaluated on NSL-KDD dataset for 2-class classification and has 99.9% detection rate with 0.06% error. Agarwal et al. [1] presented a hybrid approach which combines entropy and Support Vector Machine (SVM) for anomaly network traffic detection system. The Hybrid method outperforms the single method in terms of accuracy with 97.25% and misclassified instances of 2.75%. Chitrakar et al. [7] proposed two hybrid approaches for anomaly intrusion detection. In first approach, k-medoids clustering is combined with classifier Naïve Bayes and in second, k-medoids clustering is combined with classifier Support Vector Machine. These approaches show enhancement in the TPR and reduction in the FPR.

Sindhu et al. [26] proposed a lightweight Network IDS employing a wrapper based feature selection approach with neural ensemble of decision trees to maximize the specificity and sensitivity. The average classification rate and error of the proposed system with 16 selected features is 98.4% and 1.62% and performed better than C4.5, Naïve Bayes, Decision Stump, REP tree, Random Tree and Random Forest. A reliable and efficient IDS based on gradually feature removal method with the combination of k-mean, ant colony algorithm and support vector machine has been developed by Li et al. [20] for normal or attack detection in network. The system was evaluated on KDD Cup 99 data set. Nineteen features were selected by applying gradually feature removal method with accuracy of 98.62%. Lin et al. [21] combined support vector machine, decision tree and simulated annealing for anomaly based intrusion detection. In this, support vector machine and simulated annealing were used to obtain the best selected features using KDD dataset and decision tree and simulated annealing were used to find decision rules for new attacks which can enhance the accuracy of the method. The performance of the proposed algorithm outperforms other existing approaches with weighted average TPR and FPR of 98.3% and 1.4% respectively.

A new hybrid intrusion detection method which hierarchically integrates a misuse detection model and an anomaly detection model was proposed by Kim et al. [16]. The C4.5 decision tree algorithm has been used to build the misuse detection model. This model is then used to decompose normal training data into smaller subsets. The one-class support vector machine (SVM) is used to build the anomaly detection model for each decomposed subset. NSL-KDD dataset has been used to evaluate the proposed method. The experimental results demonstrated that method was better in term of detection rate for both known and unknown attacks and reduced the training and testing time of the model. A hybrid approach to anomaly detection using a real-valued negative selection based detector generation in the large scale dataset is presented in [13]. It uses k-mean clustering to reduce the size of the training dataset to identify good starting points for the detector generation based on multi-start metaheuristic method and genetic algorithm. The results showed that this approach outperforms other techniques by 96.1% accuracy with time of 152 s and low false positive rate of 0.033. Vahid Golmah [14] developed a hybrid method to improve the accuracy of the intrusion detection system based on C5.0 and SVM. The proposed method is evaluated on benchmark "KDD Cup 1999" dataset with full feature set. The average precision of classification for the proposed algorithm is 99.96%.

3 Related Background

3.1 Feature Selection Method

Feature selection method is commonly used to find the optimal feature subset to improve the system performance by eliminating redundant and irrelevant features from dataset. It also helps to alleviate "the curse of dimensionality". There are three approaches—*filter, wrapper* and *hybrid* for feature selection [4]. Filter approach [22] utilizes external classifier to assess the performance of selected features. The wrapper approach [17] "wrap around" the predetermined classifier to assess subset of features. This method is computationally more expensive than the filter method [9, 17]. The hybrid approach [9] combines the filter and wrapper approach to achieve the best possible performance with a specific classifier. In this work, HyFSA [3] has been used for the selection of optimal features set. A survey of several works on feature selection approaches applied on "KDD Cup 1999" dataset for IDS is presented in [2].

3.2 Classification Methods

The base classifiers selected for the ensemble are probability theory based Naïve Bayes (NB) [31], decision tree based C4.5 [24], homogeneous ensemble of decision trees based random forest (RF) [6], soft computing based Neural networks using Stochastic Gradient Descent (NN-SGD) [5], instance based k-Nearest Neighbor (kNN) [28], and rule based Repeated Incremental Pruning to Produce Error Reduction (RIPPER) [8]. These base classifiers are briefly discussed below.

3.2.1 Naïve Bayes Classifier

Naïve Bayes (NB) classifier [31] is a supervised classifier. It is based on Bayes' theorem. This classifier computes the posterior probability for each class c_j to classify an input pattern $\vec{x}_i$ and assigns the target class c* with the highest posterior probability to $\vec{x}_i$ using Equation (2). The output of the individual classifiers as a posteriori probability can be represented as $P(c_j|\vec{x}_i)$, where $\vec{x}_i$ is presented to i^{th} classifier and assigned to class c_j. For two-class classification, the posteriori probability using Bayes theorem can be calculated as

$$P(c_j|\vec{x}_i) = \frac{P(\vec{x}_i|c_j)P(c_j)}{P(\vec{x}_i)} = \frac{P(\vec{x}_i|c_j)P(c_j)}{P(\vec{x}_i|c_1)P(c_1) + P(\vec{x}_i|c_2)P(c_2)} \quad (1)$$

where j=1,2;i=1,...,L

$$c^* = arg\ max P(c_j|\vec{x}_i) \quad (2)$$

where $P(c_j|\vec{x}_i), P(c_j), P(\vec{x}_i|c_j)$ *and* $P(\vec{x}_i)$ are called the posterior probability, prior probability, likelihood, and evidence respectively. Naïve Bayes classifier can work on symbolic as well as numerical features. It exhibits high speed and accuracy and highly suitable for high dimensional large dataset [31].

3.2.2 Decision Tree

Decision tree (DT) is a supervised learning algorithm based on tree-like structure which consists of nodes and branches. Each non-terminal node represents a test on an attribute, each branch represents the outcome of the test and each leaf node represents the class label for classification of the input pattern. The classification of input pattern starts from the root node and follow the branch to reach the leaf node of the DT. A well-known algorithm for constructing decision tree is C4.5 [24]. The C4.5 algorithm is also very robust for high dimensional data and handling missing data. It works well on both numerical and symbolic features.

3.2.3 Random Forest

Random Forest (RF) [6] is an ensemble based classification techniques. It generates many unpruned decision tree by inducing different bootstrap sample using random feature selection from training dataset. It is called random forest as sampling of records is done randomly and forest of decision trees are built in the process. Final class of an input instance is made by aggregating the decisions of the individual trees in the forest by majority voting for classification. Random forest works efficiently on high dimensional large dataset and able to deal with unbalanced and missing data.

3.2.4 Neural Network

Neural network or Artificial Neural network (ANN) is computational technique that mimics the neurons of human brain. It consists of set of simple processing components called artificial neurons that are interconnected to other neuron by synapses or link. These neurons are organized into network in many ways known as topologies. The neurons in ANN are grouped into layers as input, output and hidden layer. Each link is associated with weight. Neurons in the input layer receive stimuli from outside the network, transform it into output and pass the output to the subsequent hidden or output layer. The network learns by adjusting the corresponding weight of the link in the learning phase. It helps ANN to assign the correct class label to the given input pattern. Most commonly used ANN architecture is Stochastic Gradient Descent (SGD) [5] for large scale dataset and online learning.

3.2.5 K-Nearest Neighbor

K-nearest neighbor (kNN) [28] is supervised classification method. It is simple, non-parametric, instance-based and lazy learning algorithm. Lazy learning signifies that the algorithm does not build the model until the time classification is required. This algorithm classifies the new input

pattern by calculating the similarity measure (e.g. distance function) between the new input pattern and each instance of training dataset and then uses the class labels of the k most similar neighbors to assign the class of new input pattern based on majority voting. Euclidean distance or the cosign value can be used as similarity measures to calculate the similarity between two instances. The performance of this classifier relies on the value of k.

3.2.6 RIPPER

Repeated Incremental Pruning to Produce Error Reduction (RIPPER) [8] is supervised rule-based learning algorithm. This is an extension of IREP (Incremental Reduced Error Pruning). It improved the efficiency of IREP by reducing errors by applying repeated pruning, faster training time, support missing attributes and noisy datasets. This learning algorithm searches the feature set of the training dataset and produces concise rule-sets for each class label. It works efficiently on numerical and large dataset.

3.3 Ensemble of Classifiers

An ensemble of classifiers combines multiple weak or diverse classifiers whose individual outputs are combined in some means to form a final decision [10]. The combined decisions of an ensemble generally provide better performance than the individual classifiers [25]. The main motivation of using ensemble of classifiers is to get better accuracy of the complex problem by exploiting the strengths of individual classifiers with the aim to obtain the best possible collective decision accuracy than any of the individual classifiers. Ensemble of multiple diverse classifiers has more reliable and better decision than single classifier as it reduces the chance of incorrect classification done by single classifier and also overcome the limitation of single classifier. The architectures of the ensemble of classifiers are mainly categorized into two types, i.e., parallel and serial. There are two steps for constructing an ensemble: (1) generating the base learning algorithms or classifiers, and (2) combining the decisions of base learning algorithms for maximum accuracy.

3.3.1 Generating Base Classifiers

In this step, individual classifier of the ensemble known as base classifier is generated. Methods for generating ensemble can be categorized as Homogenous and Heterogeneous ensemble. Homogeneous ensemble can be generated from the different executions of the same classifier. This ensemble can be generated by using any method from (i) different subset of training data with same classifier, (ii) different set of input training parameters available with a single classifier, (iii) different feature sets, (iv) multi-class specialized systems, or (v) manipulation of output labels. Examples are Bagging, Boosting, Option Trees, Error-correcting output codes. Heterogeneous ensemble uses different learning algorithm or classifiers on the same data set. Different methods for generating heterogeneous ensemble are: (i) Voting or fixed-rule aggregation, and (ii) Stacked generalization or meta-learning.

The classifiers in ensemble should be accurate and diverse to improve the performance of ensemble system over single classifier. Therefore, classifiers must be highly accurate and diverse to build efficient and accurate ensemble [19]. Weak classifiers in the ensemble also result in weak ensemble and hence deteriorate the accuracy of ensemble. The classifiers are said to be diverse or unique if they make distinct errors on distinct instances and combining their outputs can decrease the total error. Also diverse classifiers contribute toward uncorrelated decisions and improve the overall accuracy of the ensemble system. There are many methods to achieve classifier diversity [18]. Any method of homogeneous or heterogeneous described in section 3.3.1 can be used to generate diversity among the classifiers.

3.3.2 Combining Classifiers

The second step in ensemble method is the approach employed in combining the decision of the classifiers. The main motivation for combining multiple classifiers is to obtain a consensus decision by combining the individual decision of classifiers [18]. There are mainly two approaches in combining the decisions of different classifiers—*classifier selection* and *classifier fusion* [18]. The classifier selection approach selects a single classifier to give the final decision for a new instance while classifier fusion approach combines the decision of all classifiers. The various combination methods have been reported in [18]. The most commonly used method is elementary combiners based on algebraic combination rules. This method combines the decisions of classifiers that can be expressed as a posteriori probability. The major benefit of using this method is its simplicity as it does not need any training. It includes several methods as *Sum, Average, Product, Majority Voting, Minimum*, and *Maximum* rules.

Let $\{D_1, D_2, ..., D_L\}$ be the set of L individual classifiers and $\{c_1, c_2, ..., c_m\}$ be the set of m possible class labels. The combiner combines the decisions of all D_i to predict the final class label for the input instance $\vec{x}_i$. In order to employ the *Sum, Average, Product, Majority Voting, Minimum*, and *Maximum* methods, outputs of all D_i can be viewed as a posteriori probabilities using Bayes theorem defined in Equation (1). Let input instance $\vec{x}_i$ is finally assigned to class c, where c is one of the m possible classes. The *Sum, Average, Product, Majority Voting, Minimum,* and *Maximum* methods can be used to determine c are defined as follows:

$$\text{Sum Rule: } c = max_{j=1...m} \sum_{i=1}^{L} P(c_j|\vec{x}_i)$$

$$\text{Average Rule: } c = max_{j=1...m} \frac{1}{\mathrm{L}} \sum_{i=1}^{L} P(c_j|\vec{x}_i)$$

$$\text{Product Rule: } c = max_{j=1...m} \prod_{i=1}^{L} P(c_j|\vec{x}_i)$$

$$\text{Majority Voting Rule: } c = max_{j=1...m} \sum_{i=1}^{L} \Delta_{ji} \quad (3)$$

$$\text{Minimum Rule: } c = max_{j=1..m} min_{i=1..L} P(c_j|\vec{x}_i)$$

$$\text{Maximum Rule: } c = max_{j=1..m} max_{i=1..L} P(c_j|\vec{x}_i)$$

In the decision rule in Equation (3), $\Delta_{ji} = 1\ if\ P(c_j|\vec{x}_i) = max_{j=1..m} P(c_j|\vec{x}_i)$ and zero otherwise.

4 Performance Evaluation Measures

Several performance evaluation methods are employed to assess the accuracy and efficiency of the HyFSA-HEIC, classifiers and for comparison. The performance evaluation methods are *True Positive (TP), False Positive (FP), True Negative (TN), False Negative (FN), Receiver Operating Characteristics (ROC) or Area Under Curve (AUC), Time-span to Build the Model (TBM), Time-span to Test the Model (TTM)* and as follows:

$$True\ positive\ rate(TPR) or\ Recall(R)\ or$$

$$Sensitivity\ or\ Detection\ rate(DR) = \frac{TP}{TP+FN}$$

$$False\ positive\ rate(FPR) = \frac{FP}{FP+TN}$$

$$Accuracy\ (ACC) = \frac{TP+TN}{TP+TN+FP+FN}$$

$$Precision\ (PRE) = \frac{TP}{TP+FP}$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (P_i - T_i)^2} \quad (4)$$

where T_i is the true value, P_i is the prediction, and N is the number of observations in Equation (4).

5 Proposed System and Experimental Setup

The aim of this paper is to propose a Hybrid Feature Selection Approach – Heterogeneous Ensemble of Intelligent Classifiers (HyFSA-HEIC), a hybrid system for network intrusion detection for the classification of coming input pattern into either normal or intrusion. This system must be accurate, lightweight, low false positive rate, high detection rate and able to work in real time. The HyFSA-HEIC integrates the hybrid feature selection approach (HyFSA) with heterogeneous ensemble of intelligent classifiers (HEIC). This is a hierarchical system in which HyFSA selects the optimal feature set for the classification of normal or attack pattern. Then selected optimal feature set is provided as an input to next layer i.e. HEIC for final decision. The overall accuracy of the system relies on the accurate functionality of all layers in the system. Weka 3.7.13 [15] is used as tool for classifiers, feature selection approaches and ensembles utilized in this paper. Figure 1 illustrates the architecture of the HyFSA-HEIC. It contains following five phases:

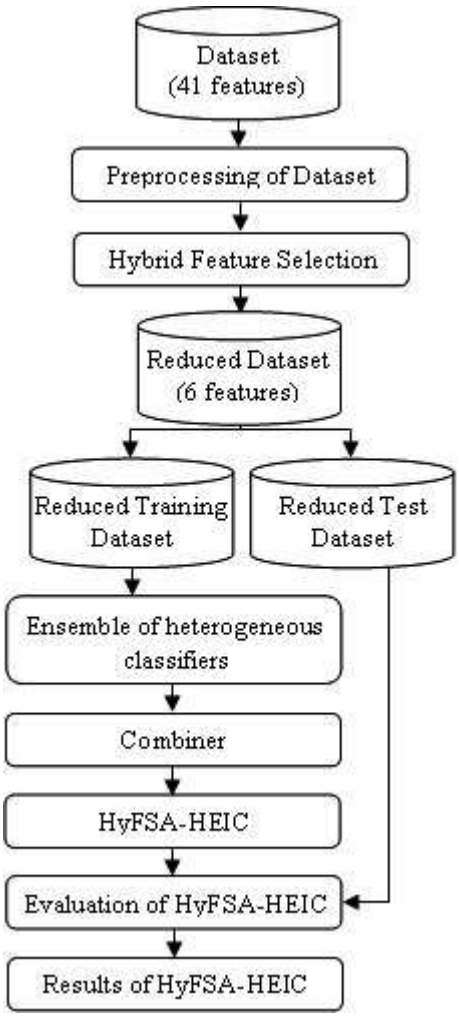


Figure 1: Framework of proposed HyFSA-HEIC

Phase I: Construction of dataset;

Phase II: Feature Selection;

Phase III: Selection of base classifiers for ensemble;

Phase IV: Ensemble and combiner method;

Phase V: Evaluation of classifiers and ensembles.

5.1 Phase I: Construction of Dataset

The "KDD Cup 1999" dataset [29] is the benchmark dataset for intrusion detection and derived from DARPA 1998 dataset. It is the most widely used comprehensive dataset used by many researchers for NIDS. "KDD

Cup 1999" is comprised of three independent datasets—"Whole KDD", "10% KDD" and "Corrected Test". The "10% KDD" has 494,021 connection records in which 97277 are normal and 396744 are attack whereas "Corrected Test" has 311,029 connection records in which 60,593 are normal and 250,436 are attacks shown in Table 1. Each connection record has 41 features (32 numerical and 9 categorical) numbered in an order of 1,2,3,4....,41 plus one class label. Each connection record has label of either normal or attack, with a specific kind of attack type. The attacks fall into one of the four main classes—Probe, User to Root (U2R), Remote to Local (R2L), and Denial of Service (DoS). The "Whole KDD" dataset consists of 22 attack types and "Corrected Test" dataset includes 17 additional attack types and hence total 39 attack types.

NSL-KDD dataset [30] is another widely used dataset for anomaly detection in NIDS. This dataset is refined version of "KDD Cup 1999" dataset. It consists of selected and non-redundant connection records with same number of features of "KDD Cup 1999" dataset. Data cleaning step of data preprocessing can be skipped by using this dataset. Data preprocessing is an important step in any decision making system. Therefore "KDD Cup 1999" dataset has been used as an experimental dataset to make a complete system from preprocessing step to final decision making step and more suitable for real time and on-line processing.

The "10% KDD" and "Corrected Test" datasets are selected as experimental dataset and preprocessed for binary class classification (normal or attack). This phase contains 4 steps. **1) Data transforming:** For binary class classification, the label must have either normal or attack for each connection. Therefore label of each connection for all types of attack are transformed into label "attack". **2) Removal of redundant records:** The converted binary class dataset also consists of large number of redundant connection records. These redundant records will cause classifiers to be influenced towards redundant records in the training dataset. It will also influence the performance of the learning algorithms. The "10% KDD" and "Corrected Test" datasets contain around 70% and 75% redundant records respectively in which the attack class has most of the redundant records than normal class. The resultant datasets are named as "Unique 10% KDD" and "Uni Corr Test". **3) Discretization of dataset:** Feature selection approaches employed in this paper work on discrete data and "10% KDD" dataset has 32 numerical features. Therefore, discretization approach presented by [12] based on Entropy Minimization is utilized. The "Dis_Unique 10% KDD" is resultant discretized dataset. **4) Construction of training and test dataset:** "Unique 10% KDD" is equally partitioned into two datasets: the training dataset ("Uni Train") and test dataset ("Uni Test") for all 41 features. Each dataset contains 72793 records in which each class comprises 50% of the data of "Unique 10% KDD". Reduced training dataset ("Red Uni Train") and reduced test dataset ("Red Uni Test") for selected 6 features in phase II are created from "Uni Train" and "Uni Test" datasets respectively. "Uni Corr Test" and "Reduced Uni Corr Test" are also employed as another test datasets for all 41 and reduced 6 features respectively. Training datasets are utilized to build and test datasets are utilized to assess the performance of the classifiers and ensembles. Table 1 illustrates the statistics of the records for normal and attack in "10% KDD", "Unique 10% KDD", "Corrected Test" and "Uni Corr Test" datasets respectively.

5.2 Phase II: Feature Selection

The accuracy and efficiency of the IDS also depends on the dimension of the dataset. Hybrid method for feature selection proposed in [3] has been used to obtain the optimal number of features for binary (normal or attack) classification. This method employs fusion of filter based feature selection approaches and wrapper method using Naïve Bayes classifier. Filter based feature selection methods used for fusion are consistency-based feature selection (CON), gain ratio (GR), correlation-based feature selection (CFS) and information gain (IG). First, the common features are selected from the feature subsets obtained by applying CFS and CON with best first search. Similarly, the common features are selected from the feature subsets obtained by applying IG and GR. Then, initial feature subset is created by adding these two common feature subsets. Another left feature set is created by adding the remaining features from the sets of CFS, CON, IG and GR. Wrapper based feature selection method employed Naïve Bayes classifier is further applied to obtain the final optimal feature subset. Linear Forward selection (LFS) is used in wrapper method. It starts with the initial feature subset and then add feature one by one from the left feature set until there is no change in the performance of current feature subset. This feature selection approach selects the relevant features and removed redundant and irrelevant features. Finally, 6 best features are selected from 41 features by employing Hybrid feature selection approach [3] on "Dis_Unique 10% KDD" dataset consisting of 41 features. These feature's name and number are {Service-3, Src-bytes-5, Dst-bytes-6, Hot-10, Num-compromised-13, Same-srv-rate-29}. For detail procedure for feature selection and performance of selected features, refer [3].

5.3 Phase III: Selection of Base Classifiers for Ensemble

In HyFSA-HEIC, a novel heterogeneous ensemble is presented that combines the decisions of diverse and accurate learning algorithms or classifiers to the problem of normal or attack detection in IDS. The main issues in the ensemble technique are accuracy and diversity of individual classifiers in ensemble. Different learning algorithms or classifiers for constructing an ensemble enforce a high level of diversity [18]. This ensemble has benefit of differ-

Table 1: Instances and percentages of division of normal and attack in "10% KDD", "Unique 10% KDD", "Corrected Test" and "Uni Corr Test" datasets for all 41 and 6 features.

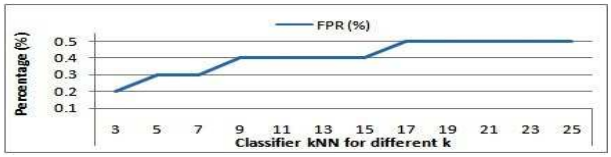
	10% KDD		Unique 10% KDD		Corrected Test		Uni Corr Test	
Type	**#Instance**	**(%)**	**#Instance**	**(%)**	**#Instance**	**(%)**	**#Instance**	**(%)**
Normal	97277	19.69	87832	60.33	60,593	19.48	47,913	61.99
Attack	396744	80.31	57754	39.67	250,436	80.52	29,373	38.01
Total	494021	100	145,586	100	311,029	100	77,286	100

ent biases for each individual classifier and also reduces the bias that can be occurred in a single learning algorithm or classifier. Also combining weak diverse classifiers in the ensemble will result in weak ensemble and lead to deteriorate the accuracy of the ensemble. Therefore, different classifiers are compared to select accurate and diverse base classifiers. The base classifiers selected for the ensemble are C4.5, Naïve Bayes (NB), Neural networks using mini-batch stochastic gradient descent (NN-SGD), K-nearest neighbor (kNN), Repeated Incremental Pruning to Produce Error Reduction (RIPPER) and random forest (RF). Motivation of selecting these different types of base classifiers leads diversity in creating ensemble classifiers. Each selected classifier has different learning hypotheses (trees, instance-based, rules and statistics) and also different inductive bias that make diverse set of classifiers for ensemble. Different learning hypotheses and inductive bias generates diversity among the classifiers. The base classifiers are briefly discussed in Section 3.2.

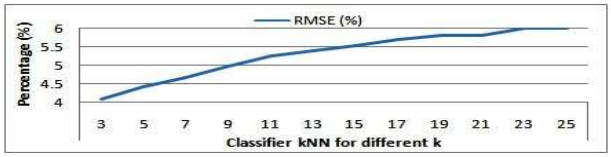
The performance of kNN classifier is influenced by the suitable selection of optimal value of parameter k. In order to select the optimal value of k for kNN classifier, 10-fold cross validation is performed on "Red Uni Train" dataset for 6 features. The value of k is varied from 1 to 25 for odd number for two-class classification to avoid tied votes. The empirical results are shown in Table 2. The value of k for kNN is selected based on the value of k which minimizes errors and maximizes predictive accuracy of the kNN classifier. As can be seen in Table 2, the kNN for $k = 3$ outperformed among all k except $k = 1$. The kNN for $k = 1$ cannot be considered in the selection of k as it badly overfits the classifier. Therefore, three nearest neighbours ($k = 3$) is selected for each instance in kNN classifier. Figure 2 shows performance comparison of kNN for different values of k in terms of FPR, RMSE and ACC respectively. Euclidean Distance as similarity measure has been used to find the nearest neighbours. Euclidean distance $d(u, v)$ between two instances u and v is defined as $d(u, v) = \sqrt{\sum_{i=1}^{N} \delta_i^2}$, where δ_i is the difference between the i^{th} feature's value of the instances u and v and N is the dimension of the dataset. The difference δ_i can be measured for numerical as well as for nominal feature as

$$\delta_i = \begin{cases} x_i - y_i, & for\ numerical\ feature \\ 0, & if\ x_i = y_i\ for\ nominal\ feature \\ 1, & if\ x_i \neq y_i\ for\ nominal\ feature \end{cases} \quad (5)$$

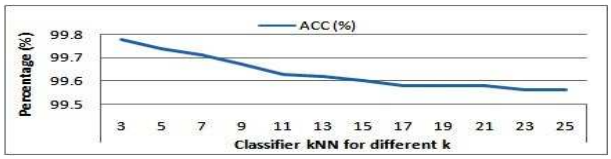
These classifiers are trained on "Uni Train" training dataset. Performance metrics used in the comparisons are TPR, FPR, ACC, PRE, ROC, TBM, TTM and RMSE. The results of these classifiers using 6 selected features and all 41 features based on different performance metrics on training datasets ("Red Uni Train" and "Uni Train") are depicted in Table 3.



(a) FPR



(b) RMSE



(c) ACC

Figure 2: Performance comparison of kNN for different k in terms of (a) FPR, (b) RMSE and (c) ACC

5.4 Phase IV: Ensemble and Combiner Method

A heterogeneous ensemble of classifiers is a collection of multiple diverse classifiers. In this, decision of individual classifiers is combined to classify input instance. A heterogeneous ensemble of classifiers will combine the strength and disagreement of all diverse classifiers and also make individual classifier disagree with each other. The strength and disagreement among the diverse classifiers are utilized by elementary combiners based on algebraic combination rules to give accurate and reliable final decision. To construct the heterogeneous ensemble, a parallel ensemble structure is employed in which each classifier is trained on "Red Uni Train" training dataset independently. Then elementary combiners based on al-

Table 2: Performance of kNN on "Red Uni Training" training dataset for 6 features using 10-fold cross validation

Evaluation Metrics	kNN Classifiers												
	k=1	k=3	k=5	k=7	k=9	k=11	k=13	k=15	k=17	k=19	k=21	k=23	k=25
TPR (%)	99.8	99.8	99.7	99.7	99.7	99.6	99.6	99.6	99.6	99.6	99.6	99.6	99.6
FPR (%)	0.2	0.2	0.3	0.3	0.4	0.4	0.4	0.4	0.5	0.5	0.5	0.5	0.5
ACC (%)	99.84	99.78	99.74	99.71	99.67	99.63	99.62	99.6	99.58	99.58	99.58	99.56	99.56
PRE (%)	99.8	99.8	99.7	99.7	99.7	99.6	99.6	99.6	99.6	99.6	99.6	99.6	99.6
ROC (%)	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9	99.9
RMSE (%)	3.88	4.09	4.42	4.68	4.98	5.26	5.39	5.53	5.68	5.8	5.8	5.99	5.99

Table 3: Performance of classifiers on "Uni Train" and "Red Uni Train" training datasets for 41 and 6 features

Evaluation Metrics	# Features	Classifiers					
		NB	NN-SGD	kNN(k=3)	RIPPER	C4.5	RF
TPR (%)	41	97.0	99.5	99.9	99.9	99.9	99.9
	6	95.1	97.2	99.9	99.8	99.9	99.9
FPR (%)	41	3.8	0.5	0.1	0.2	0.1	0.1
	6	6.1	3.7	0.1	0.2	0.2	0.1
ACC (%)	41	96.99	99.47	99.87	99.85	99.91	99.93
	6	95.12	97.16	99.87	99.83	99.88	99.9
PRE (%)	41	97.0	99.5	99.9	99.9	99.9	99.9
	6	95.2	97.2	99.9	99.8	99.9	99.9
ROC (%)	41	97.6	99.5	100	99.9	100.0	100.0
	6	99.2	96.7	100	99.8	100.0	100.0
TBM (sec)	41	2.56	541.43	0.08	171.51	40.34	952.99
	6	0.45	170.56	0.08	46.28	3.24	38.11
TTM (sec)	41	6.74	29.87	15781.1	0.64	0.58	146.05
	6	1.61	1.61	5087.73	0.21	0.35	8.85
RMSE (%)	41	17.28	7.25	2.84	3.78	2.9	2.56
	6	21.97	16.86	3.09	4.05	3.32	2.85

gebraic combination rules are utilized to fuse the decisions of these base classifiers to produce final decision. In this paper, *Average, Product, Majority Voting, Minimum,* and *Maximum* methods of *classifier fusion* approach are used to fuse the decisions of the classifiers to produce final decision. The reason for using these methods is to achieve good results, very fast computation and their simplicity. Additionally, the heterogeneous ensemble is also constructed using "Uni Train" training dataset (41 features) for comparison. The 5 base classifiers—NB, NN-SGD, RIPPER, C4.5 and RF out of 6 are utilized for the construction of ensemble. The results of these ensembles using reduced 6 and all 41 features based on different performance metrics on training datasets ("Red Uni Train" and "Uni Train") are illustrated in Table 4.

5.5 Phase V: Evaluation of the Classifiers and Ensembles

Datasets "Uni Test", "Red Uni Test", "Uni Corr Test" and "Red Uni Corr Test" have been used to test the effectiveness of classifiers and ensembles used in this paper. Performance metrics were used in the experiments are TPR, FPR, ACC, PRE, ROC, TBM, TTM and RMSE. The performance of these classifiers using all 41 features and 6 selected features based on different performance metrics were evaluated on test datasets ("Uni Test" and "Red Uni Test") are shown in Table 5 and on "Uni Corr Test" and "Red Uni Corr Test" in Table 7. The performance of constructed heterogeneous ensembles using *Average, Product, Majority Voting, Minimum,* and *Maximum* using all 41 features and 6 selected features based on different performance metrics were evaluated on test dataset ("Uni Test" and "Red Uni Test") are illustrated in Table 6 and on "Uni Corr Test" and "Red Uni Corr Test" in Table 8. The classifiers and ensembles were also evaluated on training datasets ("Uni Train" and "Red Uni Train") for 41 and 6 features using 10-fold cross validation. Table 9 illustrates the results of NB, NN-SGD, RIPPER, C4.5, RF and HEIC on training dataset ("Red Uni Train") for 6 features using 10-fold cross validation.

6 Experimental Results and Analysis

To evaluate the performance of HyFSA-HEIC proposed in Section 5 in terms of accuracy and efficiency, several

Table 4: Performance of ensembles on "Uni Train" and "Red Uni Train" training datasets for 41 and 6 features

Evaluation Metrics	# Features	Ensemble of Classifier (Vote)				
		Average	Product	Majority Voting	Minimum	Maximum
TPR (%)	41	100.0	99.8	100.0	99.8	98.1
	6	99.9	99.6	99.9	99.6	97.8
FPR (%)	41	0.0	0.2	0.0	0.2	2.8
	6	0.1	0.5	0.1	0.5	3.2
ACC (%)	41	99.97	99.47	99.97	99.47	98.15
	6	99.9	97.16	99.91	97.16	97.85
PRE (%)	41	100.0	99.8	100	99.8	98.2
	6	99.9	99.6	99.9	99.6	97.9
ROC (%)	41	100.0	99.7	100.0	99.7	100.0
	6	100.0	98.3	99.9	98.3	99.9
TBM (sec)	41	626.23	590.15	541.61	616.26	599.57
	6	227.6	227.12	226.54	264.01	253.69
TTM (sec)	41	21.5	17.86	20.34	21.08	18.08
	6	8.42	9.91	9.09	10.39	10.47
RMSE (%)	41	4.29	3.91	1.85	3.91	9.17
	6	7.49	6.14	3.06	6.14	11.59

Table 5: Performance of classifiers on "Uni Test" and "Red Uni Test" test datasets for 41 and 6 features

Evaluation Metrics	# Features	Classifiers					
		NB	NN-SGD	kNN(k=3)	RIPPER	C4.5	RF
TPR (%)	41	97.1	99.5	99.8	99.9	99.9	100.0
	6	95.2	97.2	99.9	99.9	99.8	99.9
FPR (%)	41	3.7	0.5	0.2	0.1	0.1	0.1
	6	6.1	3.6	0.2	0.2	0.2	0.1
ACC (%)	41	97.06	99.48	99.81	99.92	99.87	99.96
	6	95.17	97.23	99.84	99.86	99.85	99.92
PRE (%)	41	97.1	99.5	99.8	99.9	99.9	100.0
	6	95.2	97.3	99.8	99.9	99.8	99.9
ROC (%)	41	97.6	99.5	99.9	99.9	99.9	100.0
	6	99.3	96.8	100.0	99.9	99.9	100.0
RMSE (%)	41	17.07	7.24	3.94	2.81	3.33	2.15
	6	21.87	16.64	3.67	3.62	3.63	2.54

experiments have been performed. All experiments were conducted on an Intel (R) @ 2.13 GHz Core (TM) i3 CPU M 330 computer with 2.87 GB memory and Windows 7 Home Premium operating system in Java Environments Weka 3.7.13. Datasets were used in the experiments for training are "Uni Train" and "Red Uni Train", and for testing are "Uni Test", "Red Uni Test", "Uni Corr Test" and "Red Uni Corr Test" depicted in Table 1.

Firstly, hybrid feature selection approach (Phase II, Section 5) was utilized to obtain the optimal features set for the identification of normal or intrusion instances. These features set obtained based on performance is reduced to 15% from 41 to 6 features. Then training and testing were performed for 6 diverse classifiers and 5 ensembles built from these classifiers using reduced 6 and all 41 features sets and then compared these models using 6 features with those using 41 features on different evaluation metrics.

From Table 3, as can be seen, all 6 classifiers obtained same or better performance on original 41 features than reduced 6 features set on all evaluation metrics except TBM and TTM. It is evident that classifier RF is equally best on 41 features as well as on 6 features set with TPR(99.9%), FPR (0.1%), PRE (99.9%), and ROC (100.0%) and also outperforms other five classifiers in terms of all evaluation metrics but TBM and TTM. This classifier has ACC of 99.93% and 99.90%, and RMSE of 2.56% and 2.85% on 41 and 6 features respectively. The classifier kNN has minimum TBM of 0.08 sec because no explicit training step is required. Apart from this, among all classifiers, NB has minimum TBM of 2.56 sec and 0.45 sec on 41 and 6 features respectively, C4.5 has minimum TTM of 0.58 sec on 41 features, RIPPER has minimum TTM of 0.21 sec on 6 features as depicted in Table 3. The performance of the classifier kNN is same on both 41 and 6 features set in terms of TRP (99.9%), FPR (0.1%), ACC

Table 6: Performance of ensembles on "Uni Test" and "Red Uni Test" test datasets for 41 and 6 features

Evaluation Metrics	# Features	Ensemble of Classifier (Vote)				
		Average	Product	Majority Voting	Minimum	Maximum
TPR (%)	41	99.9	99.8	99.9	99.8	98.1
	6	99.9	99.6	99.9	99.6	97.8
FPR (%)	41	0.1	0.2	0.1	0.2	2.8
	6	0.2	0.5	0.2	0.5	3.2
ACC (%)	41	99.94	99.45	99.94	99.45	98.15
	6	99.87	97.22	99.88	97.22	97.84
PRE (%)	41	99.9	99.8	99.9	99.8	98.2
	6	99.9	99.6	99.9	99.6	97.9
ROC (%)	41	100.0	99.6	99.9	99.6	100.0
	6	100.0	98.4	99.9	98.4	99.9
RMSE (%)	41	4.44	4.44	2.49	4.44	9.19
	6	7.5	6.27	3.42	6.27	11.6

Table 7: Performance of classifiers on "Uni Corr Test" and "Red Uni Corr Test" test datasets using 41 and 6 features

Evaluation Metrics	# Features	Classifiers					
		NB	NN-SGD	kNN (k=3)	RIPPER	C4.5	RF
TPR (%)	41	91.5	92.8	94.2	94.5	94.5	94.2
	6	90.4	91.8	95.4	95.2	92.6	94.6
FPR (%)	41	12.3	10.7	9.0	8.5	8.6	9.1
	6	15.0	12.8	6.8	7.5	11.2	7.1
ACC (%)	41	91.52	92.77	94.20	94.52	94.51	94.21
	6	90.41	91.78	95.36	95.15	92.61	94.61
PRE (%)	41	91.8	93.1	94.5	94.8	94.8	94.6
	6	91.3	92.4	95.5	95.4	93.0	94.6
ROC (%)	41	93.3	91.0	93.9	93.1	94.6	99.3
	6	97.9	89.5	94.6	93.8	94.0	97.1
RMSE (%)	41	29.02	26.88	23.19	23.41	23.26	19.73
	6	30.91	28.67	20.89	22.08	25.41	20.64

Table 8: Performance of ensembles on "Uni Corr Test" and "Red Uni Corr Test" test datasets using 41 and 6 features

Evaluation Metrics	# Features	Ensemble of Classifier (Vote)				
		Average	Product	Majority Voting	Minimum	Maximum
TPR (%)	41	94.3	93.7	94.3	93.7	91.0
	6	93.5	93.2	93.6	93.2	92.1
FPR (%)	41	8.9	9.5	8.9	9.5	13.7
	6	10.3	11.0	10.1	11.0	12.7
ACC (%)	41	94.31	92.66	94.31	92.66	90.99
	6	93.48	91.12	93.60	91.19	92.06
PRE (%)	41	94.7	93.9	94.7	93.9	91.6
	6	93.9	93.8	94.0	93.8	92.8
ROC (%)	41	99.2	91.7	92.7	91.7	99.0
	6	97.9	90.3	91.8	90.3	98.0
RMSE (%)	41	21.94	25.19	23.84	25.19	21.26
	6	22.74	25.98	25.31	25.98	21.78

(99.87%), PRE (99.9%), ROC (100.0%) and TBM (0.08 sec) but TTM is reduced by 67.76% from 15781.1 sec to 5087.73 sec (Table 3). The classifiers kNN, C4.5 and RF achieved same ROC of 100.0%, the classifier C4.5 achieved same TPR (99.9%), PRE (99.9%) and ROC (100.0%) on original and reduced features set. Performance of NB and NN-SGD is slightly higher for original features set, whereas performance of RIPPER on reduced 6 features

Table 9: Performance of classifiers on "Red Uni Train" training dataset for 6 features using 10-fold cross validation

Evaluation Metrics	Classifiers					
	NB	NN-SGD	RIPPER	C4.5	RF	HEIC
TPR (%)	95.1	97.1	99.8	99.8	99.9	99.8
FPR (%)	6.1	3.7	0.2	0.2	0.1	0.2
ACC (%)	95.12	97.14	99.83	99.83	99.9	99.83
PRE (%)	95.1	97.2	99.8	99.8	99.9	99.8
ROC (%)	99.3	96.7	99.8	99.9	100	99.8
TBM (sec)	0.22	105.87	42.57	3.12	32.68	184.91
RMSE (%)	21.97	16.92	4.05	3.82	2.85	4.09

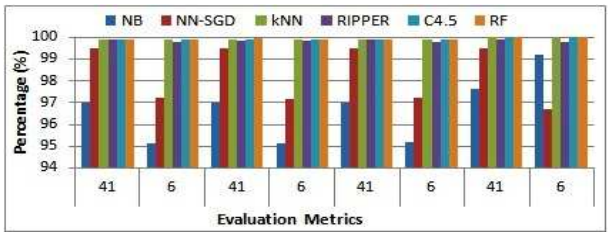


Figure 3: Performance comparison of classifiers in terms of TPR, ACC, PRE and ROC on 41 & 6 features

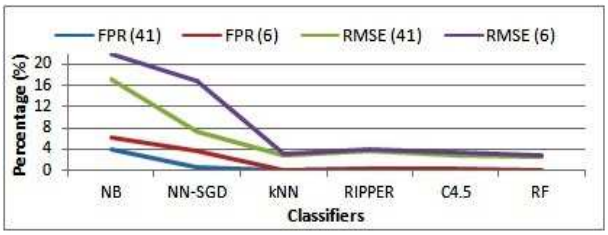


Figure 4: Performance comparison of classifiers in terms of FPR, RMSE on 41 & 6 features

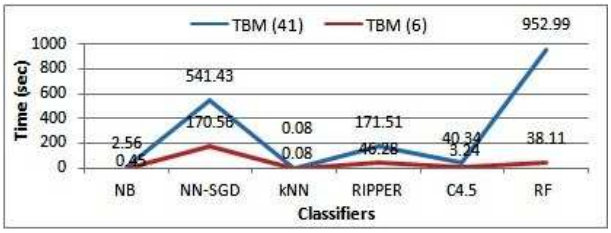


Figure 5: Comparison of TBM for 41 & 6 features in sec

set is near to that of original one. Comparative graph of the performance of classifiers for 41 and 6 features set are shown in terms of TPR, ACC, PRE, and ROC in Figure 3, FPR and RMSE in Figure 4. As can be seen from Table 3, TBM and TTM for original 41 features set are remarkably higher than that of reduced 6 features set. The TBM is reduced by approximately 68-96% except for kNN and TTM is reduced by approximately 40-94% for 6 features set. The Figure 5 and 6 show comparative graph for TBM and TTM on 41 and 6 features respectively. The performances of classifiers were evaluated using test datasets("Uni Test" and "Red Uni Test") are illustrated in Table 5. As can be seen, performances of NB and NN-SGD classifiers in testing phase (Table 5) are little bit higher with comparison to the performance of these classifiers in training phase (Table 3), whereas classifiers kNN, C4.5, RIPPER and RF performed near to equal in training and testing phase. From Figure 3, 4, 5 and 6, it can be observed that selection of optimized features set consume less computation time in training and testing phase and also maintain the same classification performance as of original features set. Therefore, feature selection approach helps to build lightweight NIDS suitable for real time and on-line processing by selecting non-redundant, informative and relevant features.

Among six classifiers, kNN yielded highest TTM (5087.73) which is remarkably very high and will increase the computation time of the ensemble and in turn degrade the performance of HyFSA-HEIC. It is also not suitable for high volume network traffic for real time and on-line processing. Therefore, it was not selected as base classifier in ensemble. As a result, 5 base classifiers—NB, NN-SGD, RIPPER, C4.5 and RF were selected to form the ensemble based on the evaluation. The decision of these 5 classifiers in ensemble is combined by using 5 algebraic combination rules of classifier fusion approach—*Average, Product, Majority Voting, Minimum,* and *Maximum* to produce final decision. Finally, five ensembles were formed. Table 4 illustrates the results of these 5 ensembles on training datasets ("Uni Train" and "Red Uni Train").

Among all 5 ensemble models, ensembles with average and majority voting combiners achieved equally best performance in terms of TPR (100.0%), FPR (0.0%), ACC (99.97%), PRE (100.0%), and ROC (100.0%) on original dataset and TPR (99.9%), FPR (0.1%), and PRE (99.9%) on reduced dataset. Majority voting combiner outperformed in ACC (99.91%), average combiner outperformed in ROC (100.0%) on reduced dataset, majority voting combiner has lowest RMSE among all ensemble models as 1.85% and 3.06% for 6 and 41 features respectively from the results of Table 4. Maximum combiner has lowest performance on all evaluation metrics except

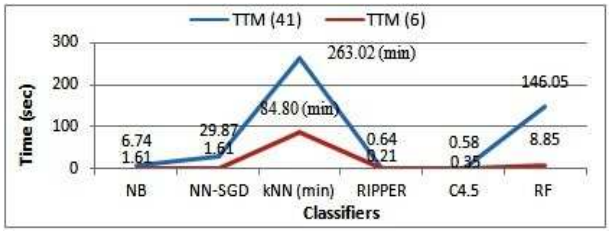


Figure 6: Comparison of TTM for 41 & 6 features (kNN in min)

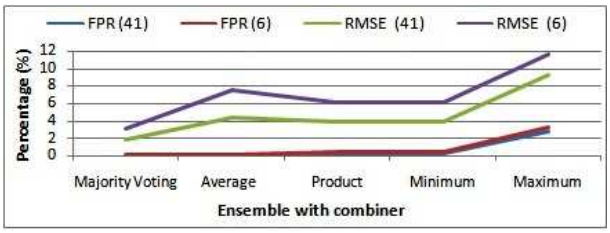


Figure 8: Performance comparison of ensembles in terms of FPR and RMSE on 41 and 6 features set.

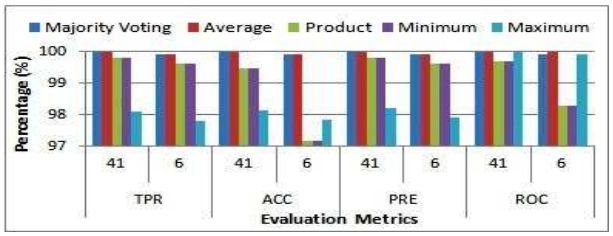


Figure 7: Performance comparison of ensembles in terms of TPR, ACC, PRE and ROC on 41 and 6 features set.

Figure 9: Comparison of Time-span to Build the Model (TBM) for ensembles on 41 and 6 features in seconds.

ROC (100.0%) on 41 features and TBM (253.69 sec) on 6 features. Whereas performances of product and minimum combiners degraded due to unclassified instances of 275 (0.38%) and 1800 (2.47%) on 41 and 6 features respectively but achieved lowest error rate except majority voting combiner (Table 4). Therefore the best performing combining rule for ensemble model is majority voting based on overall performance. TBM and TTM for all combiners of ensemble models are almost the same on 41 as well as on 6 features set. From Table 4, it is observed that TBM and TTM of ensembles are drastically reduced by approximately 58-64% and 42-56% respectively for 6 features set. Figure 7 shows performance comparisons of ensembles in terms of TPR, ACC, PRE, and ROC, Figure 8 for FPR and RMSE, Figure 9 for TBM and Figure 10 for TTM for 41 and 6 features. The performances of ensembles in testing phase on test datasets ("Uni Test" and "Red Uni Test") as illustrated in Table 6 are almost the same to the performances in the training phase (Table 4). The performances of product and minimum combiners degraded due to unclassified instances of 255 (0.35%) and 1745 (2.40%) on 41 and 6 features respectively but achieved lowest error rate except majority voting combiner (Table 6) in the testing phase also.

The experiments were also conducted to measure the performance of classifiers and ensembles on "Uni Train" and "Red Uni Train" training datasets for 41 and 6 features using 10-fold cross validation. It was found that the performance of classifiers and ensembles using 10-fold cross validation were almost same to the performance in the training and testing phase for 41 and 6 features. As can be seen from the results of Tables 3, 4, 5, 6 and 9 of NB, NN-SGD, RIPPER, C4.5, RF and HEIC for 6 features.

The performance of classifiers and ensembles also tested on "Red Uni Corr Test" test dataset for reduced 6 features are near to same the performance on "Uni Corr Test" for original 41 features on all evaluation metrics as illustrated in Tables 7 and 8. Hence the proposed system also achieves near to equal performance on reduced 6 features on this test dataset. The performance of classifiers and ensembles in training phase (Tables 3 & 5) on "Uni Train" and "Red Uni Train" training datasets and testing phase (Tables 4 & 6) on "Uni Test" and "Red Uni Test" for 41 and 6 features are at the higher side than that of testing phase (Tables 7 & 8) on "Uni Corr Test" and "Red Uni Corr Test" test datasets for 41 and 6 features. The reason for this is the test dataset ("Corrected Test") is not from the same probability distribution as the training dataset ("10 % KDD") as it includes additional 17 novel attack types that are not present in training dataset. It makes the system more realistic to perform on real time data.

On the comparison of results of 5 ensemble models (Tables 4 & 6) and 6 classifiers (Tables 3 & 5) based on different evaluation metrics with selected 6 features has shown that ensemble with majority voting combiner outperformed other ensemble models and individual classifiers and thus more reliable and capable for NIDS and hence chosen as ensemble model for HyFSA-HEIC. The performance comparison of individual classifiers NB, NN(SGD), RIPPER, C4.5, RF and HyFSA-HEIC in terms of TPR, ACC, PRE, ROC is shown in Figure 11 and in terms of FPR and RMSE in Figure 12. The results strongly indicate that by employing feature selection approach as preprocessing step and heterogeneous ensemble of intelligent

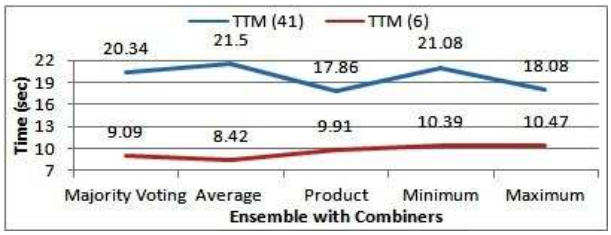


Figure 10: Comparison of Time-span to Test the Model (TBM) for ensembles on 41 and 6 features in seconds.

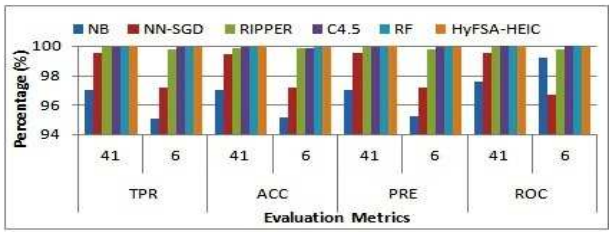


Figure 11: Performance comparison of classifiers NB, NN-SGD, RIPPER, C4.5, RF and HyFSA-HEIC.

classifiers in model building enhance the performance of HyFSA-HEIC. It has been enhanced in terms of TRP (99.9%), ACC (99.91%), PRE (99.9%), ROC (99.9%), with extremely low FPR (0.1%) & RMSE (3.06%) and has faster building and testing time than the ensemble with full features set.

7 Conclusion and Future Work

The aim of this work is to propose Hybrid Feature Selection Approach – Heterogeneous Ensemble of Intelligent Classifiers (HyFSA-HEIC) for network intrusion detection and to demonstrate that this system can enhance the accuracy and efficiency of the system as well as reduce the false positive rate, error rate, training and testing time. It hierarchically integrates hybrid feature selection approach (HyFSA) with heterogeneous ensemble of intelligent classifiers (HEIC). The main challenging issues arise in IDS are to handle large-scale high dimensional dataset and maximizing overall accuracy and less false alarm. The HyFSA-HEIC addresses these issues by incorporating hybrid feature selection approach (HyFSA) and heterogeneous ensemble of intelligent classifiers (HEIC). The heterogeneous ensemble built in this work employed five diverse accurate intelligent classifiers—NB, NN (SGD), RIPPER, C4.5 and RF and their decisions were combined by utilizing majority voting of elementary combiner based on algebraic combination rule. This ensemble was built on using only 6 selected features i.e. only 15% of original 41 features. Several experiments were performed to compare the HyFSA-HEIC with other ensembles and individual classifiers with and without applying feature selection approach. "KDD Cup 1999" and "Corrected Test"

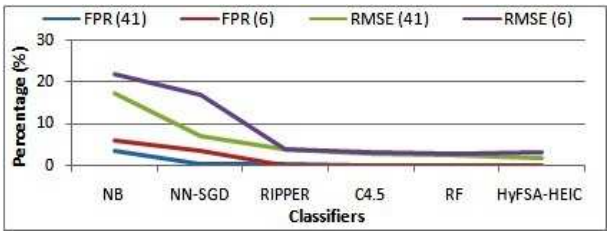


Figure 12: Performance comparison of NB, NN-SGD, RIPPER, C4.5, RF and HyFSA-HEIC.

datasets have been utilized to train, and test the methods and HyFSA-HEIC used in this work. The results show that HyFSA-HEIC outperforms other methods with true positive rate (99.9%), accuracy (99.91%), precision (99.9%), receiver operating characteristics (99.9%), and low false positive rate (0.1%) and root mean square error rate (3.06%) with minimum number of selected 6 features. It also reduces the training time by 50.79% and testing time by 55.30% on reduced features set. The classifiers used in proposed HEIC are applicable to both numerical and categorical features as well as on large dataset, which is practically advantageous for real time intrusion detection. In conclusion, integrating feature selection approach to the heterogeneous ensemble of intelligent classifier improve the detection rate, accuracy, precision, and receiver operating characteristics and reduce the false alarms and error rates with minimum computation time.

Due to continuous increase of intrusion or attack and ever-growing network traffic in computer networks, there has been an endless requirement for improvement in the performance of NIDS especially in terms of low false rate and minimum computation time. Therefore, we anticipate enhancements in the performance of the proposed heterogeneous ensemble method in terms of high accuracy, low false rate, low error rate and minimum computation time. This can be achieved by inducing higher diversity among the classifiers in ensemble as well as by investigating other classifiers for ensemble or ensemble methods. The proposed system is capable of classifying between attack and normal traffic connection, but lack in dealing with further classification of specific attack type. The extension of this work is to further classification of attack into four classes—DoS, Probe, U2R and R2L i.e. multi-class classification and determination of optimal features set for each attack type for network intrusion detection.

References

[1] B. Agarwal and N. Mittal, "Hybrid approach for detection of anomaly network traffic using data mining techniques," in *Proceedings of 2nd International Conference on Communication, Computing and Security, Procedia Technology*, vol. 6, pp. 996–1003, 2012.

[2] Amrita and P. Ahmed, "A study of feature selection methods in intrusion detection system: A survey," *International Journal of Computer Science Engineering and Information Technology Research*, vol. 2, no. 3, pp. 1–25, 2012.

[3] Amrita and P. Ahmed, "A hybrid-based feature selection approach for IDS," in *Proceedings of 5th International Conference on Networks and Communications (NETCOM'13)*, vol. 284, pp. 195–211, Chennai, India, Dec. 2013.

[4] A. L. Blum and P. Langley, "Selection of relevant features and examples in machine learning," *Artificial Intelligence*, vol. 97, no. 1-2, pp. 245–271, 1997.

[5] L. Bottou, "Large-scale machine learning with stochastic gradient descent," in *Proceedings of the 19th International Conference on Computational Statistics (COMPSTAT'10)*, pp. 177–187, Paris France, Aug. 2010.

[6] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[7] R. Chitrakar and H. Chuanhe, "Anomaly based intrusion detection using hybrid learning approach of combining k-medoids clustering and Naïve Bayes classification," in *Proceedings of 8th IEEE International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM'12)*, pp. 1–5, Sep. 2012.

[8] W. W. Cohen, "Fast effective rule induction," in *Proceedings of the 12th International Conference on Machine Learning*, pp. 115–123, Paris France, 1995.

[9] S. Das, "Filters, wrappers and a boosting-based hybrid for feature selection," in *Proceedings of the Third International Conference on Machine Learning*, pp. 74–81, Dalian, China, June 2001.

[10] T. G. Dietterich, "Ensemble methods in machine learning," *Multiple Classifier Systems, Lecture Notes in Computer Science*, vol. 1857, pp. 1–15, 2000.

[11] I. El-Henawy, H. M. El Bakry, H. M. El Hadad, "A new muzzle classification model using decision tree classifier," *International Journal of Electronics and Information Engineering*, vol. 6, no. 1, pp. 12–24, 2017.

[12] U. M. Fayyad and K. B. Irani, "Multi-interval discretization of continuous-valued attributes for classification learning," in *Proceedings of the 13th International Joint Conference on Artificial Intelligence (IJCAI'93)*, pp. 1022–1029, Paris France, 1993.

[13] T. F. Ghanem, W. S. Elkilani, and H. M. Abdul-Kader, "A hybrid approach for efficient anomaly detection using meta heuristic methods," *Journal of Advanced Research*, vol. 6, no. 4, pp. 609–619, 2014.

[14] V. Golmah, "An efficient hybrid intrusion detection system based on C5.0 and SVM," *International Journal of Database Theory and Applications*, vol. 7, no. 2, pp. 59–70, 2014.

[15] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The WEKA data mining software: An update," *SIGKDD Explorations*, vol. 11, no. 1, 2009. (`http://www.cs.waikato.ac.nz/ml/weka/`)

[16] G. Kim, S. Lee, and S. Kim, "A novel hybrid intrusion detection method integrating anomaly detection with misuse detection," *Expert Systems with Applications*, vol. 41, no. 4, pp. 1690–1700, 2014.

[17] R. Kohavi and G. H. John, "Wrappers for feature subset selection," *Artificial Intelligence*, vol. 97, no. 1-2, pp. 273–324, 1997.

[18] L. I. Kuncheva, *Combining Pattern Classifiers: Methods and Algorithms*. Wiley-Interscience, 2004.

[19] L. I. Kuncheva and C. J. Whitaker, "Measures of diversity in classifier ensembles and their relationship with ensemble accuracy," *Machine Learning*, vol. 51, no. 2, pp. 181–207, 2003.

[20] Y. Li, J. Xia, S. Zhang, J. Yan, X. Ai, and K. Dai, "An efficient intrusion detection system based on support vector machines and gradually feature removal method," *Expert Systems with Applications*, vol. 39, no. 1, pp. 424–430, 2012.

[21] S. W. Lin, K. C. Ying, C. Y. Lee, and Z. J. Lee, "An intelligent algorithm with feature selection and decision rules applied to anomaly intrusion detection," *Applied Soft Computing*, vol. 12, no. 10, pp. 3285–3290, 2012.

[22] H. Liu and H. Motoda, *Feature Selection for Knowledge Discovery and Data Mining*, Boston: Kluwer Academic, 1998.

[23] M. Panda, A. Abraham, and M. R. Patra, "A hybrid intelligent approach for network intrusion detection," in *Proceedings of the International Conference on Communication Technology and System*, vol. 30, pp. 1–9, 2012.

[24] J. R. Quinlan, *C4.5: Programs for Machine Learning*, San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1993.

[25] L. Rokach, "Ensemble-based classifiers," *Artificial Intelligence Review*, vol. 33, no. 1-2, pp. 1–39, 2010.

[26] S. S. S. Sindhu, S. Geetha, and A. Kannan, "Decision tree based light weight intrusion detection using a wrapper approach," *Expert Systems with Applications*, vol. 39, no. 1, pp. 129–141, 2012.

[27] A. Tayal, N. Mishra and S. Sharma, "Active monitoring & postmortem forensic analysis of network threats: A survey," *International Journal of Electronics and Information Engineering*, vol. 6, no. 1, pp. 49–59, 2017.

[28] S. Thirumuruganathan, *A Detailed Introduction to K-Nearest Neighbor (KNN) Algorithm*, World Press, 2010.

[29] UCI KDD, *KDD Cup 1999 Intrusion Detection Dataset*, Oct. 28, 1999. (`http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html`)

[30] UNB, *NSL-KDD Network-based Intrusion Detection Dataset*, Apr. 29, 2017. (`http://nsl.cs.unb.ca/KDD/NSL-KDD.html`)

[31] H. Zhang, "The optimality of Naïve Bayes," in *Proceedings of The Seventeenth International Florida Artificial Intelligence Research Society Conference*, pp. 17–19, Miami Beach, 2004.

Biography

Amrita is an Assistant Professor in Department of Computer Science and Engineering at Sharda University, Greater Noida, INDIA. She received her M.Tech. in Computer Science from Banasthali Vidyapith, Rajasthan. She is currently pursuing her Ph.D. in Computer Science and Engineering from Sharda University, Greater Noida (U.P.). She has more than 15 years of experience in Academics, Software Development Industry and Government Organization.

Kiran Kumar Ravulakollu is working with Sharda University of INDIA as Assistant Professor with interests in sensory networks, robotics, biologically inspired multimodal behaviour modelling, ambient intelligence, and sensory network design along with visual and auditory information processing. He has more than 5 years of research experience in Hybrid Intelligent Systems area at Centre for Hybrid Intelligent Systems, University of Sunderland, UK. He has received his Ph.D. degree from University of Sunderland, UK.

Distributed Intrusion Detection System Based on Mixed Cooperative and Non-Cooperative Game Theoretical Model

Amin Nezarat

(Corresponding author: Amin Nezarat)

Department of Computer Science & Engineering and IT, Payame Noor University
Post Box: 19395-4697, Lashkarak Highway, Nakhl street, Tehran, Iran
(Email: aminnezarat@pnu.ac.ir)

(Received Feb. 14, 2016; revised and accepted June 10 & Sept. 25, 2016)

Abstract

Intrusion Detection Systems (IDS) are systems to protect the network resources against the attacks. Considering the extent of the attacks in the internet environment and the change in the form and type of the attacks from the centralized to the distributed strategy, such systems also tend to move towards the distributed architecture. In this paper, a mobile agent based method working as suspicious movements detection sensors has been proposed. The attack detecting White Globule Agents (WGA) scatter in the network; moving from a node to another, they build a Security Overlay Network at a time and using a kind of collaborative game and communication, and after reaching to the desired Shapley value, they can detect and report the origin and level of the attack. The proposed method in this study includes a scenario in which the WGA in a non-cooperative game against the attacker element tries to develop a negotiation to calculate the Nash equilibrium point and attain maximum utility, so that meanwhile separating the attacks from the real requests, the security level of the attack is obtained with the aid from other WGAs.

Keywords: DIDS; Game Theory; Mobile Agents; Multi-Agent Systems; Nash Equilibrium; Network Security; Shapley Value

1 Introduction

The attacks and vulnerability of computerized networks are evolving and spreading day by day both from the complexity and technology perspectives, to the extent that some attacks might pull the E-commerce companies out of the business cycle [2, 21]. The intrusion detection systems must be able to propose suitable solutions in mitigating the harms and attacks to the computing networks [6, 15]. To achieve this aim the public and private organizations and businesses have taken steps towards establishment of security operation centers for analyzing and monitoring the intrusion reports and incidents.

For example, one of the security tools developed by Symantec called the 'Threat con warning sys', measuring the Security Risk value and reports the threat scale to the network manager [9, 20]. The network manager then may predict the necessary measures to control the attack based on the Security Risk value. Besides, the measurement is not configured based on the transmitted data between the attacker and defender.

The recently IDSs can be divided into two groups:

1) Reactive (signature based authentication or entry/exit permission).

2) Proactive (Secure Overlay Network, proxy network, etc.).

Secure Overlay Network systems have an architecture designed to preventing from the attacks and preparation against the distributed DOS attacks. But the method used in this architecture for establish communication between the nodes is based on permanent connection which imposes heavy burden on the network, a drawback impeding the further progress of this architecture. Although the proactive method is a more efficient and advantageous method than the reactive method, yet abundant problems remain unsolved as to implementation of total efficiency of the method.

The centralized IDS systems are highly susceptible to Single Point of Failure and can be discovered by the invaders to attack. To remove this problem a higher number of IDS systems can be used to diminish the number of undiscovered attacks; albeit requiring higher expenses [18, 22].

The power of an intrusion detection system lies in establishing equilibrium between the number of defenders and the number of detected and/or undetected errors.

For this purpose, employing a distributed intrusion detection system seems necessary for modeling the attacks and measuring Security Risk Value (attack threat) to facilitate decision making ability in response to the attacks. Mobile agents can be considered as defenders in network nodes that autonomously move among the nodes and a collection of them can make an overlay network from a multi-agent system. A mobile agent can enter into an interaction with the attacker to create a better understanding of a security attack and assessing the risk level.

In order to escalate the decision making accuracy and detecting the attack a game theoretic method can be used in interaction between the detecting agent and the attacker. Afterwards each one of the detecting agents may -in an another game- enter into negotiation with other agents inside the Overlay Network to evaluate the accuracy rate of the obtained results. This game is of the collaborative type to reach an agreement concerning the security risk value of the attack.

In the first game between the mobile WGA[1] as the defender and an attacker, a Reciprocal interaction is established which is a non-cooperation two players game. Both parties of the game seek maximization of their own utility and payoff. The attacker intends to penetrate into the network and destruct it through gathering information and the defender seeks to counter and detect greater number of attacks to the network. Both of the players have a utility function and reach equilibrium at a point where both players play in a way pretending not willing to change the game. At this point, the defending mobile agent attains at a value showing the security risk value.

To make sure whether or not an attack has taken place and to avoid issuing false alerts, the mobile agent in another collaborative game negotiates with other defender agents in a multi-agent environment. In this game the agents can form coalitions out of multiplayer groups to avoid overburden of the network. In each coalition formed for an attack issue, the agents after negotiation reach a shared compromised value called Shapley Value which determines the final network security value and the measures available for decision making against the attack issue. Considering the obtained value, the severity and level of the counterattack could be specified [9, 10].

2 Multi Agent Overlay Network

The distribution type of the mobile agents and their array of position in each time to cover the best part of the network is one of the crucial issues in establishing Overlay Network and a significant criterion for each agent's decision as regards to its mobility in the network.

For example: In Figure 1 it is assumed that an intruder intents to send a packet from the host α (intruder) to the host t (target) and a multi agent overlay network consisting of $MA^2 = \{MA_1, MA_2\}$ has been constituted. Each of these nodes is the host of a mobile agent. These agents at any moment must take position in the network nodes in a way that the maximum security coverage is established, and considering their mobility, the protective agent's movement domain must be configured so that the accumulation of the agents or vacuum in the network could be avoided. For this reason and in order to reduce/increase the error detection rate and controlling the network bandwidth, such factors as the number of protective agents, the area covered by each of them and their security sensitivity level must be determined, corrected and announced throughout the time. In this study, aiming at increasing the network protection and dealing with the Overlay Network Architecture intruders and administering the agents' mobility as well as the security game to reach the agent sensitivity value (Shapley value) has been investigated and proposed.

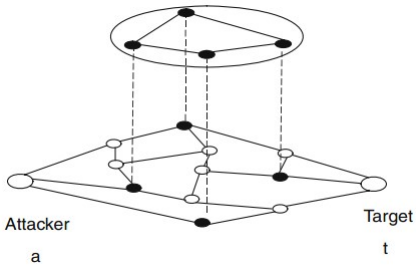


Figure 1: Overlay network

3 Architecture of the Proposed Multi-agent System

Given the extent of computer networks and inner network growth of data exchange and communication volume, establishing and maintaining perennial communication can cause overburden of the network, disturbing the main function of the institutions' application software. Security and monitoring systems must be able to fulfill their tasks with minimal turbulence in the network and functional overburden.

Mobile agents are one of the agent types. A mobile agent is an autonomous, active and movable software capable of relocation inside the network without the need for continuous control by the main server. It can perform its duties including gathering information, combating the attacks, buy and sell or negotiating with other agents, and ultimately reporting the results to the dispatching server without the requirement to return to the server and occupying the bandwidth or even sending all the information [12, 13, 14].

The agents can through constituting multi-agent communities and establishing synergistic communications with each other attain fast results with lowest human involvement possible. The agents also can learn, change

[1] White Globule Agent
[2] Mobile Agent

shape and improve the initial strategy. In this study the negotiating feature of the agents has been used to introduce the proposed architecture.

In order to prevent agents from getting enlarged and their involvement in examining the computerized packages, one of the renowned and open source IDSs, namely the SNORT has been used [19]. SNORT is a perfect and widely used IDS which through investigating the network packages and comparing them with the attack scenarios existing in its knowledge bank can detect the attacks and issue the appropriate alert message. But as was mentioned before, this system is a central IDS, and in order to detect a set of attacks to different nodes of the network, it must be installed in different points of the network and be able to communicate with each other. Currently this system lacks this capability and to achieve this objective, all the logs of the attacks must be relocated to the central IDS and be merged so that totality of the attack could be correctly recognized, something that is not accessible at the moment.

In the proposed model (Figure 2) which is a combination of mobile agents, a central server has been used to produce the agents. A number of mobile agent as White Globules have been deployed to the network nodes which can move in specified periods. The server composition includes the following parts: White Globule Agent Factory; Lymphocyte Agent Factory; SNORT; and IDS Knowledge Base.

- **White Globule Agent Factory**
  This WGAF agent is tasked with producing the new defender agents. At the beginning of the system operation a number of locations peculiar to the WGAs that are to be produced are identified by the network administrator and this agent sends each WGA to the specified node after its creation. The type of travel of the WGA to the network nodes is of the strong mobility type, accompanying any time the status of all previous events and their consequences. Each WGA agent periodically communicates with this agent and announces a report of its observations. At any moment the agent knows where the WGA is. The differences of the produced WGAs in this agent lie in that a WGA agent includes in itself a summary of the attack scenarios in the network and with the aid from the intrusion detection tool SNORT, it can identify the suspicious behaviors in the network. After identifying a suspicious behavior, it negotiates with other WGA agents and complements its knowledge or makes corrections to it. This agent has been comprised of several other agents, each tasked with a particular duty:
    - AMS[3] agent: It is the chief manager of the agents and allocates a dedicated number to each agent. It is responsible for the creation, deletion, relocation, etc. of the agents.
    - DF[4] agent: This agent has the full information of all the created and implemented agents. It knows the name, location, specifications and history of the agent. When a WGA agent wants to communicate with other agents, it sends its request to this agent.
- **Lymphocyte Agent Factory (LAF)**
  One of the duties of Network Intrusion Detection systems is the accumulation of information and identification of the attacks for reporting the attack prevention systems or IPSs[5]. In this model a kind of agent has been considered having been tasked with the attacks and removing of the attacker. After identification and confirmation of an attack by a coalition of the WGAs, a report is sent to the WGAF including the assessed security risk value.

  The WGA uses this security risk value and sends a request to this agent and the mentioned agent detects the number of agents and the locations they must be sent to.
- **SNORT**
  SNORT is an open source IDS which through considering the exchanged packets in the network and sniffing of the network enables the user to implement of some rules. In this tool a number of rules for identifying the attacks are used and when the scenario embedded in the rules is achieved, it shows the suitable reaction. In this paper the SNORT package has been used for comparing the scenario with the network packets.
- **IDS Knowledge Base**
  In this Knowledge base the scenarios relating to the security attacks in the network is depicted so that each agent to know what to do in the event a security attack occurs. Since a WGA agent should interact with the attacker when it senses a threat, using the information gathered in this Knowledge, it knows what measure it must use in the next step to compel the attacker for implementing the next stage of the attack; so by simulation of a Vulnerability it could detect the threat.

4 Methodology of the Proposed Model

As was mentioned before, this security model is comprised of different components. This section deals with the model function and method of implementation. The WGAF agent creates a number of WGA agents and dispatches them to the network nodes. In the advent of an attack from an exterior agent, the WGA positioned in that node starts interacting with the attacker. This is a

[3]Agent Management System

[4]Directory Facilitator

[5]Intrusion Prevention System

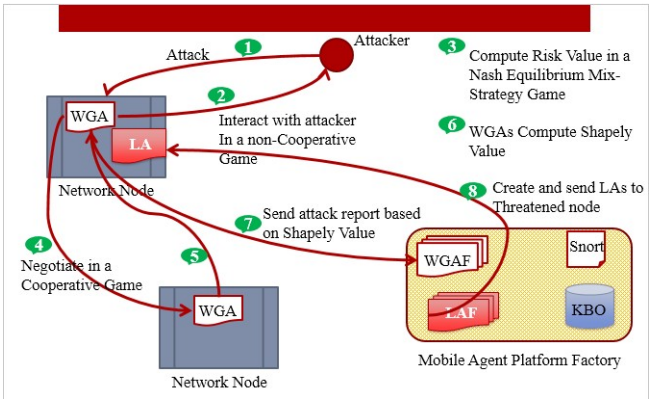


Figure 2: The proposed distributed intrusion detection system

non-collaborative type of game. The WGA agent with the help of the information obtained from the interaction with the intruder chooses a mixed strategy to reach a Nash equilibrium point. This Nash equilibrium point is considered as the risk value of that attack.

To increase in the accuracy of the risk value and determining whether or not an attack similar to this one has been identified in other nodes of the network, the WGA takes measure as to formation of a coalition with a number of other WGAs and negotiates with them in collaborative game until eventually they reach an agreed value called the Shapley value from computing the risk values obtained from other WGAs. After calculating the Shapley value, the WGA agent sends this value to WGAF. This figure is considered as a security risk of the network agreed upon by a coalition of WGAs, and the WGAF reports it to LAF and also it can decide to increase or decrease the number of WGAs or relocation of them.

The LAF agent upon receiving the new security risk of the network creates the Lymphocyte Agents (LA) and sends them to the threatened nodes. An LA agent will be sent to each threatened group and informs it of the number of its own reproductions. To prevent from occupying the network bandwidth, each LA after arriving at its destination reproduces itself to the required number so that it could suitably counter the attack considering the reported security value.

Given the new security value obtained, the number and location of the WGAs will be reviewed by the WGAF and the new number and locations is announced; in the end an overlay network consisting of the mobile agents and a new multi-agent environment is built. The criteria for this decision making are based on a number of parameters including: The threatened nodes, the threat value of each node, the traveled path of an intruder, changes of the security risk value compared with the past, etc.

5 Security Risk Game

In a general sense, by the security risk it is meant the number of the penetrations for which the possibility of occurrence exists (breach of security ratio). The security intrusion occurs when an intruder manages to enter into the system through security gaps [16, 17].

$$\textit{Security risk} = \textit{Ratio of security breach} * \textit{Average cost of each intrusion}$$

The main aim of this paper is determining the security risk value so that it could be used to adjust the Distributed IDS. In the proposed security game the following roles are defined: Attacker and WGAgent.

- **Attacker:** A network continually provides different users with the requested services, some of which could be the attacker requests. The purpose of an attacker is intruding the network to fulfill the predefined objectives (shaw, 1999). If the attack of an attacker is not detected, an expense of equal to b_2 is due to be paid by the attacker. If the attack is detected the due cost shall be equal to b_1. Also if $b_1 - b_2 \leqq 0$ then it means that in case the attacker is detected it will gain a positive payoff, and if we have $b_1 - b_2 \geqq 0$, then the attacker will gain a negative payoff. In case an attacker invades a WGA numbered i which resides on one of the network's nodes, two parameters of i_1, i_2 are assumed, the first one representing the attacker's spread and the second relates to the bandwidth the attacker occupies to invade the i-th agent. This parameter has been introduced by [1]. It can be concluded that an expanded attack will occupy λf_i of the bandwidth.
- **WGAgents:** is a set of WGA detecting agents that are represented as $N = \{WGA_1 - WGA_2 - \cdots WGA_n\}$. These agents are positioned on the network nodes and constitute a Multi Agent overlay network for detecting the intrusion and formation of a distributed IDS. Such agents are tasked with screening the ordinary users from the attackers through examining the transferring packets inside the network. Suppose the mi indicates the mobility ratio of the WGA_i agent in travel from one host to the other. The WGA_i normally may commit two types of errors: Classifying the attacker as the normal user and classifying the normal user as the attacker. In this model we seek to establish an equilibrium between such false alarms and true alarms.

For the WGA_i agent the intrusion detection ratio (detection probability) and non-detection ration are indicted by p_d and $(p_d - 1)$ respectively, considering the p_f as the false detection probability. The $-C1$ cost for detecting an attack by WGA_2 and $C2$ and $C3$ respectively are assumed as the costs the IDS must pay for false detection and non-detection cases.

On the other hand the attacker in case of intruding the system obtains the cost $-b_1$ and in case of failure, it

Table 1: Intrusion detection probabilities by the agent

IDS cost			
$-C1$	True detection	p_d	WGA_i
$C3$	Non-detection probability	$1-p_d$	
$C2$	False detection probability	p_f	

must pay the cost b_2. Issuing false alarm costs zero for the attacker.

Table 2: The cost the attacker gains

Result	Cost	
Intrusion	$-b_1$	Attacker
Not-intrusion	b_2	

Finally, each of the two players (WGA_i, attacker) shall be allocated a payoff value in Formula (1) which is calculated for each one of them as follows. For the attacker the result will be as follows:

$$payoff_{att} = r_1[p_d b_2 - b_1(1-p_d)] \tag{1}$$

where r_1 is the probability by which an attacker may show destructive behavior, intending to intrude into the system. The expected utility value WGA in IDS may also be calculated as Formula (2):

$$payoff_{WGA} = r_1 c_3 + p_d c_2 - r_1 p_d (c_1 + c_2 + c_3) \tag{2}$$

Now we model the non-cooperative game using the mixed strategy between the (WGA, attacker). The set of strategies both players may choose is as follows:

$$S_{attacker} = \{u_1, u_2, u_3\}$$
$$S_{WGA} = \{d_1, d_2\},$$

u_1 represents a full attack by the attacker with r_1 probability. u_2 is considered as the reproduction strategy of an attack in such a way that with r_2 probability it begins expansive reproducing itself.

u_3 shows that with $1-r_1-r_2$ probability an attack will not occur. As for the agent also the strategy d_1 is indicative of detecting an intrusion with q probability and issuing the necessary alert. The strategy d_2 represents a state in which the agent is showing no reaction toward the attack. The d_2 probability is shown by $1-q$. Given the above assumed strategies the expected utility of payoff for attacker and the agent is expanded as follows: The reason for expanding of the utility is the type of game for which the mixed strategy has been selected. In order to include the probability of all the assumed strategies in the final utility, the following table containing the q-mix (agent's expected utility) and r-mix (attackers expected utility) is considered [4].

This table shows the utility matrix of the two players in different strategy profile. The assumption is that each player has such a matrix and is able to calculate the other player's strategy. In other words the defender has a Belief of its strategy and the counter response from the attacker. In this stage, each player must play in a mode as if not willing to change its strategy and the competitor player also must reach to this same point (NASH equilibrium). This point that is represented by (r^*, q^*) is obtained from the matrix $3*2$ shown in Table 3 and Formula (3).

$$\begin{aligned} q^*[&-c_1(1+p_d+m_i)r_1^* - c_1(1+p_d+l_i+m_i)(r_2^*) \\ &+c_2(1+p_f+m_i)(1-r_2^*-r_1^*)] + (1-q^*) \\ &[c_3(1+(1-p_d))r_1^* + c_3(1+l_i+(1-p_d))r_2^* \\ \leq q[&-c_1(1+p_d+m_i)r_1^* - c_1(1+p_d+l_i+m_i)(r_2^*) \\ &+c_2(1+p_f+m_i)(1-r_2^*-r_1^*)] + (1-q^*) \\ &[c_3(1+(1-p_d))r_1^* + c_3(1+l_i+(1-p_d))r_2^*]. \end{aligned} \tag{3}$$

The above expression shows that the best strategy that the counter-player chooses must be lower than the selected strategy of the (main) player, q and r are both between 0 and 1.

After solving the inequality, if we assume that $r_2 = 1-r_1$, then we will have (two general modes are considered for the attacker)

if $-c_1(1+p_d+m_i)r - c_1(1+p_d+l_i+m_i)(1-r)$
$= c_3(1+(1-p_d))r + c_3(1+l_i+(1-p_d))(1-r)$
and $b_1 f_i q - b_2 f_i(1-q) = b_1(1+\lambda f_i + l_i)q$
$- b_2(1+\lambda f_i + l_i)(1-q)$ **then**

$$r^* = \{c_1(1+p_d+l_i+m_i) + c_3(1+l_i+(1-p_d))\}/$$
$$\{-c_1(1+p_d+m_i) + c_1(1+p_d+l_i+m_i)$$
$$- c_3(1+l_i+(1-p_d)) + c_3(1+(1-p_d))\}$$
$$q^* = \frac{-b_2(1+\lambda f_i+l_i) + b_2 f_i}{b_1(1+\lambda f_i+l_i) + b_2(1+\lambda f_i+l_i) + (b_1+b_2)f_i}$$

if $-c_1(1+p_d+m_i)r + c_2(1+p_f+m_i)(1-r)$
$= c_3(1+l_i+(1-p_d))(r)$
and $b_1 f_i q - b_2 f_i(1-q^*) = 0$
then $r^* = \{c_2(1+p_f+m_i)\}/\{c_1(1+p_d+m_i)$
$+ c_2(1+p_f+m_i) + c_3(1+l_i+(1-p_d))\}$

$$q^* = \frac{b_2 f_i}{(b_1+b_2)f_i}$$

if $c_2(1+p_f+m_i)(1-r) - c_1(1+p_d+l_i+m_i)(r)$
$= c_3(1+l_i+(1-p_d))(r)$
and $b_1(1+\lambda f_i+l_i)q - b_2(1+\lambda f_i+l_i)(1-q) = 0$
then $r^* = \{c_2(1+p_f+m_i)\}/\{c_2(1+p_f+m_i)$
$+ c_1(1+p_d+l_i+m_i) + c_3(1+l_i+(1-p_d))$

$$q^* = \frac{b_2(1+\lambda f_i+l_i)}{b_1(1+\lambda f_i+l_i) + b_2(1+\lambda f_i+l_i)}$$

Table 3: Payoff matrix

	WGA_i		
attack	d_1	d_2	$q-mix$
u_1	$b_i f_i,\ -c_1(1+p_d+m_i)$	$b_2 f_i,\ c_3(1+(1-p_d))$	$b_1 f_i q - b_2 f_i(1-q)$
u_2	$b_1(1+\lambda f_i+l_i)$, $-c_1(1+p_d+l_i+m_i)$	$b_2(1+\lambda f_i+l_i)$, $-c_3(1+l_i+(1-p_d))$	$b_1(1+\lambda f_i+l_i)$ $-b_2(1+\lambda f_i+l_i)(1-q)$
u_3	$0, c_2(1+p_f+m_i)$	$0, 0$	0
$r-mix$	$-c_1(1+p_d+m_i)r_1 - c_1(1+p_d+l_i+m_i)(r_2)$ $+c_2(1+p_f+m_i)(1+r_2-r_1)$	$c_3(1+(1-p_d))r_1$ $+c_3(1+l_i+(1-p_d))r_2$	

After calculating the optimum strategies and NASH equilibrium of the game, the probability vector $r^* = \{r^*(u_1), r^*(u_2), r^*(u_3)\}$ for the attacker when it selects the $\{u_1, u_2, u_3\}$ strategies and the probability vector $q^* = \{q^*(d_1), q^*(d_2), q^*(d_3)\}$ for the strategies $\{d_1, d_2\}$ selected by the defender agent are obtained.

$$v_i = \frac{q_i^*(d_1)}{q_i^*(d_2)} + \frac{r_i^*(u_1)+r_i^*(u_2)}{r_i^*(u_3)} \qquad i \in N. \tag{4}$$

6 Cooperative Game Between Agents

In this section, the collaborative game of WGAagents for formation of a coalition and determining the Shapley value using Nash equilibrium (or SRV) calculated in the previous section will be described. The Shapley value is a powerful indicator for cost allocation problem. The collaborative game in which all the players (WGAs) in a series of negotiations seek escalating the benefits of the group and the established coalition is a suitable solution for determining the security value of the overlay network [3].

First the function $R : V \to R^+$ as a one-to-one function of real numbers is considered in a way that each element of v, r is defined as $V = \{v_1, v_2, \cdots, v_j\}$. The proposed IDS security value is defined for using the L. It is defined so that

$$L = \{l_1, \cdots, l_l\}\ when\ 0 < k_1 < k_2 < \cdots < k_L$$

are considered as threshold values. The following relationship illustrates the different security level states using the agents' output vector V:

$$SL = \begin{cases} l_1 & \text{if } \sum_{i=1}^{N} R(v_i) \geqslant k_1 \\ l_j & \text{if } \sum_{i=1}^{N} R(v_i) \geqslant k_j \\ l_{j+1} & \text{if } \sum_{i=1}^{N} R(v_i) \geqslant k_{j+1} \\ l_L & \text{if } \sum_{i=1}^{N} R(v_i) \geqslant k_L \end{cases}$$

where:

$$k_1 = v_{min} + k_{in} \quad , \quad k_j = v_{min} + jk_{in},$$
$$k_{j+1} = v_{min} + (j+1)k_{in}, \cdots, k_H = v_{min} + Lk_{in}$$
$$k_{in} = \frac{v_{max} + v_{min}}{L+1}$$

Using the security values obtained by each agent (k_J) the WGAs agents are categorized in suitable groups. The SRV of the agent can be modeled using a N-player game with $X = \{1, 2, \cdots, n\}$ in which X is a set of players and each subset can be obtained as $V \subset N$, so that $\forall j \in V$ and $v_j \neq 0$ can be recognized as a coalition [5, 11]. The coalition of X agent in a group with k threshold of the security levels is indicative of the attack pattern and the L security level in the group. The aggregate value of the coalition is calculated as $R(c) = \sum_{i \in c} R(v_i)$ by sum total of SRV values of coalition members and is called the Coalition Function. Suppose $R(c) = \sum_{i \in c} R(v_i), v_i \in V, c \subset X$ is the C coalition value with C number of members. Then the Shapley value is defined as Formula (5).

$$SP(i) = \sum_{c \subset X\ ,\ i \in c} \frac{(c-1)!(n-c)!}{n!} \left[R(c) - R(c - \{i\}w)\right] \tag{5}$$

After calculating the $sp\ (i)$ for each agent obtained from their membership in the coalition, the security risk value of the attack to that agent can be classified in different security groups using different values of k_j and in accordance with each l_1, l_j, l_{j+1}, l_L group, a security risk value can be determined for each agent.

7 Simulation

For the purpose of simulating the proposed model efficiency a numerical example is needed. First some 20 assumed numbers are randomly produced in a matrix for 20 agents. In this matrix as per each agent some other numbers are also randomly produced as the attacker parameters such as b_1, b_2,etc. Next, using the GAMBIT game simulation software the NASH equilibrium point for each agent and attacker is obtained entered into the table [7, 8]. In the next step MATLAB software is used for calculating the Shapley value, through which the SRV

Table 4: The attacker agent parameters and their individual NASH equilibrium

	Attacker and Agent's params										Nash
Agent	b_1	$-b_2$	$f_i\%$	$-c_1$	c_2	c_3	$l_i\%$	$m_i\%$	$p_d\%$	$p_f\%$	v_i
N1	10	-90	0.49	-25	71	55	0.8	0.35	0.5	0.7	10.732
N2	20	-100	0.44	-81	20	10	0.3	0.9	0.8	0.2	3.97
N3	30	-50	0.5	-50	31	36	0.3	0.45	0.8	0.3	2.31
N4	40	-70	0.35	-70	40	50	0.4	0.7	0.7	0.4	2.11
N5	50	-50	0.25	-50	60	40	0.6	0.5	0.5	0.6	1.589
N6	60	-90	0.2	-60	50	60	0.5	0.6	0.6	0.5	2.547
N7	70	-40	0.6	-40	90	70	0.8	0.5	0.7	0.4	3.20
N8	80	-50	0.7	-90	60	50	0.7	0.4	0.8	0.7	1.235
N9	90	-90	0.3	-80	70	90	0.15	0.2	0.7	0.9	2.231
N10	100	-11	0.05	-5	100	90	1	0.1	0.4	0.2	0.0087
N11	5	-10	0.05	-10	100	90	1	0.1	0.5	0.2	0.96
N12	15	-85	0.44	-2	96	97	0.051	0.88	0.5	0.1	19.278
N13	25	-91	0.44	-81	15	14	0.05	0.99	0.5	0.1	5.7
N14	35	-98	0.34	-1.5	98	100	0.91	0.99	0.5	0.1	49.01
N15	45	-55	0.21	-65	33	31	0.35	0.65	0.7	0.2	2.385
N16	55	-41	0.2	-41	55	50	0.55	0.45	0.79	0.45	1.45
N17	65	-100	0.4	-10	71	85	0.8	0.21	0.5	0.3	10.02
N18	75	-25	0.7	-30	74	79	0.75	0.25	0.8	0.25	1.093
N19	85	-10	0.05	-80	15	15	0.14	0.85	0.5	0.5	0.381
N20	100	-10	0.01	-70	20	54	0.9	0.9	0.9	0.1	4.021

value of each agent is obtained and is positioned in one of L security levels. For this purpose 4 threshold value for L is considered and the agents are positioned in the suitable categories (coalitions). From the calculation of NE[6] values the vector for 20 WGAs agents are obtained. Afterwards using the k_{in} the threshold value for 4 coalition groups are calculated and using the determined figure and the v_i values, the accurate value of Shapley value is obtained. After calculating the sp value for each agent, it would be possible -with the aid from the k_{in} calculated for 4 desired security levels- to categorize the threat level for each agent in one of these 4 groups and determine the attack risk level. After classification of the agents in each of the groups the system manager can make the necessary decisions considering the obtained information. The threshold values relevant to each security level can be altered by the system manager so that the actual classification can be represented (See Table 4).

The threshold value for each one of the 4 groups is calculated and the values $\{9.08, 3.02, 20.1, 40.72\}$ for k_{in} are obtained, categorizing the calculated NASH values in 4 groups. In the next stage using the Shapley value formula and the k_{in} it would be possible to accurately calculate the $SP(i)$ value. Agents $\{n3, n4, n5, n6, n8, n9, n10, n11, n15, n16, n18, n19\}$, $\{n2, n13, n20\}$, $\{n1, n12, n17\}$, $\{n14\}$ are positioned in groups 1st, 2nd, 3rd, and 4th groups respectively. Accordingly a coalition so is formed and the security level of the groups can be considered as identical. Considering the Shapley value, the groups can be changed and different overlay networks can be formed to defend against the attacks.

In Figure 3, a number of Shapley values have been considered for different states, illustrating the formation of coalition in three states.

8 Conclusions

In this paper, a distributed model has been introduced for detecting the DIDS network intrusion base on the games theory. In traditional IDS the system administrators must continuously monitor the system and investigate all the occurrences and possible scenarios. In this study a game theory based model in distributed form in two cooperative and non-cooperative states has been proposed. In non- cooperative form (competitive) the security risk value is calculated using the Nash equilibrium for each of the agents and in the cooperative game by the agents belonging to a group, after formation of coalition and calculating the Shapley value the coalition security value is determined. Considering that in this model the agents WGAs have been distributed throughout the network, they can monitor the substantial transactions and exchanged data in the network through formation of an intrusion detecting overlay network and cover the best part of the network in distributed form. Additionally, in the event of weak or false detection, the agents through establishing coalition and collective calculation of SRV (Shapley value) the false detection percentage and issuing false alarms is

[6] Nash Equilibrium

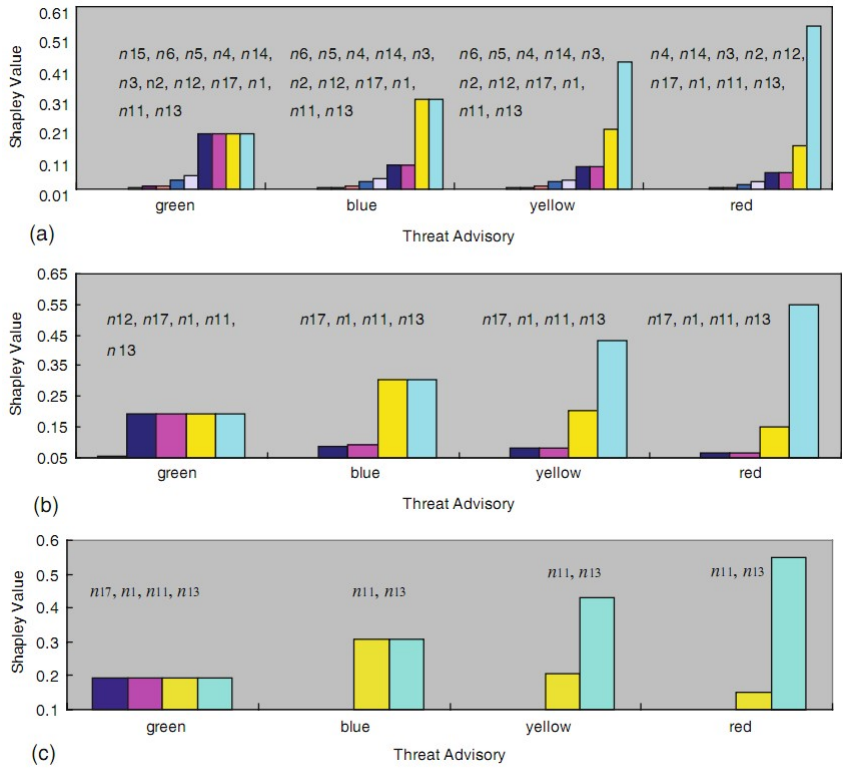


Figure 3: Three states of coalition configuration for Shapley values of (a) greater than 0.7 and (b) greater than 0.8 and (c) greater than 0.9.

minimized. In this model, based on the agent diagnosis (WGA) v_i the threshold values are updated (k_{in}) and such values are changed upon each time of coalition formation and practically the system in each instances of SRV calculation cycles performs new categorization. For the future continuation of this study, the researchers are planning to calculate the k_{in} in gradual form using Reinforcement Learning methods, meanwhile utilization of the previously gathered data.

References

[1] T. Alpcan, T. Basar, "A game theoretic approach to decision and analysis in network intrusion detection," in *IEEE Conference on Decision and Control*, pp. 2595–2600, 2003.

[2] A. Anurag, "Network neutrality: Developing business model and evidence based net neutrality regulation," *International Journal of Electronics and Information Engineering*, vol. 3, no. 1, pp. 1–9, 2015.

[3] G. Avni, T. Tamir, "Cost-sharing scheduling games on restricted unrelated machines," in *Proceedings of the 8th International Symposium on Algorithmic Game Theory (SAGT'15)*, pp. 7–9, 2015.

[4] Y. M. Chen, D. Wu, C. K. Wu, "A game theoretic framework for multi-agent deployment in intrusion detection systems," *Security Informatics*, Springer US, pp. 117–133, 2010.

[5] A. Dixit, S. Skeath, D. H. Reiley, *Games of Strategy*, W. W. Norton & Company, Paperback, 2014.

[6] R. H. Dong, D. F. Wu, Q. Y. Zhang, "The integrated artificial immune intrusion detection model based on decision-theoretic rough set," *International Journal of Network Security*, vol. 19, no. 6, pp. 880–888, 2017.

[7] Gambit: Software Tools for Game Theory, Dec. 2015. (`http://econweb.tamu.edu/gambit`)

[8] GAMBIT Software Tools for Game Theory, Dec. 2015. (`http://www.gambit-project.org`)

[9] M. Jain, D. Korzhyk, O. Vanek, M. Pechoucek, V. Conitzer, M. Tambe, "A double oracle algorithm for zero-sum security games on graphs," in *Proceedings of the 10th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'14)*, pp. 276–284, 2014.

[10] A. Jiang, Z. Yin, C. Kietkintveld, K. Leyton-Brown, T. Sandholm, M. Tambe, "Towards optimal patrol strategies for fare inspection in transit systems," in *Proceedings of the AAAI Spring Symposium on Game Theory for Security, Sustainability and Health*, pp. 145–153, 2014.

[11] P. Liu, W. Zang, and M. Yu, "Incentive-based modeling and inference of attacker intent, objectives, and strategies," *ACM Transactions on Information System Security*, vol. 8, no. 1, pp. 78–118, 2015.

[12] K. Maskat, "Mobile agents in intrusion detection system: Review and analysis," *Canadian Center of Sci-*

ence and Education, Modern Applied Science, vol. 5, no. 6, pp. 218–231, 2011.

[13] A. Nezarat, G. Dastghaibifard, "A game theoretic method for resource allocation in scientific cloud," *International Journal of Cloud Applications and Computing*, vol. 6, no. 1, pp. 15–41, 2016.

[14] A. Nezarat, Y. Shams, "A game theoretic-based distributed detection method for VM-to-hypervisor attacks in cloud environment," *Journal of Supercomputing*, doi:10.1007/s11227-017-2025-7, 2017.

[15] E. Popoola, A. O. Adewumi, "Efficient feature selection technique for network intrusion detection system using discrete differential evolution and decision," *International Journal of Network Security*, vol. 19, no. 5, pp. 660–669, 2017.

[16] S. E. Schechter, *Computer Security Strength and Risk: A Quantitative Approach*, Ph.D. Thesis, Harvard University, 2004.

[17] S. E. Schechter, "Toward econometric models of the security risk from remote attacks," *IEEE Security & Privacy*, vol. 3, no. 1, pp. 40–44, 2005.

[18] D. Singh, D. Patel, B. Borisaniya, and C. Modi, "Collaborative IDS framework for cloud," *International Journal of Network Security*, vol. 18, no. 4, pp. 699–709, 2016.

[19] SNORT, *Intrusion Detection Systems NIST Special Publication on Intrusion Detection Systems*, Dec. 2015. (`http://www.snort.org/docs/nistids.pdf`)

[20] Symantec Corporation, Dec. 2015. (`http://www.symantec.com/index.jsp`)

[21] A. Tayal, N. Mishra and S. Sharma, "Active monitoring & postmortem forensic analysis of network threats: A survey," *International Journal of Electronics and Information Engineering*, vol. 6, no. 1, pp. 49–59, 2017.

[22] X. Zhao, "The optimization research of the multimedia packets processing method in NIDS with 0/1 knapsack problem," *International Journal of Network Security*, vol. 17, no. 3, pp. 351–356, 2015.

Biography

Amin Nezarat is an assistant professor at Payam Noor University since 2013. He received his B.S. in Computer Science from the Shahid Bahonar University, M.S. in IT engineering from the Shiraz University, PhD. in software systems from Shiraz University in 2003, 2010 and 2014, respectively. His research interests include cloud and grid computing, parallel computing, resource scheduling, big data, multi agent systems, game theory, electronic commerce.

Generalized PVO-K Embedding Technique for Reversible Data Hiding

Jian-Jun Li[1], Yun-He Wu[1], Chin-Feng Lee[2] and Chin-Chen Chang[3]
(Corresponding author: Chin-Chen Chang)

Department of Computer Science and Technology & Hangzhou Dianzi University[1]
1158, No.2 Rd., Jianggan, Hangzhou 10336, China
Department of Information Management & Chaoyang University of Technology[2]
168, Jifeng E. Rd., Wufeng, Taichung 41349, Taiwan
Department of Information Engineering and Computer Science & Feng Chia University[3]
100, Wenhwa Rd., Seatwen, Taichung 40724, Taiwan
(Email: alan3c@gmail.com)
(Received Sept. 29, 2016; revised and accepted Dec. 11, 2016)

Abstract

Recently, several reversible information hiding methods based PVO (pixel value ordering) techniques have been proposed, in these methods, secret data always are embedded in pixels with largest or smallest value in the block. In order to make use of the block with multiple largest-valued (or smallest-valued) pixels, a PVO-K method was proposed, which treats K largest-valued (or smallest-valued) pixels as a unit to embed secret data, and all K pixels are modified together to embed one bit of information. In this paper, we propose a generalized PVO-K method (GePVO-K) that takes full advantage of these pixels with largest or smallest values by embedding K bits of secret data into the K pixels. As a result, the GePVO-K method has greater embedding capacity than the PVO-K method. The superiority of the GePVO-K scheme was verified by the experimental results.

Keywords: Pixel Value Ordering; Prediction Error Expansion; Reversible Data Hiding

1 Introduction

Image data hiding is the technology in which secret data, such as authentication or, copyright information, are embedded in a digital image [7, 25]. In data hiding techniques, it is critically important that the recipient be able to extract the secret data completely from the camouflage image, but, at the same time, any decreases in the quality of the image should not particularly evident, and it is especially important that, the difference cannot be detectable by the human eye.

Information hiding processes can be divided into two categories based on the technology they use. One technology is data hiding with distortion, and the other technology is non-distortion data hiding, which also is called RDH (reversible data hiding). Least Significant Bit (LSB) [4, 23], Revisited Matching [13], and Exploiting Modification Direction (EMD) [26] are well-known, non-reversible data hiding techniques that are simple and have high embedding capacity. Compared with an ordinary data hiding algorithm, RDH must take more requirements into consideration. It also requires that the original cover image be recoverable after the secret data have been extracted from the camouflaged image. That is to say, RDH is a special data hiding method that is always applied for scenarios that are sensitive to image distortion, such as processing military, medical, or remote sensing images.

To date, many reversible data hiding techniques have been proposed. The first kind of RDH method was based on lossless compression [2, 3, 5], and this method acquires embedding space through lossless compression of a specific part of the digital cover image. As a result, they usually have low embedding capacity and produce significant distortion of the image. In 2006, Tian et al. proposed an important spatial data hiding algorithm called "difference expansion" (DE) [20]. They overcame the limitations of embedding secret data through lossless compression, and they focused on diffusion of the difference between pixel pairs to embed secret data reversibly. Later, several improved methods involving DE were proposed. One method tried to decrease the size of the location map [8, 12, 24], the second method was based on integer transform [10, 16, 22], and the third method involved prediction error expansion (PEE) [1, 6, 19]. Image data have spatial redundancy that is caused by the correlation between adjacent pixels in the image. The PEE method has great embedding capacity because it can take advantage of the spatial redundancy of a digital image. The PEE

method was first proposed by Thodi et al. [19], after which Hu et al. [16] made some improvements by constructing a location map that depends on the payload, thereby decreasing the size of the compressed location map.

In addition to the DE technique, Ni et al. proposed another important RDH method, i.e., the histogram shift (HS)-based technique [14]. Later, Lee et al. [9] improved the HS method by using a histogram of the difference between adjacent pixels, and this method improved the embedding capacity and reduced image distortion.

Recently, Li et al. [11] proposed a new RDH method based on pixel value ordering (PVO). This method combines DE and HS with PEE and uses the PVO technique to embed secret data in a block-by-block manner. For each block, the pixel values are sorted in ascending order, then, the largest-valued pixel is predicted by the second-largest pixel, and the smallest-valued pixel is predicted by the second-smallest pixel. Thus, the second-largest and the second-smallest pixel values in the block remain unchanged in the embedding phase, and the largest pixel value may become larger since it always is increased, and the smallest pixel value becomes smaller after being decreased. Therefore, the order of pixel values in the block remains unchanged. Typically, the minimum prediction error is non-positive, and the maximum prediction error is non-negative, so the RDH algorithm based on PVO regards the non-negative and non-positive values, which occur most frequently, as the carriers of secret data. The prediction error in the range of "-1" to "1" is defined as the peak value, and the secret data are embedded in the peaks using HS technology. Then, the smallest and largest pixel values are modified according to the prediction error values.

However, since the local pixel values are correlative, there will be many prediction errors 0, which are ignored by the PVO method. In view of this phenomenon, Peng et al. [17] proposed an improved PVO method, i.e., IPVO, and they used the pixel location number in the blocks before ordering the pixel values to optimize the process of generating the prediction error. In addition to IPVO, another method, known as PVO-K, was proposed in [15], and its aim also was to improve the PVO. The PVO-K method treats K identical largest-valued or smallest-valued pixels as a unit to embed secret data. Compared with PVO, when the largest pixel value and the second-largest value (or the smallest pixel value and second-smallest pixel value) are equal, they are treated as a unit, so this block may be still used to embed secret data. Obviously, the PVO method is a special case of the PVO-K method, i.e., when K=1. In [15], the embedding capacity was improved by using a combination of the PVO-1 and PVO-2 methods. But in the smooth block, there often are more identical largest or smallest pixel values, and the PVO-K method may change all K pixel values to embed just one bit of secret data, so there is still room for improving the PVO-K method. Besides, a new path of methods has been proposed in 2015, they break the block restrictions of other PVO-based methods, therefore make more use of the pixels that can be utilized to embed secrete data, and significantly enhance the performance of the PVO-based method, like the PPVO [18] and the method of Wang et al. [21]. In this paper, a strategy is presented concerning ways to improve the PVO-K method so that K bits of secret data can be embedded into K largest-valued or smallest-valued pixels. We also proposed a new way to produce a special block and compared the performance with traditional treatments.

The remaining sections of this paper are organized as follows. Several PVO-based methods are introduced in Section 2. In Section 3, a new generalized PVO-K scheme is proposed. Section 4 presents relevant experiments and the analysis of the results. Our conclusions are presented in Section 5.

2 Related Works

In this section, two PVO-based reversible data hiding methods are introduced briefly, i.e., PVO [11], PVO-K [15].

2.1 RDH Method Based On PVO

The PVO method proposed by Li et al. provided a new predictor for the prediction error expansion, with both largest and smallest pixel values being used in a block for embedding data. The embedding process is firstly divide the cover image into blocks of pixels, and number the pixels in each block, i.e.,$(x_1, x_2, ..., x_{n1\times n2})$. Then, sort the pixels in ascending order to get an ordered sequence $(x_{\pi(1)}, x_{\pi(2)}, ..., x_{\pi(n1\times n2)})$.

After that, count two prediction errors according Equation (1), wherein, the non-negative integer $d_{\max}$ represents the difference between the largest pixel value and the second-largest pixel value; and a non-positive integer $d_{\min}$ represents the difference between the smallest pixel value and the second-smallest pixel value.The secret message $b \in \{0, 1\}$can be embedded when the maximum prediction error is 1 or the minimal prediction error is -1. Prediction errors are modified according to Equation (2)and Equation (3).At last, revise the largest and smallest pixel values using Equation (4) and proceed to the next block until all blocks have been processed or all secret data have been embedded.

$$\begin{cases} d_{\max} = x_{\pi(n1\times n2)} - x_{\pi(n1\times n2-1)} \\ d_{\min} = x_{\pi 1} - x_{\pi 2} \end{cases} \tag{1}$$

$$d'_{\max} = \begin{cases} d_{\max}\ if d_{\max}{=}0 \\ d_{\max}{+}b\ if d_{\max}{=}\ 1 \\ d_{\max}{+}1\ if d_{\max} > 1 \end{cases}, \tag{2}$$

$$d'_{\min} = \begin{cases} d_{\min}\ if d_{\min}{=}0 \\ d_{\min} - b\ if d_{\min}{=}-1 \\ d_{\min} - 1\ if d_{\min} < -1 \end{cases}. \tag{3}$$

$$\begin{cases} {x'}_{\pi(n1\times n2)} = x_{\pi(n1\times n2-1)} + {d'}_{\max} \\ {x'}_{\pi(1)} = x_{\pi(2)} + {d'}_{\min} \end{cases}. \tag{4}$$

Because the order of the pixel values remains unchanged after embedding the secret data, the secret data can be extracted in the extraction phase from the largest-valued and smallest-valued pixels according to reverse process of the embedding procedure. At the same time, the pixel values can be changed back to the original values.

2.2 RDH Method Based On PVO-K

Like IPVO, PVO-K also was proposed for the purpose of using the prediction error "0", which is discarded in the PVO method, but the difference is that PVO-K treats the largest or smallest pixel values as a unit for embedding secret data. Similar to these methods, we take the procedure of embedding secret data into a maximum number of pixels as an example; assume that the sorted pixel values in a block are:$x_{\pi(1)} \leq ... \leq x_{\pi(n1\times n2-K)} < x_{\pi(n1\times n2-K+1)} = ... = x_{\pi(n1\times n2)}$,where K is the number of largest-valued pixels, and the prediction error is calculated using Equation (5).

$$d_{\max} = x_{\pi(n1\times n2-K+1)} - x_{\pi(n1\times n2-K)}. \quad (5)$$

When the prediction error is "1", one bit of secret data can be embedded; otherwise, the K pixel values are shifted. The prediction errors are modified by Equation (6).

$$d'_{\max} = \begin{cases} d_{\max}+b \ if d_{\max}= 1 \\ d_{\max}+1 \ if d_{\max} > 1 \end{cases}. \quad (6)$$

Then the largest pixel values are modified by Equation (7), where, $i \in \{n1 \times n2 - K + 1, n1 \times n2 - K + 2, ..., n1 \times n2\}$.

$$x'_{\pi(i)} = x_{\pi(i)} + d'_{\max}. \quad (7)$$

The common factor of PVO-K and other PVO-based methods is that the original block sorting remains constant after embedding the secret data, which makes the extraction process more convenient.

3 Proposed Scheme

In this section, we propose a generalized scheme for the PVO-K method with respect to embedding capacity, and it is called GePVO-K. First, we introduce how to embed one bit of secret data in each largest-valued pixel by modifying the largest and the second-largest pixel values. Some examples are provided to demonstrate our approach. Then, the process of embedding secret data in each smallest-valued pixel is presented. Finally, we show the detailed steps of the embedding and extraction procedures.

3.1 Embedding Secret Data in Largest-valued Pixels and Data Extraction Procedure

As mentioned in the previous sections, if the PVO-K algorithm embeds one bit of secret data in a block that has K largest-valued or smallest-valued pixels, all of the K pixels must be modified in the same way. Ou et al. [15] indicated that when PVO-1 and PVO-2 are used together to increase the embedding capacity of traditional PVO-based methods; however if $K > 2$, the block should not be used to embed secret data, because a larger K will lead to a greater distortion caused by more changes in the pixels values. In nature images, especially in the blocks of the smooth region, K is often greater than 2, so the smooth region always is ignored, which makes less embedding capacity. For this phenomenon, we propose an improved method that still utilizes the largest-valued and smallest-valued pixels in the block to embed secret data, but one bit of secret data can be embedded in each pixel. Here, we present the details of embedding secret data in the largest-valued pixels as well as the extracting procedure.

3.1.1 Embedding Secret Data in Largest-valued Pixels

First, the cover image should be divided into blocks. Let the size of block B be $n_1 \times n_2$. Then, each block is visited in a zigzag manner to establish a location map, and the rules for establishing the map are as follows. If the block has the pixel values that may overflow/underflow, such as "0," "1," "254," "255," the block's position is recorded as "2;" if all of the pixel values in the block are the same, the block's position is recorded as "1;" the remaining blocks are normal blocks, and their positions are recorded as "0s."

Next, we deal with each block depending on the following cases:

Case 1: If the position number of the block in the location map is $LM(B) = 2$, the block is not used to embed secret data, and it is skipped.

Case 2: If the position number of the block in the location map is $LM(B) = 1$, i.e., all pixel values are equal in the block B, we keep the first pixel value unchanged and then embed the secret data in the remaining pixels and the pixel values are modified by Equation (8) in a zigzag manner. It states that if a to-be-embedded bit $b = 0$, do not change the pixel value; if $b = 1$, increase the pixel value by one.

$$x'_{\pi(i)} = \begin{cases} x_{\pi(i)} \ if i= 1 \\ x_{\pi(i)} + b_{i-1} \ if i= 2,3,...,\mathrm{n1} \times \mathrm{n2} \end{cases}. \quad (8)$$

Case 3: If the position number of the block in location map $LM(B) = 0$, we number the pixels in a zigzag scanning order to get $B\,(x_1, x_2, ..., x_{n1\times n2})$, and then we sort the pixel values in ascending order to obtain

a sorted block $B_\pi\left(x_{\pi(1)}, x_{\pi(2)}, ..., x_{\pi(n1\times n2)}\right)$. Assume that the ordering result is:$x_{\pi(n1\times n2-K-L)} < x_{\pi(n1\times n2-K-L+1)} = ... = x_{\pi(n1\times n2-K)} < x_{\pi(n1\times n2-K+1)} = ... = x_{\pi(n1\times n2)}$. That is, there are K largest pixels $x_{\pi(n1\times n2-K+1)} = x_{\pi(n1\times n2-K+2)} = ... = x_{\pi(n1\times n2)}$, and L second-largest pixels $x_{\pi(n1\times n2-K-L+1)} = x_{\pi(n1\times n2-K-L+2)} = ... = x_{\pi(n1\times n2-K)}$. Then, we calculate the maximum prediction error using Equation (9).

$$d_{\max} = x_{\pi(n1\times n2-K+1)} - x_{\pi(n1\times n2-K)}. \quad (9)$$

Case 3-1: If $d_{\max} > 1$, this block is not fit to be used to embed secret data, and all largest pixel values should be increased by one as Equation (10).Where,$i \in \{n1\times n2-K+1, n1\times n2-K+2, ..., n1\times n2\}$.

$$x'_{\pi(i)} = x_{\pi(i)} + 1. \quad (10)$$

Case 3-2: if $d_{\max} = 1$, K bits secret data can be embedded into the largest-valued pixels. In order to correctly find which pixels were embedded secret data during extracting process, the difference between the second-largest pixels and the third-largest pixels need to be expanded also. Specifically, after embedding secret data, the original largest-valued pixels may be still kept in the position of largest-valued pixels or some of them may change to the second-largest pixels. To distinguish these two situations for the purpose of extracting secret data and recovering the original pixels, the difference between the second-largest pixels and the third-largest pixels can be used as a judgment condition and its detailed usage will be presented in the extracting procedure. So in this embedding case, first we increase the K largest pixel values and the L second-largest pixel values by one; then we embed the secret data, $b_r \in \{0,1\}(r = 1, 2, ..., K)$, into the largest-valued pixels in the numbering order. If $b_r = 0$, keep the largest pixel value unchanged; if $b_r = 1$, increase the largest pixel value by one. In summary, the pixel values are modified by Equation (11). Where, $b_i \in \{0,1\}$, $i \in \{n1\times n2-K+1, n1\times n2-K+2, ..., n1\times n2\}$ and $j \in \{n1\times n2-K-L+1, n1\times n2-K-L+2, \cdots, n1\times n2-K\}$.

$$\begin{cases} x'_{\pi(i)} = x_{\pi(i)} + b_{i-n1\times n2+K} + 1 \\ x'_{\pi(j)} = x_{\pi(j)} + 1 \end{cases}, \quad (11)$$

3.1.2 Extracting Secret Data From Larger-valued Pixels and Restoring the Pixel Values

As can be seen from Equation (11), for a normal block (i.e., its recorded number in the location map is 0), the original ordering may be changed after the secret data have been embedded. Because some largest-valued pixel of the original block may become the second-largest after embedding the secret data, so, in the camouflage block, information can be hidden only in the largest-valued or the second-largest pixels. We can determine whether the secret data are completely hidden in the largest-valued pixels or in both the largest-valued and the second-largest pixels. Therefore, we can completely extract the secret data revise the pixel values according as follows.

First, we divide the camouflage image into blocks as we did in the embedding procedure, then, we handle each block depending on the following cases:

Case 1: If the position number of the camouflage block in location map $LM(B) = 2$, there are no hidden secret data, and the original block is the same as the camouflage block.

Case 2: If the position number of the camouflage block in location map $LM(B) = 1$, we extract the secret data starting from the second pixel. First, we calculate the prediction error d_i according to Equation (12). If $d_i = 0$, extract secret data $b_{i-1} = 0$, keeping the pixel value unchanged; if $d_i = 1$, extract the secret data $b_{i-1} = 1$, decreasing the pixel value by one, as shown in Equation (13). Where, $i \in \{2, 3, ..., n1\times n2\}$.

$$d_i = x'_{\pi(i)} - x'_{\pi(1)}, \quad (12)$$

$$\begin{aligned} x_{\pi(1)} &= x'_{\pi(1)}, \\ x_{\pi(i)} &= \begin{cases} x'_{\pi(i)}, b_{i-1}{=}\,0 \ if d_i{=}0 \\ x'_{\pi(i)} - 1, b_{i-1}{=}\,1 \ if d_i{=}\,1 \end{cases}. \end{aligned} \quad (13)$$

Case 3: If the position number of the camouflage block in location map $LM(B) = 0$, we sort the pixels in ascending order to obtain $B'_\pi\left(x'_{\pi(1)}, x'_{\pi(2)}, ..., x'_{\pi(n1\times n2)}\right)$, assuming that the ordering results are $x'_{\pi(n1\times n2-R-S-T)} < x'_{\pi(n1\times n2-R-S-T+1)} = ... = x'_{\pi(n1\times n2-R-S)} < x'_{\pi(n1\times n2-R-S+1)} = ... = x'_{\pi(n1\times n2-R)} < x'_{\pi(n1\times n2-R+1)} = ... = x'_{\pi(n1\times n2)}$. Which means there are R largest-valued pixels (m1), S second-largest pixels (m2), and T third-largest pixels (m3).

m1: $x'_{\pi(n1\times n2-R+1)} = x'_{\pi(n1\times n2-R+2)} = ... = x'_{\pi(n1\times n2)}$;

m2: $x'_{\pi(n1\times n2-R-S+1)} = x'_{\pi(n1\times n2-R-S+2)} = ... = x'_{\pi(n1\times n2-R)}$;

m3: $x'_{\pi(n1\times n2-R-S-T+1)} = x'_{\pi(n1\times n2-R-S-T+2)} = ... = x'_{\pi(n1\times n2-R-S)}$.

Two prediction errors are calculated according to Equation (14). If T does not equal to 0, it's very plain that the prediction errors $d1 \geq 1$ and $d2 \geq 1$; if T equals to 0which means that there are not the third-largest pixels, only one prediction error $d1 \geq 1$

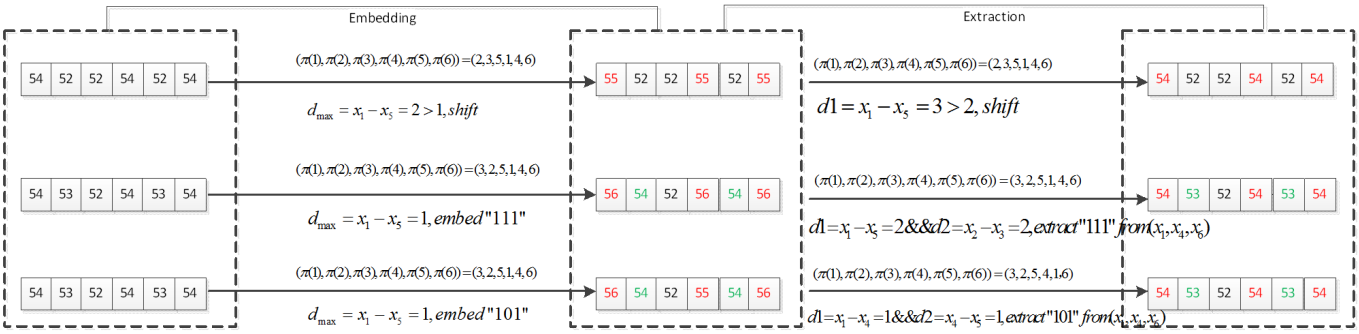


Figure 1: Example of embedding secret data in largest-valued pixels and extracting procedure

is obtained. So we process the block on the basis of the following Cases 3-1–3-3.

$$\begin{cases} d1 = x'_{\pi(n1\times n2-R+1)} - x'_{\pi(n1\times n2-R)} \\ d2 = x'_{\pi(n1\times n2-R-S+1)} - x'_{\pi(n1\times n2-R-S)} \end{cases} \quad (14)$$

Case 3-1: If $d1 > 2$, there are no hidden secret data, and we decrease R largest pixel values by one, the pixel values are revised according to Equation (15).

$$x_{\pi(i)} = x'_{\pi(i)} - 1, \quad i \in \{n1\times n2 - R + 1, n1\times n2 - R + 2, ..., n1\times n2\}. \quad (15)$$

Case 3-2: If $d1 \leq 2$ and $d2 = 1$, the secret data were embedded in the m1 and m2 zones. First, we decrease the value of pixels in the m1, m2, and m3 zones by one. (Notice that d1 and d2 remain unchanged in keeping with Equation (16).

$$x_{\pi(i)} = x'_{\pi(i)} - 1, \quad i \in \{n1\times n2 - R - S - T + 1, n1\times n2 - R - S - T + 2, ..., n1\times n2\}. \quad (16)$$

Then, we extract the secret data $S(B) = \{b_i | b_i \in \{0,1\}, i = 1, 2, ..., R+S\}$ from the pixels in the m1 and m2 zones, depending on the numbering order. We count D_i one by one, where D_i is the difference in the values between the pixels in the m1 or m2 zone and the pixels in the m3 zone, as indicated in Equation (17). If $D_i = 2$, decrease the pixel value by one and extract secret data $b_i = 1$; if $D_i = 1$, let the pixel value remain unchanged and extract secret data $b_i = 0$. The extraction procedure depends on Equation (18). Where, $i \in \{n1\times n2 - R - S + 1, n1\times n2 - R - S + 2, ..., n1\times n2\}$.

$$D_i = x'_{\pi(i)} - x'_{\pi(n1\times n2-R-S)}, \quad i \in \{n1\times n2 - R - S + 1, n1\times n2 - R - S + 2, ..., n1\times n2\}. \quad (17)$$

$$x_{\pi(i)} = \begin{cases} x'_{\pi(i)}, b_i = 0 \ if D_i = 1 \\ x'_{\pi(i)} - 1, b_i = 1 \ if D_i = 2 \end{cases} \quad (18)$$

Case 3-3: If $d1 \leq 2$ and $(d2 \geq 2 || T = 0)$, this means secret data were embedded in m1 zones, and the secret data fragment $S(B)$is a binary string only including 1 (in this case, $d1 = 2$), or an all 0 binary string (in this case, $d1 = 1$). So, first, we decrease the value of pixels in the m1 and m2 zones by one, and then extract the secret data from m1 in the numbering order. The secret data are determined by the value D_i (Equation (19)), which is the difference between the pixels in m1 and the pixels in m2. If $D_i = 2$, decrease the pixel value by one and extract secret data $b_i = 1$; if $D_i = 1$, keep the pixel value unchanged and extract secret data $b_i = 0$. In short, we can recover the original pixel value and extract secret data as Equation (20). Where, $j \in \{n1\times n2 - R - S + 1, n1\times n2 - R - S + 2, ..., n1\times n2\}$ and $i \in \{n1\times n2 - R + 1, n1\times n2 - R + 2, ..., n1\times n2\}$.

$$D_i = x'_{\pi(i)} - x'_{\pi(n1\times n2-R)}, \quad i \in \{n1\times n2 - R + 1, n1\times n2 - R + 2, ..., n1\times n2\}. \quad (19)$$

$$x_{\pi(j)} = x'_{\pi(j)} - 1, \quad x_{\pi(i)} = \begin{cases} x'_{\pi(i)}, b_i = 0 \ if D_i = 1 \\ x'_{\pi(i)} - 1, b_i = 1 \ if D_i = 2 \end{cases}. \quad (20)$$

3.1.3 Example of Embedding and Extraction Procedures

For a better illustration, there are several examples to demonstrate the above steps in Figure 1. We assume that the block size is $n1 = 2$ and $n2 = 3$. As can be seen from Figure 1, three blocks are selected as examples. All the pixel values in these three blocks are numbered and sorted firstly, for instances, the original pixel sequence in first block is (54, 52, 52, 54, 52, 54), the sorted pixel values are (52, 52, 52, 54, 54, 54) and their numbers in the original block are (2, 3, 5, 1, 4, 6). There are three largest pixels and three second-largest pixels in first block, according embedding rules Case 3 in 3.1.1 part of Section 3.1, we calculate the prediction error $d_{\max} = x_1 - x_5 = 2$, because

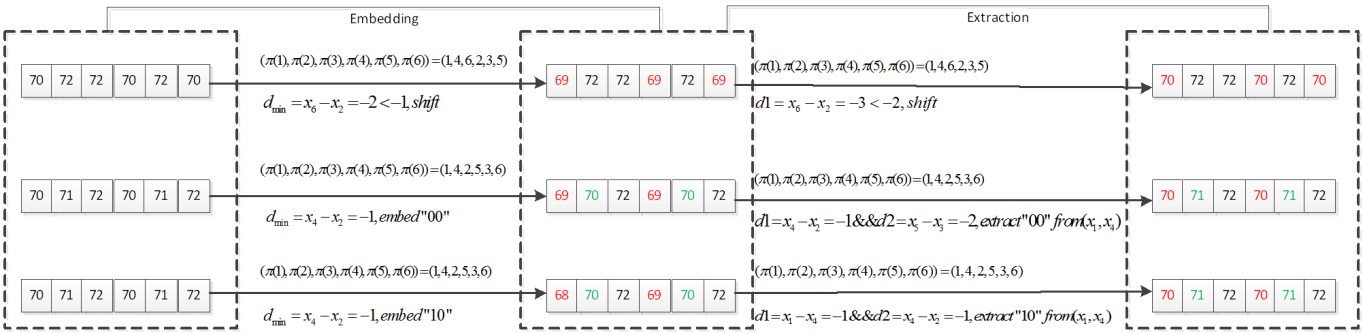


Figure 2: Example of embedding secret data in smallest-valued pixels and extracting procedure

$d_{\max} > 1$, this block is not fit to be used to embed secret data, and all largest pixel values should be increased by one to get the pixel values (52, 52, 52, 55, 55, 55); after that, the pixels need to be moved to their original position according to their number (2, 3, 5, 1, 4, 6). Finally, the pixel values of first block become to (55, 52, 52, 55, 52, 55).

The embedding procedures of the other two blocks are similar. For the second block, the original pixel sequence is (54, 53, 52, 54, 53, 54). After pixel numbering and sorting, the pixels become (52, 53, 53, 54, 54, 54) and their number are (3, 2, 5, 1, 4, 6). There are three largest pixels and two second-largest pixels, we calculate the prediction error $d_{\max} = x_1 - x_5 = 1$. According to the embedding rules in 3.1.1 part of Section 3.1, three bits secret data can be embedded into three largest pixels, the three largest pixels and two second-largest pixels need to be firstly increased by one to get (52, 54, 54, 55, 55, 55), and then three secret bits (we choose111) are embedded into three largest pixels to get (52, 54, 54, 56, 56, 56). Next, move them to their original position according to the number sequence (3, 2, 5, 1, 4, 6), so the final pixel sequence is (56, 54, 52, 56, 54, 56). For the third block, its pixels are the same as the second block. The difference is that we embed the secret data that has both 0 and 1 into three largest pixels for better demonstration of different extraction cases below.

The first step of the extraction procedures is numbering and sorting the pixels in the block too, then the blocks are handled according the prediction errors. As shown in Figure 1, for the first block, the sorted pixel sequence is (52, 52, 52, 55, 55, 55), and the prediction error $d_1 = x_1 - x_5 = 3 > 2$, which means there are no secret data and three largest pixels need to be decreased by one to get (52, 52, 52, 54, 54, 54), then move them to their original position to get the original block.

For the second block, the sorted sequence is (52, 54, 54, 56, 56, 56), prediction error $d_1 = x_1 - x_5 = 2 \leq 2$. In this case, we easily know that there are embedded secret data, but we cannot determine whether it is hidden within all the largest pixels or both the largest pixels and the second-largest pixels. Because if the embedded secret data is an all "0" string or all "1" string, all the largest pixels remains largest after embedding procedure, such as the second block; if the embedded secret data is a string with both "0" and "1", some of the largest pixels becomes to the second largest pixels after embedding procedure, such as the third block. Thus, we need to calculate a prediction error between the second-largest pixels and the third-largest pixels to distinguish these two cases. As presented in Figure 1, another prediction error of the second block $d_2 = x_2 - x_3 = 2 \geq 2$, according Case 3-3 in B part of Section 3.1, all the secret data are embedded in the largest pixels. Therefore, we extract the three bits of secret data 111 and recover the pixels to get (52, 53, 53, 54, 54, 54), then we can obtain the original sequence (54, 53, 52, 54, 53, 54) by moving them to their original positions. For the third block, the sorted sequence is (52, 54, 54, 55, 56, 56), the prediction errors $d_1 = x_1 - x_4 = 1 \leq 2$ and $d_2 = x_4 - x_5 = 1$, which means that the secret data is in both two largest pixels and one second-largest pixels. Therefore we can extract the secret data 101 and recover the original pixels (54, 53, 52, 54, 53, 54), according to Case 3-2 in 3.1.2 part of Section 3.1.

3.2 Embedding Secret Data in Smallest-valued Pixels and Data Extraction Procedure

Except for the special blocks ($LM(B) = 1$ or $LM(B) = 2$), we also used the smallest-valued pixels in the normal block to embed secret data. In order to further demonstrate the proposed methods, Figure 2 presents several examples of embedding data into the smallest-valued pixels and the data extraction procedure. Similar to the example in Figure 1, there are three embedding cases. Case 1: The first block is not fit to embed secret data; Case 2: The second block embeds an all "0" or "1" string; Case 3: The third block embeds a string with both "0" and "1".

The specific steps of embedding secret data in the smallest pixels and extraction procedures are significant similar to Figure 1, in order to avoid duplication, we just take the complex cases (the third block) as an example to illustrate the procedure. For the third block, the original sequence (70, 71, 72, 70, 71, 72) is numbered and

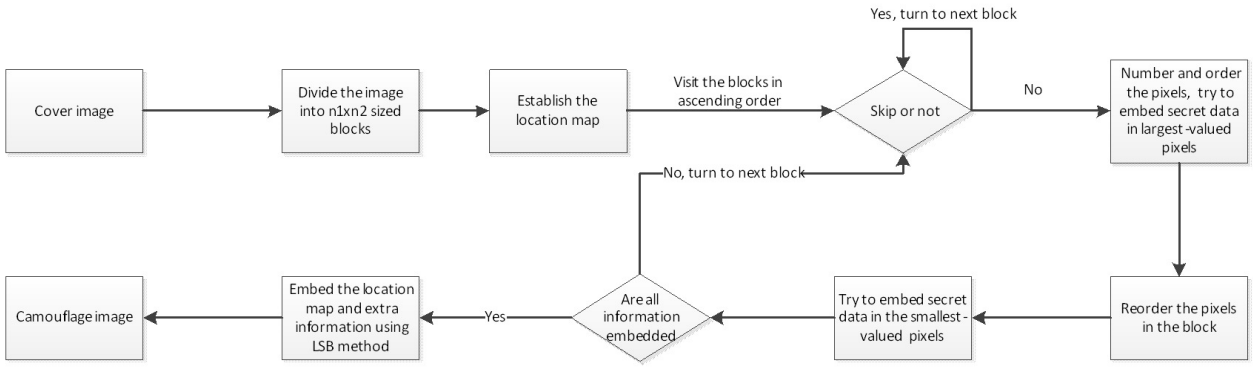


Figure 3: The flowchart of proposed data embedding procedure

sorted to get (70, 70, 71, 71, 72, 72), where their number are (1, 4, 2, 5, 3, 6). There are two smallest pixels and two second-smallest pixels, and the prediction error $d_{\min} = x_4 - x_2 = -1$. Thus two bits of secret data can be embedded in two smallest pixels. The two smallest pixels and two second-smallest pixels need to be decreased by one to get (69, 69, 70, 70, 72, 72), and then we embed "10" in two smallest pixels in two smallest pixels to get (68, 69, 70, 70, 72, 72). The last step of the embedding procedure is to move them to their original positions to get (68, 70, 72, 69, 70, 72). During extraction procedures, the first step is the same as embedding procedure. After numbering and sorting, we can get the sequence (68, 69, 70, 70, 72, 72). Next, we compute the prediction errors $d_1 = x_1 - x_4 = -1$, and $d_2 = x_4 - x_2 = -1$. It is evident that two secret data bits are embedded: one in the smallest pixel and one in the second-smallest pixel. Then, the extract procedure gets "10" from the smallest pixel and the second-smallest pixel and then the recovery process increases the smallest pixel by two, the second-smallest pixel by one, and two third-smallest pixels are also increased by one to get the sequence (70, 70, 71, 71, 72, 72). The last step is also to move the sequence (70, 70, 71, 71, 72, 72) according to their original positions; finally, we can obtain the original sequence (70, 71, 72, 70, 71, 72).

3.3 Proposed Data Embedding Procedure

In this section, we describe the detailed steps of the embedding phase in the proposed scheme. For the normal block, first, we embed the information into the largest-valued pixels, and, then, we reorder the pixels in the block; next, the smallest-valued pixels also are utilized to embed secret data. Like PVO-K, there also is the possibility of an overflow/underflow situation. So, we build a location map to record the position of the pixels that may overflow/underflow. In addition, we record some auxiliary information, as shown in Table 1. We assume that the size of the cover image is $H \times W$ and that the minimal and maximal block sizes are 2×2 and 4×4, respectively.

The detailed steps of the embedding phase are shown as follows, and Figure 3 is the embedding flowchart.

Step 1. Divide the cover image into $n1 \times n2$ -sized blocks and then visit each block in a zigzag manner to establish the location map according to the rules in Section 3.1. After that, two binary bits are used to represent a value in the location map and then the location map is compressed using arithmetic coding, a lossless data compression, to reduce its length.

Step 2. For each block B, if $LM(B) = 2$, skip; if $LM(B) = 1$, embed secret data in the pixels except for the first one; if $LM(B) = 0$, number and order the pixels in the block, try to embed secret data in the largest-valued pixels, and then reorder the pixel values in the block and try to embed the secret data in the smallest-valued pixels.

Step 3. When embedding the secret data is completed, embed first $2 \times \log_2((H \times W)/(n1 \times n2)) + \log_2(H \times W) + L_1 + 4$ least significant bits (LSB) of the pixels in the cover image into the remaining blocks and then record the last embedding position.

Step 4. Use the LSB method [4] to embed extra information and the compressed location map into the cover image from the first pixel.

3.4 Proposed Data Extraction Procedure

The corresponding proposed data extraction procedure is presented in this section; its flowchart is shown in Figure 4.

Step 1. First, extract the extra information and the compressed location map by using LSB method, and then decompress the location map to obtain LM.

Step 2. Next, divide the camouflaged image into blocks depending on the size that was extracted from Step1. Then, visit the blocks in reverse order, which means the extraction procedure must start from the last embedding position. For each block, if $LM(B) = 2$,

Table 1: The extra information

Extra information	Purposes	Memory required
The compressed location map	Record the special blocks	L_1 bits
The length of compressed location map	Correctly extract the location map	$2\times\log_2((H\times W)/(n1\times n2))$ bits
Block size $n1\times n2$	Divide the camouflage image	4 bits
The last position of embedding secret data i and j	Reversely extracting the secret data.	$\log_2(H\times W)$ bits

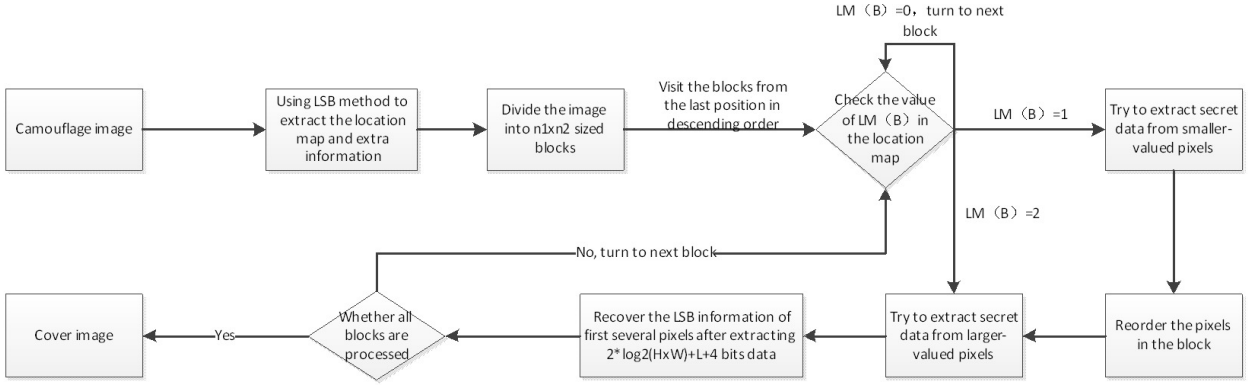


Figure 4: The flowchart of proposed data extraction procedure

skip; if $LM(B)=1$, extract secret data from the pixels in zigzag order except the first one; if $LM(B)=0$, number and order the pixels and try to extract the secret data from the smaller-valued pixels and then reorder the pixels and try to extract data from the larger-valued pixels.

Step 3. After extracting $2\times\log_2((H\times W)/(n1\times n2))+\log_2(H\times W)+L_1+4$ bits of data, revise the LSB information of the first several pixels that were modified to embed the compressed location map and extra information. Then, continue to extract secret data from the remaining blocks until all blocks have been processed.

To better demonstrate the embedding and data extraction procedures, a detailed example is given in Figure 5.

4 Experimental Results

In this section, first, we describe the experimental environment and the evaluation criteria in Section 4.1. Then, the results are discussed in five parts. Section 4.2 is the self-analysis of the proposed scheme in which we discuss the relationship between embedding capacity and the image quality; Section 4.3 compares the performance of GePVO-K with that of PVO-K by using a different block size; in Section 4.4, we analyze two methods of dealing with special blocks; since the proposed scheme is aimed mainly at enhancing the embedding capacity, several PVO-based methods were compared with the proposed scheme in terms of maximum embedding capacity in Section 4.5; Section 4.6 shows the multi-level embedding performance of several PVO-based methods, including GePVO-K.

4.1 Experimental Environment and the Evaluation Criteria

We tested eight cover images using the MATLAB R2010a Platform; all of the cover images were 8-bit grayscale images, and the size of each image was 512×512. Refer to Figure 6. During the experiment, the secret message was a randomly-generated string composed of 0s and 1s.

In the experiment, we used EC (embedding capacity), bpp (bits per pixel), and PSNR (peak signal-to-noise ratio) to evaluate the proposed scheme. EC means the number of embedded bits in the camouflaged image; bpp represents the average embedded bits per pixel; PSNR was used to evaluate the quality of the camouflaged image. PSNR is defined as Equation (21), where H and W are the image height and width, respectively, and MSE (mean square error) is defined as Equation (22).

$$PSNR = 10\times\log_{10}\left(\frac{255^2}{MSE}\right), \tag{21}$$

$$MSE = \frac{1}{H\times W}\sum_{i=1}^{H}\sum_{j=1}^{W}(\text{CoverImage}_{(i,j)}-\text{StegoImage}_{(i,j)})^2. \tag{22}$$

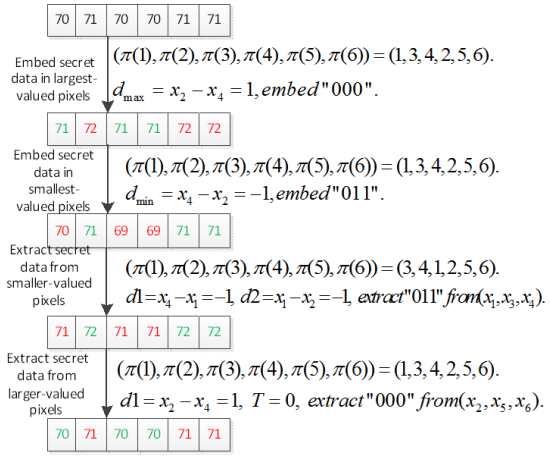


Figure 5: A detailed example of embedding and data extraction procedures

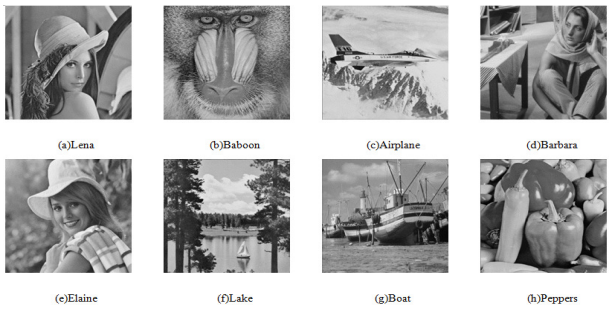


Figure 6: The test cover images

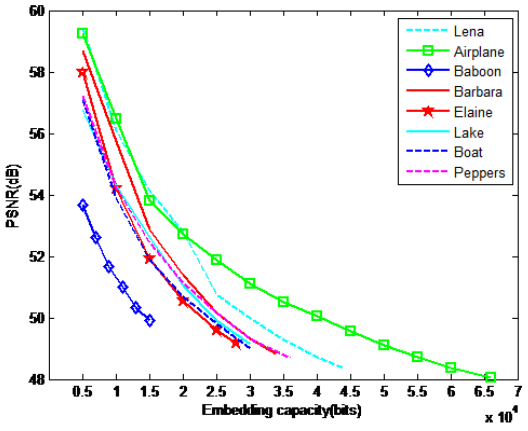


Figure 7: Performance of the proposed scheme

4.2 The Analysis of Proposed Scheme

In this section, we evaluate the proposed scheme by using eight standard grayscale images as the cover image. Figure 7 shows the relationship between embedding capacity and the quality of the camouflaged image. Figure 7 shows that, like all data-hiding algorithms, the quality of the different images gradually declined as the embedding capacity increased. However, comparing the graphs of the different images, it is apparent that the effects caused by increasing the embedding capacity for the different kinds of images are still different. For example, from the test results of images "Baboon" and "Airplane", it is apparent that the degradation of the quality of image "Airplane" is greater than that of image "Baboon" when the embedding capacity was increased from 5000 to 15,000 bits. This occurred because there are more smooth blocks in image "Airplane", and the quality of the image is more sensitive to changes in these blocks. In addition, since we embedded secret data in the same largest and smallest values, the image with more smooth blocks would have the higher embedding capacity. From the experimental results as shown in Figure 7, the maximum embedding capacity of image "Baboon" was the smallest, i.e., just 15,000 bits. Compared to image "Baboon", smoother images, such as images "Lena" and "Elaine" can hide between 20,000 and 30,000 more bits of secret data, and the embedding capacity of smoothest image, i.e., image "Airplane", had 51,000 more bits than that of image "Baboon".

4.3 Comparison of the Performances of PVO-K and GePVO-K

Similar to PVO-K, the size of the block will affect the performance of the proposed scheme. Therefore, we used different block sizes to test the performance of the GePVO-K and PVO-K methods. Taking the local correlation of digital image into account, block sizes that are too large will reduce the maximum embedding capacity drastically. So, five kinds of blocks were tested, the sizes of which were 2×2, 2×3, 3×3, 3×4, and 4×4, respectively. Tables 2-3 show the experimental results of "Lena", "Airplane".

As can be seen from the results in the Tables 2-3, in the process of increasing the block size, the maximum embedding capacity will be reduced because the relevance of the pixel values in the block will be reduced. However, the image quality will be enhanced. In the case of different blocks, GePVO-K greatly improved the maximum embedding capacity. The PVO-K method embeds one bit of secret data to move K bits, so the modification to each pixel is, at most, 1, while the GePVO-K method can embed K bits of data into K largest-valued pixels, and, of course, each pixel value may be expanded by 1 or 2; in addition, the second-largest pixels are modified by one, but when the block is small, the modified pixels may be shifted to the original value in the procedure of embedding secret data in the smallest-valued pixels. As shown in Example 3.4, we embedded six bits of secret data, but

the total modification to all pixels was only 2. Therefore, although GePVO-K inevitably decreases the image quality while improving the embedding capacity, the negative impact is not too high.

Table 2: Performance comparison between PVO-K and proposed GePVO-K (Lena)

Lena	**PVO-K**			**Proposed**			
B size	EC	Payload	PSNR	EC	Payload	PSNR	Gain in EC
2x2	37000	0.14	51.36	44000	0.17	48.37	7000
2x3	28000	0.11	52.27	37300	0.14	49.32	9300
3x3	20000	0.08	54.32	31000	0.12	50.45	11000
3x4	16300	0.06	55.52	25700	0.10	51.41	9400
4x4	12700	0.05	56.87	21200	0.08	52.28	8500

In addition, we can determine from the experiment results, for different types of images, the performance in improving the EC of the proposed scheme was not the same. for the smoother images, i.e., "Lena", the average EC was increased by about 5,800 bits; for the smoothest image, "Airplane," the average EC was increased by about 21,000 bits. These experimental data indicate that the performance of the proposed scheme was better for the smoother images.

Table 3: Performance comparison between PVO-K and proposed GePVO-K (Airplane)

Lena	**PVO-K**			**Proposed**			
B size	EC	Payload	PSNR	EC	Payload	PSNR	Gain in EC
2x2	47000	0.18	51.76	66000	0.25	48.01	19000
2x3	35800	0.14	53.36	58000	0.22	48.49	22200
3x3	24600	0.09	55.41	47900	0.18	49.31	23300
3x4	18900	0.07	56.15	41000	0.16	49.92	22100
4x4	14200	0.05	57.37	33500	0.13	50.71	19300

4.4 Handling Special Blocks

In all PVO-based methods, the overflow/underflow problem must be taken into consideration, because the boundary pixel valued "0" and "255" will be beyond the gray scope after the expansion. The usual practice is to modify the boundary pixel values to a safe range and then construct a map to record the location of the modified pixels. In the proposed scheme, the overflow/underflow may occur in the pixels which are "0," "1," "254," and "255." If these pixels are handled using traditional methods, we must use two bits to record a modified pixel, which leads to a doubling of the size of the location map. In addition, since the block with same pixel values and the normal block cannot be used in the same manner to embed secret data, if data are embedded in the block with the same pixel values, we must establish another location map to record these blocks; however, this treatment often outweigh the benefits. In response to this phenomenon, we proposed an alternative way of handling special blocks; our location map is no longer constructed in a pixel unit, but it records information of each block instead. If the block may overflow/underflow, record it as "2"; if all pixel values in the block are equal, "1" is recorded, the rest blocks are recorded as "0."

Table 4 shows a comparison of the performance of using two overflow/underflow handling methods in the proposed scheme. The first method is to use the traditional way to build a pixel location map, and we used two bits to record the location of a modified pixel and abandoned embedding secret data in the block with the same pixel values; the second is to create a block location map, use two bits to record a special block position, and do nothing with the block that may overflow/underflow. Also, the block with same pixel values is used to embed the secret information.

As can be seen from the data in Table 4, the second method has better performance in both PSNR and maximum embedding capacity. Therefore, it was selected by the proposed scheme. In addition, note that the location map of the selected method is smaller than that of the PVO-K scheme when the block size is greater than 2×2.

Table 4: Performance comparison of two overflow/underflow handling methods

Images	**Use pixel LM**		**Use block LM**	
	EC	**PSNR**	**EC**	**PSNR**
Lena	42600	48.25	44000	48.37
Baboon	14700	49.79	15000	49.90
Airplane	63000	48.14	66000	48.13
Barbara	33000	48.85	34000	48.84
Elaine	26300	49.17	28000	49.17
Lake	29700	49.03	30000	49.14
Boat	29500	48.98	30000	49.00
Peppers	35000	48.58	36000	48.67
Average	34225	48.85	35375	48.90

4.5 Comparison of the Performances of the Proposed Scheme and Several PVO-based Methods

In this section, we evaluate our proposed GePVO-K scheme by comparing it with several PVO-based methods. Because our GePVO-K scheme focuses on increasing the maximum embedding capacity of the cover image, the quality of the camouflaged image may be affected to some degree. Table 5 shows the comparison result of PVO, IPVO, PVO-K, as well as the proposed scheme by test-

ing eight standard gray-scale images; the block size in the experiments was 2×2.

From Table 5, we can conclude that the proposed scheme successfully improved the performance of the three compared methods in embedding capacity. The average embedding capacity of the proposed scheme was 9,625 bits more than PVO, 5,625 bits more than IPVO, and 6,375 bits more than PVO-K. Note that the improvement in the embedding capacity was more obvious for the images with more smooth area. As can be seen from the experimental results of "Airplane," the maximum embedding capacity increased by nearly 20,000 bits compared with PVO-K. With this significant improvement in embedding capacity, the quality of the image will inevitably be affected. But, in the case of substantially increasing the EC, the PSNR remained at a high level in our proposed scheme. In "Airplane," for example, after increasing the embedding capacity by nearly 20,000 bits, the PSNR was still 48.13 dB. Moreover, compared with two most advanced PVO-based RDH methods, PVO-K and PPVO, the proposed scheme also shows excellent performance in terms of embedding capacity under the premise of ensuring the quality of the stego-image. It achieves greatest average EC while maintaining the average PSNR in 48.91dB.

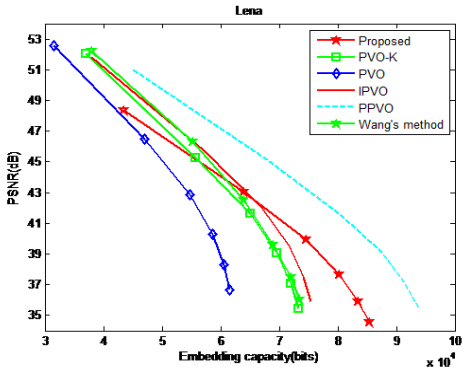
4.6 Multilevel Embedding Analysis

Like majority of the reversible data hiding algorithms, GePVO-K also supports multilevel embedding, that is, treating the camouflaged image as the new cover image to continue embedding secret data. Fig. 8 shows the multilevel embedding performance of PVO, IPVO, PVO-K, and the proposed GePVO-K by testing "Lena" and "Airplane."

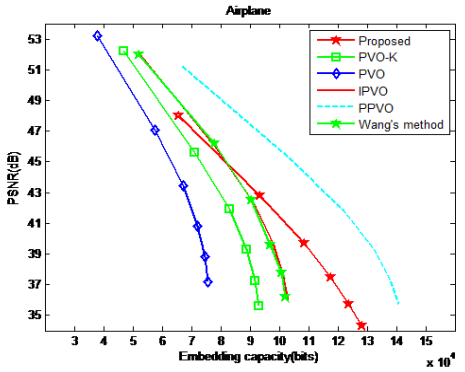
For the proposed scheme, the largest and smallest pixel values in the block were used to embed secret data. Since the secret data were randomly generated, assuming that the numbers of "0" and "1" were basically the same. Then, after embedding the secret data, the previous amount of largest-valued and smallest-valued pixels was halved, which means the embedding capacity of the next level also will be halved, and because the pixels are moved further, the quality of the camouflaged image will decrease step by step. Since the prediction error will be extended by embedding 1, the other three PVO-based methods will have a similar trend.

As can be seen from Fig. 8, the GePVO-K method was better than the embedding capacity of the other three methods (PVO, IPVO, PVO-K) at each embedding level, and the total EC has more obvious differences with the increasing of the embedding level. For image "Lena", the proposed method embedded a total of about 10,000 more bits of secret data than PVO-K and IPVO, and it increased EC by nearly 25,000 bits compared with PVO; for "Airplane," the total EC of our proposed method has 27,000 more bits than IPVO, 37,000 more bits than PVO-K, and nearly 54,000 more bits than PVO. At the same time, we can notice that when using multiple levels to embed the secret data, in the case of same EC (when EC is greater than 65,000 bits for "Lena", 82,000 bits for "Airplane"), the quality of the camouflaged image in the proposed scheme was better and the degree of image distortion did not decline sharply as it did in the other methods.

Since our proposed scheme aims to enhance the performance of the embedding capacity of PVO-K by embedding K bits of secret data into K largest-valued/smallest-valued pixels, the block constrains of PVO-K still not be broken. Due to the reason that the PPVO scheme not only breaks the block constrains but also reuses the pixels to embed secret data, it achieves best multilevel embedding performance. However, compared with another PVO-based scheme, i.e., Wang et al.s method, the proposed scheme has better performance in each embedding level.



(a) Lena



(b) Airplane

Figure 8: Multilevel embedding performance comparison of proposed and other five PVO-base methods

5 Conclusions

In this paper, we presented a new PVO-based RDH method, which is the enhancement on the basis of PVO-K method. It is an extension of PVO-K, and it no longer treats the largest and smallest values in the block as a unit to embed secret data, but it embeds the data in each of

Table 5: Performance comparison between proposed scheme and several PVO-based methods

Images	PVO		IPVO		PVO-K		PPVO		Wang's method		Proposed	
	EC	PSNR	EC	PSNR	EC	PSNR	EC	PSNR	EC	PSNR	EC	PSNR
Lena	32000	52.32	38000	52.35	37000	52.02	44000	51.25	38000	52.15	44000	48.37
Baboon	13000	51.75	13000	51.80	13000	52.00	15000	51.50	13000	51.85	15000	49.90
Airplane	38000	53.12	52000	52.50	47000	52.24	69000	50.95	52000	52.00	66000	48.13
Barbara	27000	52.24	29000	52.05	29000	52.16	33000	51.55	31000	52.45	34000	48.84
Elaine	21000	52.05	24000	52.00	23000	51.87	28000	50.00	25000	51.78	28000	49.17
Lake	23000	52.43	26000	51.79	26000	51.78	29000	52.65	26000	51.85	30000	49.14
Boat	24000	52.00	26000	51.80	26000	51.82	29000	51.20	26000	51.95	30000	49.00
Peppers	28000	52.05	30000	52.00	31000	51.92	33000	51.30	30000	52.00	36000	48.67
Average (payload)	25750 (0.098)	52.25	29750 (0.113)	52.04	29000 (0.111)	51.98	35000 (0.133)	51.30	30125 (0.115)	52.00	35375 (0.135)	48.91

the largest-valued and smallest-valued pixels. Therefore, the maximum embedding capacity has been improved significantly. Also, since it modifies more pixel values, the quality of the image will be subject to a certain decrease, but because of the modified pixel values may be shifted towards the original position in the procedure of embedding secret data in smallest-valued pixels, the quality of the camouflaged image can still be maintained at a high level in the case of maximum embedding capacity. And when embedding secret data with multiple levels, the proposed method achieved better performance in both EC and image quality than the other four PVO-based methods.

The proposed scheme also established a block location map instead of a pixel-based location map, and it processed each block in the different cases. It skipped the blocks that may overflow/underflow, and utilized blocks with the same pixel values, which further increased the embedding capacity; in addition, the quality of the camouflaged image was improved to some extent due to the reduced size of the location map and the smaller modification of the amplitudes of the pixel values.

Acknowledgments

This research work was partially supported by the Ministry of Science and Technology of the Republic of China under the Grant No. MOST105-2221-E-324 -014 and MOST 103-2632-E-324-001-MY3.

References

[1] K. Bharanitharan, C. C. Chang, H. R. Yang, and Z. H. Wang, "Efficient pixel prediction algorithm for reversible data hiding," *International Journal of Network Security*, vol. 18, no. 4, pp. 750–757, 2016.

[2] M. U. Celik, G. Sharma, A. M. Tekalp, *et al.*, "Lossless watermarking for image authentication: a new framework and an implementation," *IEEE Transactions on Image Processing*, vol. 15, no. 4, pp. 1042–1049, 2006.

[3] M. U. Celik, G. Sharma, A. M. Tekalp, and E. Saber, "Lossless generalized-lsb data embedding," *IEEE Transactions on Image Processing*, vol. 14, no. 2, pp. 253–266, 2005.

[4] C. K. Chan and L. M. Cheng, "Hiding data in images by simple lsb substitution," *Pattern Recognition*, vol. 37, no. 3, pp. 469–474, 2004.

[5] J. Fridrich, M. Goljan, and R. Du, "Lossless data embedding: new paradigm in digital watermarking," *EURASIP Journal on Applied Signal Processing*, vol. 2002, no. 1, pp. 185–196, 2002.

[6] Y. Hu, H. K. Lee, and J. Li, "De-based reversible data hiding with improved overflow location map," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 19, no. 2, pp. 250–260, 2009.

[7] B. Jana, D. Giri and S. K. Mondal, "Dual-image based reversible data hiding scheme using pixel value difference expansion," *International Journal of Network Security*, vol. 18, no. 4, pp. 633–643, 2016.

[8] L. Kamstra and H. J. Heijmans, "Reversible data embedding into images using wavelet techniques and sorting," *IEEE Transactions on Image Processing*, vol. 14, no. 12, pp. 2082–2090, 2005.

[9] S. K. Lee, Y. H. Suh, and Y. S. Ho, "Reversiblee image authentication based on watermarking," in *2006 IEEE International Conference on Multimedia and Expo*, pp. 1321–1324, IEEE, 2006.

[10] F. Li, Q. Mao, and C. C. Chang, "A reversible data hiding scheme based on iwt and the sudoku method," *International Journal of Network Security*, vol. 18, no. 3, pp. 410–419, 2016.

[11] X. Li, J. Li, B. Li, and B. Yang, "High-fidelity reversible data hiding scheme based on pixel-value-ordering and prediction-error expansion," *Signal Processing*, vol. 93, no. 1, pp. 198–205, 2013.

[12] M. Liu, H. S. Seah, C. Zhu, W. Lin, and F. Tian, "Reducing location map in prediction-based difference expansion for reversible image data embedding," *Signal Processing*, vol. 92, no. 3, pp. 819–828, 2012.

[13] J. Mielikainen, "Lsb matching revisited," *IEEE Signal Processing Letters*, vol. 13, no. 5, pp. 285–287, 2006.
[14] Z. Ni, Y. Q. Shi, N. Ansari, and W. Su, "Reversible data hiding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 16, no. 3, pp. 354–362, 2006.
[15] B. Ou, X. Li, Y. Zhao, and R. Ni, "Reversible data hiding using invariant pixel-value-ordering and prediction-error expansion," *Signal Processing: Image Communication*, vol. 29, no. 7, pp. 760–772, 2014.
[16] F. Peng, X. Li, and B. Yang, "Adaptive reversible data hiding scheme based on integer transform," *Signal Processing*, vol. 92, no. 1, pp. 54–62, 2012.
[17] F. Peng, X. Li, and B. Yang, "Improved pvo-based reversible data hiding," *Digital Signal Processing*, vol. 25, pp. 255–265, 2014.
[18] X. Qu and H. J. Kim, "Pixel-based pixel value ordering predictor for high-fidelity reversible data hiding," *Signal Processing*, vol. 111, pp. 249–260, 2015.
[19] D. M. Thodi and J. J. Rodríguez, "Expansion embedding techniques for reversible watermarking," *IEEE Transactions on Image Processing*, vol. 16, no. 3, pp. 721–730, 2007.
[20] J. Tian, "Reversible data embedding using a difference expansion," *IEEE Transactions on Circuits Systems for Video Technology*, vol. 13, no. 8, pp. 890–896, 2003.
[21] X. Wang, J. Ding, and Q. Pei, "A novel reversible image data hiding scheme based on pixel value ordering and dynamic pixel block partition," *Information Sciences*, vol. 310, pp. 16–35, 2015.
[22] X. Wang, X. Li, B. Yang, and Z. Guo, "Efficient generalized integer transform for reversible watermarking," *IEEE Signal Processing Letters*, vol. 17, no. 6, pp. 567–570, 2010.
[23] Y. L. Wang, J. J. Shen, M. S. Hwang, "An improved dual image-based reversible hiding technique using LSB matching", *International Journal of Network Security*, vol. 19, no. 5, pp. 858–862, 2017.
[24] Z. H. Wang, X. Zhuang, C. C. Chang, C. Qin, and Y. Zhu, "Reversible data hiding based on geometric structure of pixel groups," *International Journal of Network Security*, vol. 18, no. 1, pp. 52–59, 2016.
[25] S. Zhang, T. Gao, L. Yang, "A reversible data hiding scheme based on histogram modification in integer DWT domain for BTC compressed images," *International Journal of Network Security*, vol. 18, no. 4, pp. 718–727, 2016.
[26] X. Zhang and S. Wang, "Efficient steganographic embedding by exploiting modification direction," *IEEE Communications Letters*, vol. 10, no. 11, pp. 781–783, 2006.

Biography

Jian-Jun Li received the B.Sc. degree in information engineering from Xian University of Electronic Science and Technology, Xian, China, and the M.Sc. and Ph.D. degrees in electrical and computer from The University of Western Ontario and University of Windsor, Canada separately. He is currently working at HangZhou Dianzi University as a chair professor. His research interests include micro-electronics, audio, video and image processing algorithms and implementation.

Yun-He Wu is in his Third year of the master program at Hangzhou Dianzi University, Zhejiang, China, in 2016. His major is Computer Science and Technology. He received his BS degree in Computer Science and Technology in 2014 in Henan University of Science and Technology, Henan, China. His research interests include data hiding, video coding/decoding and image processing.

Chin-Feng Lee received Ph.D. degree in Computer Science and Information Engineering in 1998 from National Chung Cheng University in Taiwan. She is currently a professor in the Department of Information Management at Chaoyang University of Technology, Taiwan. Her research interests include steganography, image processing, and data mining.

Chin-Chen Chang received the B.S. degree in applied mathematics and the M.S. degree in computer and decision sciences from National Tsing Hua University, Hsinchu, Taiwan, R.O.C., in 1977 and 1979, respectively. He received the Ph.D. degree in computer engineering from National Chiao Tung University, Hsinchu, in 1982. From July 1998 to June 2000, he was Director of the Advisory Office, Ministry of Education, R.O.C. From 2002 to 2005, he was a Chair Professor at National Chung Cheng University. From February 2005, he has been a Chair Professor at Feng Chia University. In addition, he was severed as a consultant to several research institutes and government departments. He is currently a Fellow of IEEE and a Fellow of IEE, UK. His current research interests include database design, computer cryptography, image compression, and data structures.

An Efficient Confidentiality Preserving Scheme Using Certificateless Encryption with High Trust Level

Rui Guo[1,2], Huixian Shi[3]
(Corresponding author: Rui Guo)
National Engineering Laboratory for Wireless Security, Xi'an University of Posts and Telecommunications[1]
Xi'an 710121, P.R. China
State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications[2]
Beijing 100876, P. R. China
School of Mathematics and Information Science, Shaanxi Normal University[3]
Xi'an 710062, P. R. China
(Email: guorui@xupt.edu.cn)
(Received Sept. 4, 2016; revised and accepted Jan. 15, 2017)

Abstract

Certificateless public key cryptography with bilinear pairing needs costly operations, which is not appropriate for a practical application. In this paper, we present a certificateless public key encryption scheme without bilinear pairing. This encryption scheme achieves Girault's trust Level 3 as in traditional public key infrastructure based cryptography, and which is proved to be secure in the random oracle model under the hardness assumption of computational Diffie-Hellman problem. Compared with the related schemes, the performance analysis and simulation show that our scheme is more efficient than others. It takes account of security and efficiency.

Keywords: Certificateless Public Key Encryption; Wireless Sensor Networks; IND-CCA Secure; Without Pairing

1 Introduction

In a traditional public key cryptography (PKC), anyone who wants to communicate over a public channel does not share any secret key to each other. A public key infrastructure (PKI) is used to provide an assurance between a public key and the holder of the corresponding private key through the certificates issued by a certification authority (CA). However, a PKI is responsible for managing the certificate, including revocation, storage, distribution and the computational cost of certificate verification, which places a computational burden on the entity.

To simplify the key management and avoid the digital certificates, Shamir [24] proposed the notion of identity based public key cryptography (ID-PKC) which alleviates the existing problems in PKI by getting rid of certificates. In ID-PKC, a user's public key is derived directly from its identity information, such as E-mail address and IP address. The corresponding private key is fully generated by a trusted third party called private key generator (PKG). Nevertheless, there is an obvious drawback that the user must trust the PKG unconditionally (even if a malicious one), which leads to the key escrow problem. Therefore, the hostility PKG could impersonate a user and decrypt the corresponding ciphertext.

In order to resolve the inherent key escrow problem while preserving the advantages of ID-PKC, Al-Riyami and Paterson [2] introduced a new paradigm called certificateless public key cryptography (CL-PKC). Compared with ID-PKC, CL-PKC not only inherits the advantages of it but also resolves the key escrow problem. Specifically, the user combines a secret value picked by itself with the partial private key obtained from the trusted authority (called key generation center, KGC) to generate the full private key. Consequently, the KGC can not obtain the user's private keys to decrypt his ciphertext any more, which has a certain practical application value.

Wireless sensor networks (WSNs) [1] are typically composed of a large number of inexpensive, small and battery-powered sensor nodes [13]. In WSNs, all sensor nodes collaborate together to collect and process certain information such as environment monitoring, health monitoring, military sensing tracking, etc. [9, 14, 17]. In practical application, many situations require the sensor node to be employed in unprotected and even hostile environments, and therefore bring lots of research challenges. One of the important issues is security.

Unfortunately, due to the inherent characteristics of

WSNs mentioned above, adversaries can eavesdrop and distort the transmitted information and disseminate misleading messages into the networks [11]. Hence, encryption mechanisms need to be implemented to protect the message from the malicious attack.

1.1 Related Works

In Al-Riyami and Paterson's work [2], a certificateless public key cryptography was introduced firstly. Based on this work, several certificateless public key encryption (CL-PKE) schemes were proposed in [6, 8, 15, 25, 26, 30]. Cheng and Comley [6] proposed a more efficient encryption scheme on the bases of Al-Riyami and Paterson's work [2]. An interesting feature of this scheme compared with the original CL-PKE is that it can verify the legitimate of the ciphertext in the process of decryption. Libert and Quisquater [15] fixed the model of Al-Riyami and Paterson and gave a method to achieve generic CL-PKE constructions which are provably choose-ciphertext attacks secure (CCA-secure) in the random oracle model. Based on the algebraic properties of Weil pairing, Shi and Li [25] proposed a CL-PKE scheme worked in a kind of parallel model. In [30], Yum and Lee provided a generic secure construction of CL-PKE. In order to resist the strong adversaries in the standard model, Dent et al. [8] presented the first strongly secure CL-PKE scheme in 2008. In 2010, Sun and Li [26] constructed a short-ciphertext CL-PKE scheme in standard model, which achieved adaptive chosen ciphertext security (CCA2-secure). However, the above schemes suffered low efficiency due to the bilinear pairing [5]. Being aware of the above problem of the current constructions of CL-PKE scheme, Baek et al. [3] proposed a CL-PKE scheme without pairing firstly in 2005. In 2011, Lai et al. [12] modified Baek et al.'s scheme to get a Girault's trust Level 3 [10].

Encryption in WSNs is an important mechanism to guarantee the confidence of the transmitted information. Earlier research mainly focused on designing symmetric key based encryption schemes, such as AES-CCM* [31], RC6 [21], Skipjack [19] and so on. There are common issues of key storage and distribution in these schemes, which causes some security threats for WSNs. Considering the security of these symmetric key based encryption schemes, several public key based schemes [4, 7, 20, 29] in WSNs have been proposed to avoid these drawbacks. In 2008, Baek et al. [4] introduced a technique called "indexing", whose performance in WSNs in terms of computation and communication costs is significant superior to normal public key encryption. Watro et al. designed a public key encryption protocol that allows authentication and key agreement between a sensor network and a third party as well as between two sensor networks in [29]. Chu et al. and Oliveria et al. [7, 20] supposed the identity-based encryption scheme for WSNs respectively, which eliminated the problem of certificate management and reduced the costly computation and communication obviously in the public key encryption schemes. Nevertheless, public key based encryption scheme also has a general shortcoming that it is more complicated than symmetric key based one. Therefore, it becomes vitally important to construct the efficient public key based encryption scheme in WSNs for data computation while preserving the level of its security.

1.2 Our Contributions

In this paper, for the reason of applying in WSNs, we propose a CL-PKE scheme that avoid the use of bilinear pairing. Provided that the computational Diffie-Hellman (CDH) problem is intractable, we also prove that the proposal is secure in the random oracle model. Compared with the existing related schemes by simulation, our scheme offers better performances on running time, energy consumption and communication bandwidth. Furthermore, this work promotes the trust level of KGC to the highest Level 3, which strengthen the security of the whole system.

The remainder of this paper is organized as follows. In Section 2, we give some preliminaries such as the definition of the Girault's trust level, the model and the security definitions of CL-PKE and some computational problems. In Section 3, we propose a CL-PKE scheme without bilinear pairing. In Section 4, we analyze the security of our proposal. In Section 5, we make a performance analysis of this scheme. Finally, we conclude the paper in Section 6.

2 Preliminaries

In this section, we review the definition of the Girault's trust level, the model and the security definition of CL-PKE as well as some computational problems which form the basis of the security in our scheme.

2.1 Girault's Trust Level

The Girault's trust level [10] provides the trust hierarchy for public key cryptography, which can be used to evaluate the creditability of the authority.

Level 1. The authority (e.g. the CA in a PKI, the KGC in an identity-based or certificateless cryptography) knows (or can easily compute) users' secret keys. Therefore, the authority can impersonate any user at any time without being detected.

Level 2. The authority does not know (or cannot easily compute) users' secret keys. Nevertheless, it can still impersonate users by generating false guarantees (e.g. false certificates in a PKI, false public keys in a certificateless cryptography).

Level 3. The authority cannot compute users' secret keys, and it can be proven that it generates false guarantees of users if it does so.

According to these definitions, we can easily find that the original certificateless cryptography falls into Level 2, and a traditional PKI achieves Level 3.

2.2 Definition of CL-PKE

The model of CL-PKE in our proposal is similar to that of [3] but with a crucial difference which makes the scheme reach Girault's trust Level 3.

A CL-PKE scheme consists of seven probabilistic, polynomial time (PPT) algorithms: **Setup**, **User-Key-Generation**, **Partial-Key-Extract**, **Set-Private-Key**, **Set-Public-Key**, **Encrypt** and **Decrypt**.

Setup: Taking security parameter k as input, the KGC returns a randomly chosen master secret key msk and a list of public parameters $param$.

User-Key-Generation: Taking a list of public parameters $param$ as input, the user returns a secret key sk and a public key pk.

Partial-Key-Extract: Taking $param, msk$, a user's identity ID and pk received from the user as input, the KGC returns a partial private key D_{ID} and a partial public key P_{ID}.

Set-Private-Key: Taking $param, D_{ID}$ and sk as input, the user returns a private key SK_{ID}.

Set-Public-Key: Taking $param, P_{ID}$ and pk as input, the user returns a public key PK_{ID}.

Encrypt: Taking a plaintext M, a list of parameters $param$, a receiver's identity ID and PK_{ID} as inputs, the sender returns a ciphertext C.

Decrypt: Taking $param$, SK_{ID} and the ciphertext C as input, the receiver runs this deterministic algorithm and returns a decryption δ, which is either a plaintext message or a "*Reject*" message.

The algorithm of **User-Key-Generation** is similar to the algorithm of **Set-Secret-Value** in Baek et al.'s definition [3]. However, the **User-Key-Generation** in this definition must be run prior to the **Partial-Key-Extract** algorithm. According to this operation model, the scheme enjoys the same trust level as the traditional PKI.

2.3 Security Model for CL-PKE

In CL-PKE scheme, as defined in [2], there are two types of adversary with different capabilities. We assume Type I adversary, $\mathcal{A}_I$ acts as a dishonest user who does not have the KGC's master secret key msk but it can replace public keys of arbitrary identities with other public keys of its own choices. While Type II adversary, $\mathcal{A}_{II}$ acts as a malicious KGC who knows the master secret key msk (hence it can compute partial secret key by itself) and is allowed to obtain full secret keys for arbitrary identities but cannot replace any user's public key.

Definition 1. A CL-PKE scheme Π is said to be secure against adaptive chosen ciphertext attack (IND-CCA secure) if no polynomially bounded adversary $\mathcal{A}$ of Type I and Type II has a non-negligible advantage in the following game played against the challenger:

Setup: The challenger $\mathcal{C}$ takes a security parameter k as input and runs the **Setup** algorithm in Section 2.2, then sends the resulting system parameters $param$ to $\mathcal{A}$. If $\mathcal{A}$ is of Type I, $\mathcal{C}$ keeps the master secret key msk to itself. Otherwise (e.g. if $\mathcal{A}$ is of Type II), it gives msk to $\mathcal{A}$.

Phase 1: $\mathcal{A}$ is given access to the following oracles:

1) **Public-Key-Request-Oracle**: Upon receiving a public key query for a user's identity ID, $\mathcal{C}$ computes (sk, pk) and (P_{ID}, D_{ID}), then computes PK_{ID} and returns it to $\mathcal{A}$.
2) **Partial-Key-Extract-Oracle**: Upon receiving a partial key query for a user's identity ID and pk, $\mathcal{C}$ computes (P_{ID}, D_{ID}) and returns them to $\mathcal{A}$. (Note that it is only useful to Type I adversary.)
3) **Private-Key-Request-Oracle**: Upon receiving a private key query for a user's identity ID, $\mathcal{C}$ computes (sk, pk) and (P_{ID}, D_{ID}), then computes SK_{ID} and returns it to $\mathcal{A}$. It outputs $\perp$ (denotes failure) if the user's public key has been replaced (in the case of Type I adversary).
4) **Public-Key-Replace-Oracle**: For identity ID and a valid public key, $\mathcal{A}$ replaces the associated user's public key with the new one of its choice (this is only for Type I adversary). The new value will be recorded and used by $\mathcal{C}$ in the coming computations or responses to the adversary's queries.
5) **Decryption-Oracle**: On input a ciphertext and an identity, return the correct decryption of ciphertext using the private key corresponding to the current value of the public key associated with the identity of user, even if the corresponding public key for the user ID has been replaced.

Challenge Phase: Once $\mathcal{A}$ decides that **Phase 1** is over, it outputs and submits two messages (M_0, M_1), together with a challenge identity ID^* of uncorrupted secret key. Note that $\mathcal{A}$ is not allowed to know the private key of ID^* in anyway. The challenger $\mathcal{C}$ picks a random bit $\beta \in \{0, 1\}$ and computes C^*, which is the encryption of M_β under the current public key PK_{ID^*} for ID^*. If the output of the encryption is $\perp$, $\mathcal{A}$ immediately losses the game. Otherwise, C^* is delivered to $\mathcal{A}$.

Phase 2: Now $\mathcal{A}$ issues the second sequence of queries as in **Phase 1**. A decryption query on the challenge

ciphertext C^* for the combination of ID* and PK_{ID^*} is not allowed.

Guess: Finally, $\mathcal{A}$ outputs its guess β' for β. The adversary wins the game if $\beta' = \beta$ and the advantage of $\mathcal{A}$ in this game is defined to be $Adv(\mathcal{A}) = |\Pr(\beta' = \beta) - \frac{1}{2}|$. The adversary $\mathcal{A}$ breaks an IND-CCA secure CL-PKE scheme Π with $(t, q_{H_i}, q_{par}, q_{pub}, q_{prv}, q_D, \epsilon)$ if and only if the guessing advantage of $\mathcal{A}$ that makes q_{H_i} times to random oracles H_i, q_{par} times **Partial-Key-Extract-Oracle**, q_{pub} times **Public-Key-Request-Oracle**, q_{prv} times **Private-Key-Request-Oracle** and q_D times **Decryption-Oracle** queries is greater than ϵ within running time t. The scheme Π is said to be $(t, q_{H_i}, q_{par}, q_{pub}, q_{prv}, q_D, \epsilon)$- IND-CCA secure if there is no attacker $\mathcal{A}$ that breaks IND-CCA secure scheme Π with $(t, q_{H_i}, q_{par}, q_{pub}, q_{prv}, q_D, \epsilon)$.

2.4 Computational Problems

Now, it will be introduced the Discrete Logarithm (DL) problem and Computational Diffie-Hellman (CDH) problem that are needed in the security analysis of our scheme.

Definition 2 (DL problem). Let G be a cyclic additive group of prime order p and P be a generator of G. Define $Q = xP$ for uniformly chosen $x \in Z_p^*$. Given (P, Q), adversary $\mathcal{A}$ tries to find the value of x.

Definition 3 (CDH problem). Let G be a cyclic additive group of prime order p and P be a generator of G. Define $Q = xP, R = yP$ for uniformly chosen $x, y \in Z_p^*$. Given (P, Q, R), adversary $\mathcal{A}$ tries to find the value of xyP.

Let $\mathcal{A}$ be a CDH adversary. $\mathcal{A}$'s advantage to solve the CDH problem is defined as $Adv(\mathcal{A}) = |\Pr[\mathcal{A}(P, xP, yP) = xyP]|$ and the probability is measured over random choices of $x, y \in Z_p^*$ and the point P. $\mathcal{A}$ solves the CDH problem with (t, ϵ) if and only if the advantage of $\mathcal{A}$ is greater than ϵ within running time t. The CDH problem is said to be (t, ϵ)-intractable if there is no adversary $\mathcal{A}$ that solves the CDH problem with (t, ϵ).

3 Our Construction

In this section, we propose a new CL-PKE scheme without pairing in a cyclic additive group G which performs well.

The notations used throughout this paper are listed in Table 1. Our proposed CL-PKE scheme consists of the following seven PPT algorithms.

Table 1: Notation defined and used in this scheme

ID	the public identity of the entity
$H_i(\cdot)$	the collision-resistant hash function (i=1, 2, 3)
p	the large prime
G	the cyclic additive group
P	the generator of G
x	the master secret key
X	the master public Key
P_{ID}	the entity ID's partial public key
D_{ID}	the entity ID's partial secret key
PK_{ID}	the entity ID's public key
SK_{ID}	the entity ID's secret key
$\|$	the concatenation operation
$\oplus$	the bitwise XOR operation
N	the set of positive integer

Setup: Generate a large prime p, which makes the CDH problem in the cyclic additive group G with generator P of order p be intractable. Pick $x \in Z_p^*$ uniformly at random and compute $X = xP$. Choose hash function $H_1 : \{0,1\}^* \times G^* \times G^* \to Z_p^*$, $H_2 : \{0,1\}^{l_0} \times \{0,1\}^{l_1} \to Z_p^*$, $H_3 : G^* \times G^* \to \{0,1\}^l$, where $l = l_0 + l_1 \in N$. Return $param = \{p, P, G, X, H_1, H_2, H_3\}$ and master secret key $msk = \{x\}$.

User-Key-Generation: Pick $y \in Z_p^*$ at random and compute $Y = yP$. Return $(sk, pk) = (y, Y)$.

Partial-Key-Extract: Pick $s \in Z_p^*$ at random and compute $\omega = sP$ and $d = (s + xH_1(\text{ID}\|\omega\|pk))$, where ID is a user's identity. Return $(P_{ID}, D_{ID}) = (\omega, d)$.

Set-Private-Key: Set $SK_{ID} = (sk, D_{ID}) = (y, d)$. Return SK_{ID}.

Set-Public-Key: Set $PK_{ID} = (pk, P_{ID}) = (Y, \omega)$. Return PK_{ID}.

Encrypt: Let the bit-length of M be l_0. Parse PK_{ID} as (Y, ω), pick $\sigma \in \{0,1\}^{l_1}$ at random and compute $r = H_2(M\|\sigma)$. Compute

$$\begin{aligned} Q_{ID} &= H_1(\text{ID}\|\omega\|pk)X + P_{ID}, \\ C &= (c_1, c_2) = (rQ_{ID}, H_3(z_1\|z_2) \oplus (M||\sigma)), \end{aligned}$$

where $z_1 = rY$, $z_2 = Q_{ID}$ (Note that the bit-length of $(M||\sigma)$ is equal to $l = l_0 + l_1$). Return $C = (c_1, c_2)$.

Decrypt: To decrypt $C = (c_1, c_2)$, compute

$$M||\sigma = H_3(d^{-1}yc_1\|dP) \oplus c_2.$$

If $H_2(M\|\sigma)dP = c_1$, return M. Else, return "*Reject*".

The above decryption algorithm is consistent if and only if (c_1, c_2) is the valid ciphertext of M, then we have:

$$
\begin{aligned}
&H_3(d^{-1}yc_1\|dP) \oplus c_2 \\
&= H_3(d^{-1}yrQ_{ID}\|dP) \oplus c_2 \\
&= H_3(d^{-1}yr(H_1(\mathrm{ID}\|\omega\|pk)X + P_{ID})\|dP) \oplus c_2 \\
&= H_3(d^{-1}yr(H_1(\mathrm{ID}\|\omega\|pk)xP + sP)\|dP) \oplus c_2 \\
&= H_3(d^{-1}yrdP\|Q_{ID}) \oplus c_2 \\
&= H_3((ryP)\|Q_{ID}) \oplus H_3(z_1\|z_2) \oplus (M||\sigma) \\
&= H_3((rY)\|Q_{ID}) \oplus H_3(z_1\|z_2) \oplus (M||\sigma) \\
&= M||\sigma.
\end{aligned}
$$

The intuition behind this construction is as follows. According to the key issuing technique of the Schnorr signature [22], our scheme only requires addition and scalar multiplication instead of transforming the addition group into multiplication group, which avoids the use of the bilinear pairing and raises efficiency of this protocol. Simultaneously, the **User-Key-Generation** algorithm operates prior to the **Partial-Key-Extract** algorithm. In this way, the **Partial-Key-Extract** algorithm includes pk generated by the user as input, and the creditability of the authority achieves to Girault's trust Level 3. Specifically, provided that KGC replaces a user's key pk, there will exist two working keys pk and pk' for this user. Furthermore, two working public keys PK_{ID} and PK'_{ID} binding an identity ID can result from two partial private keys, and only the KGC has the ability to generate these two partial private keys. Therefore, the KGC's forgery is easily detected.

4 Security Analysis

In this subsection, we will show that the scheme described in the previous is secure in the random oracle model.

Theorem 1. *The CL-PKE scheme is IND-CCA secure in the random oracle model, assuming that the CDH problem is intractable.*

Proof. In order to prove this theorem, we prove that our CL-PKE scheme is secure against the Type I and Type II attackers ($\mathcal{A}_I$ and $\mathcal{A}_{II}$) whose behaviors are as described in **Definition 1**.

Assuming there exists an adversary $\mathcal{A}$. Suppose that another PPT $\mathcal{B}$ can make use of $\mathcal{A}$ to solve the CDH problem with probability at least ϵ' and in the time at most t'.

Stage 1: Suppose that $\mathcal{A}$ in this stage is the Type I adversary $\mathcal{A}_I$ and $\mathcal{B}$ is given (p, P, aP, xP) as an instance of the CDH problem. In order to solve the problem by using of $\mathcal{A}_I$, $\mathcal{B}$ needs to simulate a challenger to execute each phase of IND-CCA game for $\mathcal{A}_I$ as follows:

Setup: $\mathcal{B}$ sets $X = xP$, where $x \in Z_p^*$ is the master secret key which is unknown to $\mathcal{B}$, then gives $\mathcal{A}_I$ (p, P, X, H_1, H_2, H_3) as $param$, where H_1, H_2, H_3 are random oracles. Adversary $\mathcal{A}_I$ may make queries of all random oracles at anytime during its attacks as follows:

H_1 **queries:** On receiving a query (ID, ω, Y) to H_1:

1) If $\langle(\mathrm{ID}, \omega, Y), e\rangle$ exists in H_1 **List** L_1, return e as answer.
2) Otherwise, pick $e \in Z_p^*$ at random, add $\langle(\mathrm{ID}, \omega, Y), e\rangle$ to L_1 and return e as answer.

H_2 **queries:** On receiving a query (M, σ) to H_2:

1) If $\langle(M, \sigma), r\rangle$ exists in H_2 **List** L_2, return r as answer.
2) Otherwise, pick $r \in Z_p^*$ at random, add $\langle(M, \sigma), r\rangle$ to L_2 and return r as answer.

H_3 **queries:** On receiving a query (z_1, z_2) to H_3:

1) If $\langle(z_1, z_2), R\rangle$ exists in H_3 **List** L_3, return R as answer.
2) Otherwise, pick $R \in \{0,1\}^l$ at random, add $\langle(z_1, z_2), R\rangle$ to L_3 and return R as answer.

Phase 1: $\mathcal{A}_I$ can issue the following oracle queries.

Partial-Key-Extract: On receiving a query ID:

1) If $\langle \mathrm{ID}, (\omega, d)\rangle$ exist in **PartialKeyList**, return (ω, d) as answer.
2) Otherwise, pick $d, e \in Z_p^*$ at random and compute $\omega = dP - eX$. Add (ID, ω, e) to L_1 (That is, e is defined to be $H_1(\mathrm{ID}\|\omega\|Y)$.) and $\langle \mathrm{ID}, (\omega, d)\rangle$ to **PartialKeyList**, return (ω, d) as answer.

Note that we have $\omega + XH_1(\mathrm{ID}\|\omega\|pk) = dP$ in the above simulation which holds in the real attack too.

Public-Key-Request: On receiving a query ID:

1) If $\langle \mathrm{ID}, (\omega, Y), coin\rangle$ exists in **PublicKeyList**, return $PK_{ID} = (\omega, Y)$ as answer.
2) Otherwise, pick $coin \in \{0,1\}$ at random so that $\Pr[coin = 0] = \delta$. (δ will be determined later.)
3) If $coin = 0$, do the following:
   a. If $\langle \mathrm{ID}, (\omega, d)\rangle$ exists in **PartialKeyList**, pick $y \in Z_p^*$ at random and compute $Y = yP$, add $\langle \mathrm{ID}, (y, d)\rangle$ to **PrivateKeyList** and $\langle \mathrm{ID}, (\omega, Y), coin\rangle$ to **PublicKeyList**, return $PK_{ID} = (\omega, Y)$ as answer.
   b. Otherwise, run the above simulation algorithm for partial key extraction taking ID as input to get partial key (ω, d), pick $y \in Z_p^*$ at random and compute $Y = yP$, add$\langle \mathrm{ID}, (y, d)\rangle$ to **PrivateKeyList** and $\langle \mathrm{ID}, (\omega, Y), coin\rangle$ to **PublicKeyList**, return (ω, Y) as answer.

4) Otherwise (if $coin = 1$), pick $y, s \in Z_p^*$ at random and compute $\omega = sP, Y = yP$, add $\langle \mathrm{ID}, (y, *), s\rangle$ to **PrivateKeyList**, and $\langle \mathrm{ID}, (\omega, Y), coin\rangle$ to **PublicKeyList**, return $PK_{ID} = (\omega, Y)$ as answer.

Private-Key-Request: On receiving a query ID:

1) Run **Public-Key-Request** on ID to get a tuple $\langle \mathrm{ID}, (\omega, Y), coin\rangle \in$ **PublicKeyList**.
2) If $coin = 0$, search **PrivateKeyList** for a tuple $\langle \mathrm{ID}, (y, d)\rangle$ and return $SK_{ID} = (y, d)$ as answer.
3) Otherwise, return "Abort" and terminate.

Decryption: On receiving a query $(\mathrm{ID}, PK_{ID}, C)$, where $C = (c_1, c_2)$ and $PK_{ID} = (\omega, Y)$:

1) Search **PublicKeyList** for a tuple $\langle \mathrm{ID}, (\omega, Y), coin\rangle$.
2) If such a tuple exists and $coin = 0$,
   a. Search **PrivateKeyList** for a tuple $\langle \mathrm{ID}, (y, d)\rangle$. (Note that $\langle \mathrm{ID}, (\omega, Y), coin\rangle$ must exist in **PublicKeyList** and when $coin = 0$, $\langle \mathrm{ID}, (y, d)\rangle$ exist in **PrivateKeyList**.)
   b. Compute $M||\sigma = H_3(d^{-1}yc_1\|dP) \oplus c_2$.
   c. If $c_1 = H_2(M\|\sigma)dP$, return M and "*Reject*" otherwise.
3) Else if such a tuple exists and $coin = 1$,
   a. Run H_1 **queries** to get a tuple $\langle(\mathrm{ID}, \omega, Y), e\rangle$.
   b. If there exist $\langle(M, \sigma), r\rangle \in L_2$ and $\langle(z_1, z_2), R\rangle \in L_3$ such that $c_1 = rQ_{ID}, c_2 = R \oplus (M||\sigma), z_1 = rY, z_2 = Q_{ID}$, return M and "*Reject*" otherwise. The pair $\langle(M, \sigma), r\rangle$ satisfies the above condition uniquely exists in L_2 as the encryption function is injective with respect to $\langle \mathrm{ID}, \omega, Y\rangle$.
4) Else if such a tuple does not exist (This is the case when the public key of a target user is replaced by $\mathcal{A}_I$),
   a. Run H_1 **queries** to get a tuple $\langle(\mathrm{ID}, \omega, Y), e\rangle$.
   b. If there exist $\langle(M, \sigma), r\rangle \in L_2$ and $\langle(z_1, z_2), R\rangle \in L_3$ such that $c_1 = rQ_{ID}, c_2 = R \oplus (M||\sigma), z_1 = rY, z_2 = Q_{ID}$, return M and "*Reject*" otherwise.

Challenge: Once $\mathcal{A}_I$ decides that **Phase 1** is over, then it outputs two messages (M_0, M_1) and a challenge identity ID^*. On receiving a challenge query $\langle \mathrm{ID}^*, (M_0, M_1)\rangle$, $\mathcal{B}$ does the following:

1) Run **Public-Key-Request** on ID^* to get a tuple $\langle \mathrm{ID}^*, (\omega^*, Y^*), coin\ \rangle \in$ **PublicKeyList**.
2) If $coin = 0$, return "Abort" and terminate.
3) Otherwise, do the following:
   a. Search **PrivateKeyList** for a tuple $\langle \mathrm{ID}^*, (y^*, *), s^*\rangle$. (In this case, we know that $Y^* = y^*P, \omega^* = s^*P$.)
   b. Pick $\sigma^* \in \{0,1\}^{l_1}$, $c_2^* \in \{0,1\}^l$ and $\beta \in \{0,1\}$ at random.
   c. Set $c_1^* = aQ_{ID^*}$, $\Omega = a\omega^*$ and $e^* = H_1(\mathrm{ID}^*\|\omega^*\|Y^*)$.
   d. Define $a = H_2(M_\beta\|\sigma^*)$ and $H_3(aY^*\|\omega^* + XH_1(\mathrm{ID}^*\|\omega^*\|Y^*)) = c_2^* \oplus (M_\beta||\sigma^*)$. (Note that $\mathcal{B}$ does not know "a".)
4) Return $C^* = (c_1^*, c_2^*)$ as a target ciphertext.

Phase 2: In this phase, $\mathcal{B}$ answers $\mathcal{A}_I$'s queries in the same way as it have done in **Phase 1**. Note that there is no **Partial-Key-Extract** query or **Private-Key-Request** query on ID^* to be issued. Also, no **Decryption** query should be made on C^* for the combination of ID^* and PK_{ID^*} that encrypted plaintext M_β.

Guess: $\mathcal{A}_I$ outputs a guess β'. Now $\mathcal{B}$ returns the set

$S = \{\frac{1}{e^*}(az_{2i} - \Omega) | z_{2i}$ is the second component of queries to H_3 for $i \in [1, q_{H_3}]$ such that $e^* = H_1(\mathrm{ID}^*\|\omega^*\|Y^*)\ \}$.

Then, $\mathcal{B}$ will be able to solve the CDH problem by picking $\frac{1}{e^*}(az_{2i} - \Omega)$ from S.

Analysis: From the construction of H_1, it is clear that the simulation of H_1 is perfect. As long as $\mathcal{A}_I$ does not query (M_β, σ^*) to H_2 nor aY^* and $\omega^* + Xe^*$ to H_3 where $e^* = H_1(\mathrm{ID}^*\|\omega^*\|Y^*)$, the simulations of H_2 and H_3 are perfect. By $\mathsf{Ask}H_3^*$ we denote the event that $(aY^*, \omega^* + Xe^*)$ has been queried to H_3. Also, by $\mathsf{Ask}H_2^*$ we denote the event that (M_β, σ^*) has been queried to H_2. If it happens, $\mathcal{B}$ will be able to solve the CDH problem by choosing a tuple $\langle(z_1, z_2), R\rangle$ from L_3 and computing $\frac{1}{e^*}(az_{2i} - \Omega)$ with the probability at least $\frac{1}{q_{H_3}}$. Hence, we have $\epsilon' \geq \frac{1}{q_{H_3}}\Pr[\mathsf{Ask}H_3^*]$.

It is easy to notice that if $\mathcal{B}$ does not abort, the simulations of **Partial-Key-Extract**, **Public-Key-Request**, **Private-Key-Request** and the simulated target ciphertext is identically distributed as the real one from the construction.

Now, we evaluate the simulation of the decryption oracle. If a public key PK_{ID} has not been replaced or PK_{ID} has not been produce under $coin = 1$, the simulation is perfect as $\mathcal{B}$ knows the private key SK_{ID} corresponding to PK_{ID}. Otherwise, simulation error may occur if $\mathcal{B}$ runs the decryption oracle simulation specified above. Let DecErr be this event. Suppose ID, PK_{ID} and C, where $C = (c_1, c_2)$ and $PK_{ID} = (\omega, Y)$, have been issued as a valid decryption query. Even if C is valid, there is a possibility that C can be produced without querying (rY, Q_{ID}) to H_3, where $r = H_2(M\|\sigma)$. Let Valid be an event that

C is valid, $\mathsf{Ask}H_3$ and $\mathsf{Ask}H_2$ respectively be events that $(aY, \omega + Xe)$ has been queried to H_3 and (M, σ) has been queried to H_2 with $C = (c_1, c_2) = (rQ_{ID}, H_3(rY \| Q_{ID}) \oplus (M||\sigma))$ and $PK_{ID} = (\omega, Y)$, where $r = H_2(M\|\sigma)$. Since DecErr is an event that $\mathsf{Valid}|\neg\mathsf{Ask}H_3$ happens during the entire simulation and q_D decryption oracle queries are made, we have $\Pr[\mathsf{DecErr}]=q_D\Pr[\mathsf{Valid}|\neg\mathsf{Ask}H_3]$. However,

$$\begin{aligned}\Pr[\mathsf{Valid}|\neg\mathsf{Ask}H_3] &\leq \Pr[\mathsf{Valid} \wedge \mathsf{Ask}H_2|\neg\mathsf{Ask}H_3] \\ &\quad + \Pr[\mathsf{Valid} \wedge \neg\mathsf{Ask}H_2|\neg\mathsf{Ask}H_3] \\ &\leq \Pr[\mathsf{Ask}H_2|\neg\mathsf{Ask}H_3] \\ &\quad + \Pr[\mathsf{Valid}|\neg\mathsf{Ask}H_2 \wedge \neg\mathsf{Ask}H_3] \\ &\leq \frac{q_{H_2}}{2^{l_1}} + \frac{1}{p}\end{aligned}$$

The event $(\mathsf{Ask}H_3^* \vee (\mathsf{Ask}H_2^*|\neg\mathsf{Ask}H_3^*) \vee \mathsf{DecErr})|\neg\mathsf{Abort}$ be denoted by E, where Abort denotes an event that $\mathcal{B}$ aborts during the simulation. The probability $\neg\mathsf{Abort}$ that happens is given by $\delta^{q_{prv}}(1-\delta)$ which is maximized at $\delta = 1 - \frac{1}{q_{prv}+1}$. Hence, we have $\Pr[\neg\mathsf{Abort}] \leq \frac{1}{e(q_{prv}+1)}$, where e denotes the base of the natural logarithm.

If E does not happen, it is clear that $\mathcal{A}_I$ does not gain any advantage greater than $\frac{1}{2}$ to guess β due to the randomness of the output of the random oracle H_3. Namely, we have $\Pr[\beta' = \beta|\neg\mathsf{E}] \leq \frac{1}{2}$.

By definition of ϵ, we have

$$\begin{aligned}\epsilon &< |\Pr[\beta' = \beta] - \frac{1}{2}| \\ &= |\Pr[\beta' = \beta|\neg\mathsf{E}]\Pr[\neg\mathsf{E}] + \Pr[\beta' = \beta|\mathsf{E}]\Pr[\mathsf{E}] - \frac{1}{2}| \\ &\leq |\frac{1}{2}\Pr[\neg\mathsf{E}] + \Pr[\mathsf{E}] - \frac{1}{2}| = |\frac{1}{2}(1 - \Pr[\mathsf{E}]) + \Pr[\mathsf{E}] - \frac{1}{2}| \\ &= \frac{1}{2}\Pr[\mathsf{E}] \\ &\leq \frac{1}{2\Pr[\neg\mathsf{Abort}]}(\Pr[\mathsf{Ask}H_3^*] + \Pr[\mathsf{Ask}H_2^*|\neg\mathsf{Ask}H_3^*] + \Pr[\mathsf{DecErr}]) \\ &\leq \frac{e(q_{prv}+1)}{2}(q_{H_3}\epsilon' + \frac{q_{H_2}}{2^{l_1}} + \frac{q_D q_{H_2}}{2^{l_1}} + \frac{q_D}{p}).\end{aligned}$$

Consequently, we obtain $\epsilon' > \frac{1}{q_{H_3}}(\frac{2\epsilon}{e(q_{prv}+1)} - \frac{q_{H_2}}{2^{l_1}} - \frac{q_D q_{H_2}}{2^{l_1}} - \frac{q_D}{p})$. The running time of the CDH adversary $\mathcal{B}$ is $t' > t + 2(q_{pub} + q_{prv})t_{sm} + q_{par}t_{sm} + 2q_D q_{H_2} q_{H_3} t_{sm} + 3t_{sm}$, where t_{sm} denotes the time for computing scalar multiplication on the cyclic addition group G.

Stage 2: Suppose that $\mathcal{A}$ in this stage is the Type II adversary $\mathcal{A}_{II}$ and $\mathcal{B}$ is given (p, P, aP, bP) as an instance of the CDH problem. In order to solve the problem by using of $\mathcal{A}_{II}$, $\mathcal{B}$ needs to simulate a challenger to execute each phase of IND-CCA game for $\mathcal{A}_{II}$ as follows:

Setup: $\mathcal{B}$ picks $x \in Z_p^*$ at randomly and computes $X = xP$, where x is the master key, then gives $\mathcal{A}_{II}$ (p, P, X, H_1, H_2, H_3) as *param*, where H_1, H_2, H_3 are random oracles. Adversary $\mathcal{A}_{II}$ may make queries of all random oracles at anytime during its attacks as **Stage 1**.

Phase 1: $\mathcal{A}_{II}$ can issue the following oracle queries.

Public-Key-Request: On receiving a query ID:

1) If $\langle \mathrm{ID}, (\omega, Y), coin\rangle$ exists in **PublicKeyList**, return $PK_{ID} = (\omega, Y)$ as answer.
2) Otherwise, pick $coin \in \{0, 1\}$ at random, so that $\Pr[coin = 0] = \delta$. (δ is the same as it in the proof of **Stage 1**.)
3) If $coin = 0$, pick $y, s \in Z_p^*$ at random and compute $Y = yP$, $\omega = sP$ and $d = s + xH_1(\mathrm{ID}\|\omega\|pk)$, add $\langle \mathrm{ID}, (y, d)\rangle$ to **PrivateKeyList** and $\langle \mathrm{ID}, (\omega, Y), coin\rangle$ to **PublicKeyList**, return $PK_{ID} = (\omega, Y)$ as answer.
4) Otherwise (if $coin = 1$), pick $y, s \in Z_p^*$ at random and compute $Y = yP$, $\omega = s(bP)$, add $\langle \mathrm{ID}, (y, *)\rangle$ to **PrivateKeyList** and $\langle \mathrm{ID}, (\omega, Y), coin\rangle$ to **PublicKeyList**, return $PK_{ID} = (\omega, Y)$ as answer.

Private-Key-Request: On receiving a query ID:

1) Run **Public-Key-Request** on ID to get a tuple $\langle \mathrm{ID}, (\omega, Y), coin\rangle \in$ **PublicKeyList**.
2) If $coin = 0$, search **PrivateKeyList** for a tuple $\langle \mathrm{ID}, (y, d)\rangle$ and return $SK_{ID} = (y, d)$ as answer.
3) Otherwise, return "Abort" and terminate.

Decryption: On receiving a query $(\mathrm{ID}, PK_{ID}, C)$, where $C = (c_1, c_2)$ and $PK_{ID} = (\omega, Y)$:

1) Search **PublicKeyList** for a tuple $\langle \mathrm{ID}, (\omega, Y), coin\rangle$. If $coin = 0$, search **PrivateKeyList** for a tuple $\langle \mathrm{ID}, (y, d)\rangle$ (Note that $\langle \mathrm{ID}, (\omega, Y), coin\rangle$ must exist in **PublicKeyList** and when $coin = 0$, $\langle \mathrm{ID}, (y, d)\rangle$ exist in **PrivateKeyList**). Then, set $SK_{ID} = (y, d)$ and run **Decrypt**. Finally, return the results of **Decrypt** algorithm.
2) Otherwise (if $coin = 1$), run H_1 **queries** to get a tuple $\langle (\mathrm{ID}, \omega, Y), e\rangle$. If there exist $\langle (M, \sigma), r\rangle \in L_2$ and $\langle (z_1, z_2), R\rangle \in L_3$ such that $c_1 = rQ_{ID}, c_2 = R \oplus (M||\sigma)$, $z_1 = rY$, $z_2 = Q_{ID}$, return M and "*Reject*" otherwise.

Challenge: $\mathcal{A}_{II}$ then outputs two messages (M_0, M_1) and a challenge identity ID^*. On receiving a challenge query $\langle \mathrm{ID}^*, (M_0, M_1)\rangle$:

1) $\mathcal{B}$ runs **Public-Key-Request** taking ID^* as input to get a tuple $\langle \mathrm{ID}^*, (\omega^*, Y^*), coin\rangle \in$ **PublicKeyList**.
2) If $coin = 0$, return "Abort" and terminate.
3) Otherwise, do the following,
   a. Search **PrivateKeyList** for a tuple $\langle \mathrm{ID}^*, (y^*, *), s^*\rangle$.(In this case, we know that $Y^* = y^*P, \omega^* = s^*bP$.)
   b. Pick $\sigma^* \in \{0, 1\}^{l_1}$, $c_2^* \in \{0, 1\}^l$ and $\beta \in \{0, 1\}$ at random.

Table 2: Comparison of the related schemes

Schemes	Enc	Dec	Sec-Lev	Sec-Ass
[2]	3P+1S+1E	1P+1S	IND-CCA	GBDHP
[6]	1P+2S+1E	1P+2S	IND-CCA	BDH
[25]	3S+1E	1P+3S	IND-CCA	K-BDHI
[8]	1P+3S+1E	4P	Strong Type I/II	3-DDH
[26]	2P+2S+2E	2P+1S	IND-CCAII	BDH
[3]	4E	3E	IND-CCA	CDH
[12]	2S+3E	2E	IND-CCA	CDH
[27]	6E	3E	IND-CCA	CDH
Ours	4S	3S	IND-CCA	CDH

c. Set $c_1^* = aQ_{ID^*}$, $e^* = H_1(\mathrm{ID}^*\|\omega^*\|Y^*)$.

d. Define $a = H_2(M_\beta\|\sigma^*)$ and $H_3(aY^*\|\omega^* + XH_1(\mathrm{ID}^*\|\omega^*\|Y^*)) = c_2^* \oplus (M_\beta||\sigma^*)$. (Note that $\mathcal{B}$ does not know "a".)

4) Return $C^* = (c_1^*, c_2^*)$ as a target ciphertext.

Phase 2: $\mathcal{B}$ repeats the same algorithms that it operated in **Phase 1** of **Stage 2**.

Guess: $\mathcal{A}_{II}$ outputs a guess β'. Now $\mathcal{B}$ returns the set

$S = \{\frac{1}{s^*}(az_{2i} - (aP)xe^*)|z_{2i}$ is the second component of queries to H_3 for $i \in [1, q_{H_3}]$ such that $e^* = H_1(\mathrm{ID}^*\|\omega^*\|Y^*)\ \}$.

Then, $\mathcal{B}$ will be able to solve the CDH problem by picking $\frac{1}{s^*}(az_{2i} - (aP)xe^*)$ from S.

Analysis: Similar to **Analysis** in the proof of **Stage 1**.

Consequently, we obtain $\epsilon' > \frac{1}{q_{H_3}}(\frac{2\epsilon}{e(q_{prv}+1)} - \frac{q_{H_2}}{2^{l_1}} - \frac{q_D q_{H_2}}{2^{l_1}} - \frac{q_D}{p})$. The running time of the CDH adversary $\mathcal{B}$ is $t' > t + 2(q_{pub} + q_{prv})t_{sm} + 2q_D q_{H_2} q_{H_3} t_{sm} + 3t_{sm}$.

To sum up the two stages above, we complete the proof of **Theorem 1**. □

5 Performance Analysis

In this subsection, we compare the proposed scheme with other existing CL-PKE schemes on the computation complexity of encryption(**Enc**), decryption(**Dec**), security level(**Sec-Lev**) and security assumption(**Sec-Ass**). Without considering the additional of two points and hash function in the cyclic additive group, each scheme has three major types of operation, i.e., Pairing(P), Scalar Multiplication(S) and Exponentiation(E). From Table 2, we can see that our scheme calculates four scalar multiplications in **Encrypt** and three scalar multiplications in **Decrypt**, which denotes that it needs a lower computation cost than others.

We simulate the cryptographic operations by using of MIRACL (version 5.6.1, [23]) on a laptop using the Intel Core i5-2400 at a frequency of 3.10 GHz with 3GB memory and a Windows XP operation system, and then obtain the average running time in Table 3. For pairing-based schemes, considering to be implemented in practice efficiently, we use the Fast-Tate-Pairing in MIRACL, which is defined over the MNT curve E/F_q [18] with embedding degree 4 and q is a 160 bits prime, and its security level achieves the difficulty of discrete log problem in 640 bits. For ECC-based scheme, we employed the parameter secp192k1 [28], where $p = 2^{192} - 2^{32} - 2^{12} - 2^8 - 2^7 - 2^6 - 2^3 - 1$. Moreover, the length of an element in multiplication group is set to be 1024 bits.

Table 3: Cryptographic operations time (in millisecond)

Fast-Tate -Pairing	Exponentiation	Scalar Multiplication
2.65	3.91	0.78

Based on these settings above, we can simulate the total running time of one round of **Encrypt-Decrypt** in different schemes respectively as shown in Table 4. As for energy consumption, it can be calculated as $W = U \times I \times t$ based on the execution time (t), the voltage (U) and current draw (I). Suppose that the voltage is 3v and the current draw is 8mA in sensor platform MICAz [16], the energy consumption of one round in every protocol is also demonstrated in Table 4. For instance, in Al-Riyami and Paterson's work [2], it needs 4 pairing operations, 2 scalar multiplications and 1 exponentiation altogether, then the total running time in MICAz is $4 \times 2.65 + 2 \times 0.78 + 1 \times 3.91 = 16.07$ ms, and the energy consumption is $3 \times 8 \times 16.07 = 385.68$ μJ.

The communication cost of different schemes are compared in terms of bandwidth of transmitted ciphertext. Assuming the output of one way Hash function is 192 bits and the symmetric cipher is 128 bits (such as AES). In our protocol and [25], each ciphertext contains one point and one Hash value, thus the bandwidths of our protocol and [25] are (192+192)/8=48 bytes respectively. In [2] and [6], each ciphertext contains one point and two Hash values, thus the bandwidths of [2] and [6] are (192+192×2)/8=72 bytes respectively. In [8], the ciphertext contains one pairing and three modular exponentia-

tions, thus the bandwidth of [8] is (160+1024×3)/8=404 bytes. In [26], the ciphertext contains one point and one symmetric cipher, thus the bandwidth of [26] is (192+128)/8=40 bytes. In [3], [12] and [27], each ciphertext contains one modular exponentiations and one Hash value, thus the bandwidths of [3], [12] and [27] are (1024+192)/8=152 bytes respectively. The detailed comparison results are listed in Table 4.

Table 4: Comparison of the CL-PKE Schemes

Schemes	Running Time (ms)	Energy Consumption (μJ)	Bandwidth (byte)
[2]	16.07	385.68	72
[6]	12.33	295.92	72
[25]	11.24	269.76	48
[8]	19.5	468	404
[26]	20.76	498.24	40
[3]	27.37	656.88	152
[12]	21.11	506.64	152
[27]	35.19	844.56	152
Ours	5.46	131.04	48

To sum up, our protocol is more suitable to be applied in WSNs with the characteristics of low cost and low power sensor nodes that are small in size and communicate wirelessly with each other nodes in short distances.

6 Conclusions

In this paper, we propose a CL-PKE scheme that does not depend on the pairing and prove that the scheme is IND-CCA secure in the random oracle model, relative to the hardness of the CDH problem. Besides, this scheme can achieve the highest trust Level 3. The comparison and simulation in Section 5 illustrate that our proposed scheme is advantageous over the related schemes on computation cost, communication overhead and energy consumption. Due to the appealing properties, the proposal could be applied in WSNs.

Acknowledgments

This study was supported by the National Nature Science Foundation of China under grant NSFC 11501343, and the Open Foundation of State key Laboratory of Networking and Switching Technology (Beijing University of Posts and Telecommunications) under grant SKLNST-2016-2-11. The authors gratefully acknowledge the anonymous reviewers for their valuable comments.

References

[1] I. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci, "A survey on sensor networks," *IEEE Communications Magazine*, vol. 40, pp. 102–114, 2002.

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

[3] J. Baek, R. Safavi-Naini and W. Susilo, "Certificateless public key encryption without pairing," in *8th International Conference on Information Security*, LNCS 3650, pp. 134–148, Springer, 2005.

[4] J. Baek, H. C. Tan, J. Zhou and J. W. Wong, "Realizing stateful public key encryption in wireless sensor network," in *The IFIP TC-11 23rd International Information Security Conference*, vol. 278, pp. 95–107, 2008.

[5] P. S. L. M. Barreto, H. Y. Kim, B. Lynn and M. Scott, "Efficient algorithms for pairing-based cryptosystems," in *Advances in Cryptology (CRYPTO'02)*, LNCS 2442, pp. 354–368, Springer, 2002.

[6] Z. H. Cheng and R. Comley, "Efficient certificateless public key encryption," *IACR Cryptology ePrint Archive*, 2005. (`http://eprint.iacr.org/2005/012.pdf`)

[7] C. K. Chu, J. K. Liu, J. Zhou, F. Bao and R. H. Deng, "Practical ID-based encryption for wireless sensor network," in *Proceedings of the 5th ACM Symposium on Information, Computer and Communications Security*, pp. 337–340, 2010.

[8] A. W. Dent, B. Libert and K. G. Paterson, "Certificateless encryption schemes strongly secure in the standard model," in *11th International Conference on Theory and Practice in Public-Key Cryptography*, LNCS 4939, pp. 344–359, Springer, 2008.

[9] T. H. Feng, W. Li and M. S. Hwang, "A false data report filtering scheme in wireless sensor networks: A survey," *International Journal of Network Security*, vol. 17, pp. 229–236, 2015.

[10] M. Girault, "Self-certified public keys," in *Advances in Cryptology (EUROCRYPTO'91)*, vol. 547, pp. 490–497, Springer, 1992.

[11] M. Kumar, K. Dutta, I. Chopra, "Impact of wormhole attack on data aggregation in hieracrchical WSN," *International Journal of Electronics and Information Engineering*, vol. 1, no. 2, pp. 70–77, 2014.

[12] J. Lai, W. Kou and K. Chen, "Self-generated-certificate public key encryption without pairing and its application," *Information Sciences*, vol. 181, pp. 2422–2435, 2011.

[13] C. T. Li, 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, Dec. 2011.

[14] G. Li, Q. Jiang, Y. Shi and F. Wei, "Anonymous network information acquirement protocol for mobile users in heterogeneous wireless networks," *In-*

ternational Journal of Network Security, vol. 18, pp. 193–200, 2016.

[15] B. Libert and J. Quisquater, "On constructing certificateless cryptosystems from identity based encryption," in *9th International Conference on Theory and Practice in Public-Key Cryptography*, LNCS 3958, pp. 474–490, Springer, 2006.

[16] A. Liu and P. Ning, "TinyECC: A configurable libraty for elliptic curve cryptography in wireless sensor networks," in *2008 International Conference on Information Processing in Sensor Networks*, pp. 245–256, 2008.

[17] Y. Lu, X. Yang and X. Wu, "A secure anonymous authentication scheme for wireless communications using smart cards," *International Journal of Network Security*, vol. 17, pp. 237–245, 2015.

[18] A. Miyaji, M. Nakavayashi and S. Takano, "New explicit conditions of elliptic curve traces for FR-reduction," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E84-A, no. 5, pp. 1234–1243, 2001.

[19] NIST, *SkipJack and KEA Algorithm Specifications*, National Institute of Standards and Technology, May 29, 1998. (`http://csrc.nist.gov/groups/ST/toolkit/documents/skipjack/skipjack.pdf`)

[20] L. B. Oliveira, R. Dahab, J. López, F. Daguana and A. A. F. Loureiro "Identity-based encryption for sensor networks," in *The Fifth Annual IEEE International Conference on Pervasive Computing and Communications Workshops*, pp. 290–294, 2007.

[21] R. L. Rivest, M. J. B. Robshaw, R. Sidney and Y. L. Yin, *The RC6 Block Cipher v1.1*, Aug. 20, 1998. (`https://people.csail.mit.edu/rivest/pubs/RRSY98.pdf`)

[22] C. P. Schnorr, "Efficient identifications and signature for smart cards," in *Advances in Cryptology (CRYPTO'89)*, LNCS 435, pp. 239–251, Springer, 1990.

[23] M. Scott, *Miracl Library*, Sept. 4, 2017. (`http://certivox.com/`)

[24] A. Shamir, "Identity-based cryptosystems and signature schemes," in *Advances in Cryptology (CRYPTO'84)*, LNCS 196, pp. 47–53, Springer, 1985.

[25] Y. Shi and J. Li, "Provable efficient certificateless public key encryption," *IACR Cryptology ePrint Archive*, 2005. (`http://eprint.iacr.org/2005/287.pdf`)

[26] Y. Sun and H. Li, "Short-ciphertext and BDH-based CCA2 secure certificateless encryption," *Science China: Information Sciences*, vol. 53, pp. 2005–2015, 2010.

[27] Y. Sun, F. Zhang and J. Baek, "Strongly secure certificateless public key encryption without pairing," in *6th International Conference on Cryptology and Network Security*, vol. 4856, pp. 194–208, 2007.

[28] The Certicom Research, *SEC2: Recommended Elliptic Curve Domain Parameters*, Version 2.0, Jan. 27, 2010. (`http://www.secg.org/sec2-v2.pdf`)

[29] R. J. Watro, D. Kong, S. Cuti, C. Gardiner, C. Lynn and P. Kruus, "TinyPK: Securing sensor networks with public key technology," in *2nd ACM Workshop on Security of Ad Hoc and Sensor Networks*, pp. 59–64, 2004.

[30] D. H. Yum and P. J. Lee, "Generic construction of certificateless encryption," *International Conference on Computational Science and Its Applications (ICCSA'04)*, LNCS 3040, pp. 802–811, Springer, 2004.

[31] ZigBee Alliance, *ZigBee Specifications*, Document 053474r17, 2008.

Biography

Rui Guo received the Ph.D degrees in the Department of State Key Laboratory of Networking and Switch Technology, Beijing University of Posts and Telecommunications in 2014. Now, he is a lecturer in National Engineering Laboratory for Wireless Security, Xi'an University of Posts and Telecommunications. His present research interests include cryptography, information security and WSN.

Huixian Shi received the B.S. and Ph.D degrees in Department of Mathematics and Information Science from Shaanxi Normal University, Xi'an, China, in 2007 and 2013, respectively. Now she is a associate professor in Department of Department of Mathematics and Information Science in Shaanxi normal University. Her present research interests include model checking, fuzzy logic and uncertainty reasoning.

Provably Secure Quantum Key Distribution By Applying Quantum Gate

V. Padmavathi[1], B. Vishnu Vardhan[2], and A. V. N. Krishna[3]
(Corresponding author: V. Padmavathi)
Associate Professor, Department of Computer Science & Engineering, Sreenidhi Institute of Science & Technology[1]
Hyderabad, Telangana, India
Professor, Department of Computer Science & Engineering, JNTUH College of Engineering[2]
Karimnagar, Telangana, India
Professor, Department of Computer Science & Engineering, Christ University[3]
Bengaluru, Karnataka, India
(Email: chpadmareddy1@gmail.com)
(Received Nov. 26, 2016; revised and accepted Feb. 21 & Mar. 31, 2017)

Abstract

The need for Quantum Key Distribution (QKD) is strengthening due to its inalienable principles of quantum mechanics. QKD commences when sender transforms bits into qubits or quantum states by applying photon polarization and sends to the receiver. The qubits are altered when measured in incorrect polarization and cannot be reproduced according to quantum mechanics principles. BB84 protocol is the primary QKD protocol announced in 1984. This paper introduces a new regime of secure QKD using Hadamard quantum gate named as PVK16 QKD protocol. Applying quantum gate to QKD makes tangle to the eavesdroppers to measure the qubits. For a given length of key, it is shown that the error rate is negligible. Also, the authentication procedure using digital certificates prior to QKD is being performed which confers assurance that the communicating entities are legitimate users. It is used as a defensive mechanism on man in the middle attack.

Keywords: Authentication; Quantum Cryptography; Hadamard Gate; QKD; Qubits

1 Introduction

In 1969, Stephen Wiesner identified that quantum mechanics has the prospective to play a vital role in the field of cryptography [26]. This has become a breakthrough idea to recommend a new research field known as Quantum cryptography. It is a promising research area where the secure communications happens by means of principles of quantum mechanics and quantum computing. The central unit of information in quantum computing is quantum bit in short known as qubit. The idea of Quantum cryptography was put forward to project a process for QKD by Wiedemann [25]. In 1984, Bennet and Brassard proposed first QKD and popularly known as BB84 protocol [2]. It requires two communication channels specifically a classical channel and a quantum channel. The classical information is communicated by classical channel and qubits by quantum channel [6].

The two key principles of quantum mechanics are: 1) the principle of Heisenberg Uncertainty; 2) the principle of photon polarization. The Heisenberg Uncertainty principle states that the simultaneous measurement of physical properties which are related, cannot be made [2]. The principle of photon polarization states that the qubits cannot be replicated according to the theorem of no-cloning [27]. Hence, the accomplishment of principles of quantum mechanics is one of the reasons why quantum cryptography is so powerful.

In this paper, a new paradigm of quantum key distribution scheme using quantum gates named as PVK16 QKD protocol with authentication of two communicating entities prior to key distribution is proposed. Authentication implies secure communication. A noteworthy characteristic of any protocol is that the information is authenticate, if authenticate entities participate in the communication. Without authentication, there is always scope of an eavesdropping attack and the entities are fooled by impersonation.

The quantum gate applied on one-qubit can be illustrated by 2 by 2 matrix [19]. Hence, the time complexity of quantum gates is 2^n which is exponential. It is proved that the exponential complexities are secure. By embodying quantum gate into QKD protocol it becomes cumbersome for an eavesdropper to measure the qubits. Besides QKD, this scheme has also adopted the process of authentication to detect man in the middle attack.

Section 2 will brief about the related work in QKD and

authentication. The implementation of proposed protocol is elucidated in section 3. Section 4 explicates the analysis of results. Section 5 highlights on security analysis. Section 6 presents conclusions.

2 Related Work

The last two decades has marked the evolution of Quantum cryptography in productive ways. With various theoretical proposals and demonstrations using experiments, the Quantum cryptography was applied in distributing secret key. Since then, QKD has been used as a solution to the problem of key distribution. A quantum channel was developed using an optical fiber in order to implement QKD in a secure fashion for over different distances like 14 kms, 24 kms [3, 7, 18]. In 2015, a highly secured network switches with QKD systems was developed [5].

In recent years, Quantum cryptography is attracting substantial attention among researchers due to major theoretical and practical research projects related to the implementation of QKD. The QKD was successfully demonstrated to resist eavesdropping attacks, photon number splitting attacks, etc. using diverse protocols like Ekert's entanglement, weak coherent pulses, decoy states and optical fiber [8, 13, 21, 22, 29, 30].

The process of authentication was also carried out to ensure that the communicating entities are alleged users from past years. It was accomplished using several concepts of quantum mechanics namely entanglement, one qubit, Bell states and unitary transformations [1, 14, 28]. Also the authentication process was performed by means of classical XOR operation [4]. In 2009, by using quantum superposition states the authentication protocol was proposed [9]. The public key authentication scheme with trusted server based on discrete logarithms for cryptosystems was also implemented [10, 11].

3 Implementation

PVK16 QKD protocol is introduced to fulfill quantum key distribution using Hadamard gate. It is implemented using MatLab R2014a [15]. The authentication procedure is carried out through digital certificate using OpenSSL tool [20] to detect man in the middle attack.

3.1 Authentication

Authentication avoids Sender and Receiver being fooled from illicit user. In order to provide authentication, we use the notion of digital certificate. It embodies the hashed user's public key which is encrypted with the private key of a trusted certification authority (CA) to form digital signature. The CA issues certificate to the communicating entities upon request. They verify each other's certificate by decrypting signature with CA's public key. With this they will get assurance that each other is communicating with legitimate user. We have implemented authentication using OpenSSL tool which provides X.509 authentication service. X.509 is an important yardstick to grant authentication because the pile of certificate and authentication protocols stated are used in wide range of applications. It is being formed on the use of public key cryptography and digital signatures which entail the use of a hash function [12, 23, 24]. Figure 1 depicts the generation of digital certificate.

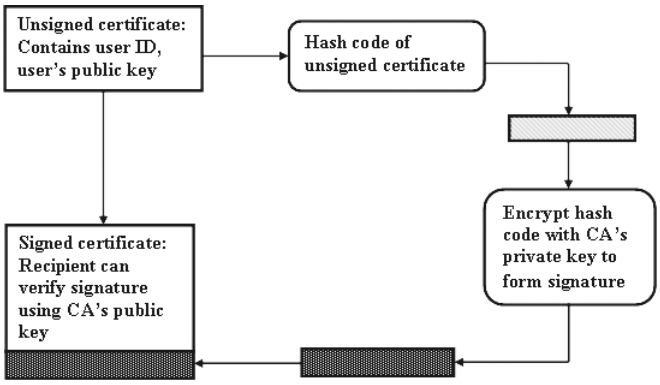


Figure 1: Digital certificate

The authentication process is carried out as shown in the following steps and described in Figure 2.

Step 1. Sender and Receiver requests digital certificate from Trusted CA before the commencement of communication.

Step 2. CA issues digital certificates to Sender and Receiver.

Step 3. Sender sends certificate to the Receiver for verification.

Step 4. Receiver sends certificate to the Sender for verification.

Step 5. Sender verifies Receiver's certificate whether he/she is alleged user or not by using CA's public key and also Receiver verifies Sender's certificate.

3.2 PVK16 QKD Protocol - Quantum Key Distribution Using Hadamard Gate

Hadamard gate is a one-qubit quantum gate. It takes either qubit $|0\rangle$ or $|1\rangle$ as input and gives $1/\sqrt{2}(|0\rangle + |1\rangle)$ or $1/\sqrt{2}(|0\rangle - |1\rangle)$ as output respectively [2]. The Hadamard gate is denoted as —H—.

The matrix representation of quantum gates is derived using tensor product. The tensor product combines two vector spaces to form a larger one. If U and V are vector spaces of dimensions m and n respectively then $U \otimes V$ is a space on mn. The elements of $U \otimes V$ are linear combinations of tensor products $|u> \otimes |v>$ of elements of $|u>$ of U and $|v>$ of V. Suppose A is a m by m matrix and

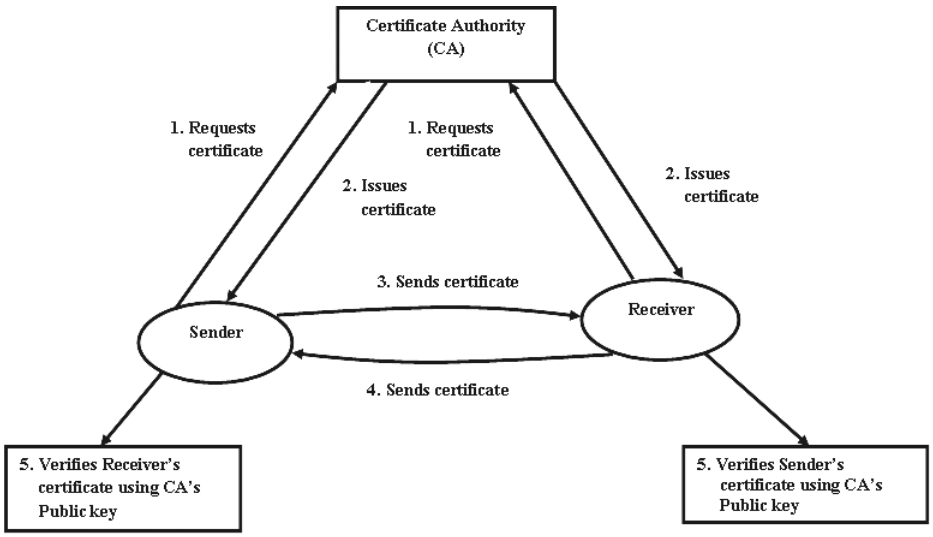


Figure 2: The process of issuing and verifying digital certificate

B is a p by q matrix, then the matrix representation is given as

$$A \otimes B = \begin{bmatrix} A_{11}B & A_{12}B & \dots & A_{1n}B \\ A_{21}B & A_{22}B & \dots & A_{2n}B \\ \vdots & \vdots & \vdots & \vdots \\ A_{m1}B & A_{m2}B & \dots & A_{mn}B \end{bmatrix}$$

Hence, to represent Hadamard gate which is one-qubit gate, a 2 by 2 matrix is needed as shown below [2, 17].

$$H = 1/\sqrt{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$|0\rangle \rightarrow H = 1/\sqrt{2}(|0\rangle + |1\rangle)$$

$$|1\rangle \rightarrow H = 1/\sqrt{2}(|0\rangle - |1\rangle)$$

The procedure of new protocol of QKD using Hadamard gate namely PVK16 QKD protocol is depicted in Figure 3.

Step 1. Sender's preparation of qubits.

Sender selects random bits 'b' in sequence manner where $b \in (\{0,1\}^n)$, n is the original length of the secret key. Then he/she prepares qubits for those bits individually in one of the four states of BB84 protocol using rectilinear '+' and diagonal 'X' bases shown in Figure 4. The BB84 protocol states are denoted as $|\phi\rangle(0,+)$, $|\phi\rangle(0,X)$, $|\phi\rangle(1,+)$ and $|\phi\rangle(1,X)$ symbolized to photons polarized at 0, 45, 90 and 135 degrees respectively [2, 16] which is depicted in Figure 5. For a given string of random bits $b \in (\{0,1\}^n)$ and string of random bases $\theta \in (\{+,X\}^n)$, $|\phi\rangle(b,\theta)$ is denoted as a state for n photons that encodes the bits $b[x]$ in the bases $\theta[x]$ for every $x \in n$.

To carry out the procedure we acquire the following notation:

- $b(\{0,1\}^n)$: Sender's bit string, n is the original length of the secret key.
- $d(\{0,1\}^m)|m \leq n$: Receiver's bit string.
- $\theta_a(\{+,X\}^n)$: Sender's bases in string.
- $\theta_b(\{+,X\}^m)|m \leq n$: Receiver's bases in string.
- $R\{0,1,\theta_b\}^m|m \leq n$: Receiver's measurement string.
- U: Undetected positions.

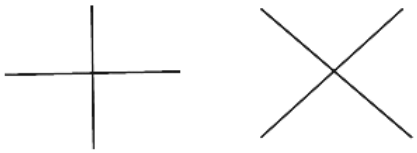


Figure 4: Rectilinear and diagonal bases

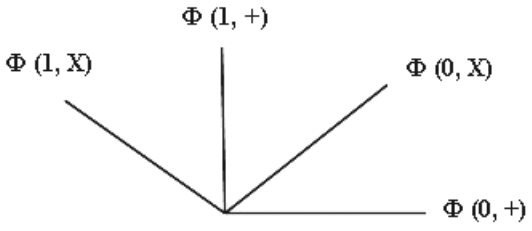


Figure 5: Four states of BB84 protocol

Step 2. Sender applies Hadamard gate.

Sender encodes the qubits $|\phi\rangle(b,\theta_a)$ using Hadamard gate. The outcome is qubits which is denoted as H and sends to Receiver.

Step 3. Receiver's measurement of basis.

Receiver receives H and applies Hadamard gate on

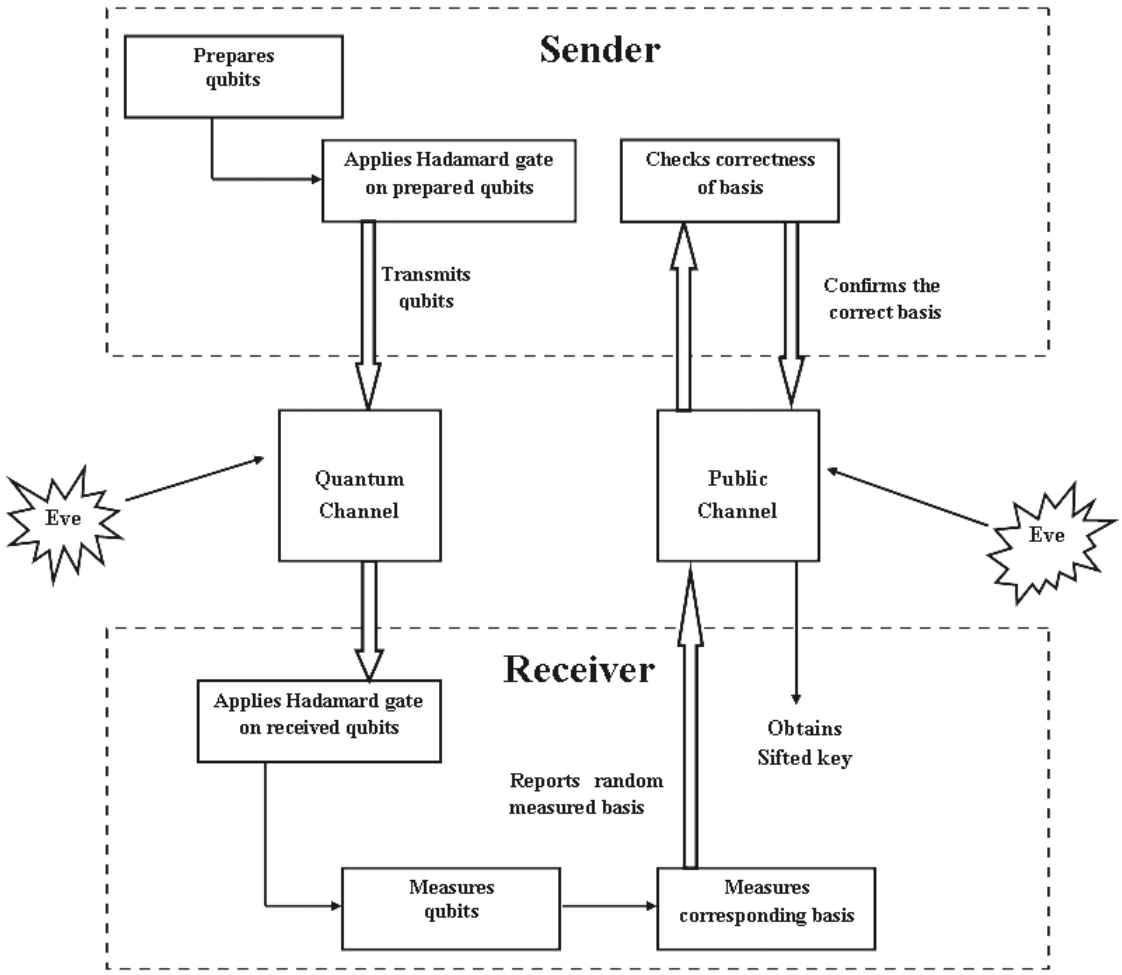


Figure 3: PVK16 QKD protocol

H to get qubits. To restore a qubit which is received to its original state, two Hadamard gates in series fashion are applied [17] which is shown in Figure 6. He/she measures each photon $|\phi\rangle(d, \theta_b)$ using either the rectilinear basis $\{|\phi\rangle(0,+), |\phi\rangle(1,+)\}$ and diagonal basis $\{|\phi\rangle(0,X), |\phi\rangle(1,X)\}$. If Receiver detects a photon i.e. qubit at position x with $|\phi\rangle(d, \theta_b)$ matches with $|\phi\rangle(b, \theta_a)$ i.e. same basis as of Sender's, then the associated outcome is denoted as R, if he/she could able to decode the same Sender's bits for the corresponding basis.

Figure 6: Two Hadamard gates in series restore a quit to its original state

Step 4. Receiver chooses tested basis.
Receiver randomly selects subset of outcome from R and reports to the Sender which is Q i.e. $Q = R \cap U$.

Step 5. Sender says the correctness of basis.
Sender checks $Q[x]$ with $\theta_a[x]$ as well as $b[x]$. If it is correct, gives the confirmation of basis denoted as K otherwise he/she will come to conclusion that eavesdropping/error E has occurred for those basis in position x. $E = U \cup R - k$.

Step 6. Sifted key.
Then both Sender and Receiver shares sifted key $K \in (\{0,1\}^m)|m \leq n$ which is the secret key.

3.3 Error Detection and Calculation

The error is detected when Sender checks $Q[x]$ with $\theta_a[x]$ and $b[x]$. If bits are unmatched, they are discarded. And these bits do not fall under the set of sifted key.

The probability of error is given by $P_e = P_{dx}/P_{bx}$.

- P_e - probability of error.
- P_{dx} - probability of Receiver's measuring bit in x position.
- P_{bx} - probability of Sender's bit position.

Table 1: The comparison of rate of error in key distribution between BB84 and PVK16 QKD protocol

No. of qubits transmitted		16	32	64	128
Rate of error (in %)	BB84 QKD Protocol	56	52	46	43
	PVK16 QKD Protocol	37	33	28	26

4 Results

The results are analyzed by carrying out comparison between BB84 QKD protocol and PVK16 QKD protocol which is displayed in Table 1.

The length of qubit is considered as parameter. The error rate is calculated for qubit sizes 16, 32, 64 and 128. From the observation it is noted that the error rate is less in PVK16 QKD protocol. It is less than 40% in all cases. Therefore, with the obtained sifted key, it is sufficient to carry out further communication using proposed protocol. The retransmission of qubits is certainly not necessary. It is plotted in Figure 7.

The basis for less error rate is that PVK16 QKD protocol is being realized with quantum gates. The complexity of quantum gate is 2^n which is exponential. It is determined that the exponential complexities are secure. It becomes tangle for an eavesdropper to know the exact qubit by measuring with random basis and also becomes a challenge to know the sifted key even with high computational resources.

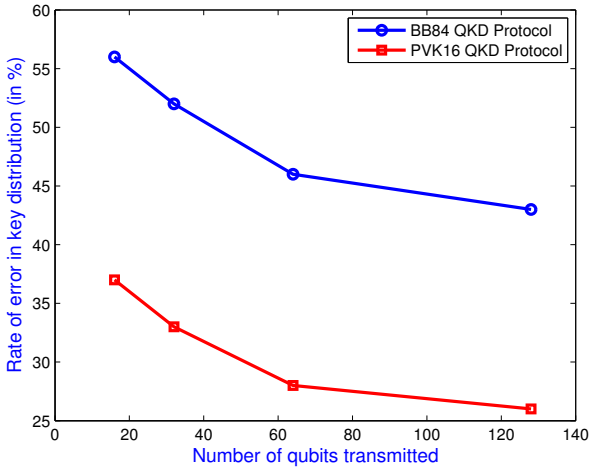


Figure 7: Rate of error in key distribution between BB84 and PVK16 QKD Protocol

5 Security Analysis

The security is analyzed for authentication as well as Quantum key distribution. Authentication is carried out to avoid eavesdropping attack and impersonation. The OpenSSL tool [20] is implemented to ensure that the communicating entities are alleged and both knows that they are communicating with each other, the one whom they have claimed. It was proven that OpenSSL tool [20] is a framework to issue and verify digital certificates. It is used as a defensive mechanism on man in the middle attack and grants trouble free communication.

It is shown that the new regime of protocol that is PVK16 QKD protocol is provably secure in distributing the secret key. The parameter considered is length of qubits. The time complexity of quantum gates is 2^n and which is exponential. Hence, we can convey that PVK16 QKD protocol is more secure. The devised scheme is tailored by reducing few steps from existing. By this the processing period is also reduced as well. It also ensures that it encounters eavesdropping attack. The length of sifted key is more as that of existing scheme which is sufficient to endure further communication. Hence, it avoids retransmission of key.

6 Conclusions

The authentication process through digital certificate is used as a measure for man in the middle attack which is becoming a cumbersome in the path of communication. The fact that the quantum gate has complexity of 2^n assisted in rendering secure communication. It is evidently depicted that PVK16 QKD protocol had essentially encountered man in the middle attack, eavesdropping attack. It is shown that the rate of error is less than 40% and as a result the key size is sufficient to undergo further communication avoiding retransmission of qubits. We thus ascertained the security of using a sifted key distributed by our PVK16 QKD protocol will endow a basis for future systems.

References

[1] H. N. Barnum, "Quantum secure identification using entanglement and catalysis," *arXiv preprint quant-ph/9910072,* 1995.

[2] C. H. Bennet, and G. Brassard, "Quantum cryptography: Public key distribution and coin tossing," in *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing,* pp. 175–179, 1984.

[3] G. Brassard, and C. Crépeau, "25 years of quantum cryptography," *ACM Sigact News,* vol. 27, no. 3, pp. 13–24, 1996.

[4] Y. Chang, C. Xu, and et al., "Quantum secure direct communication and authentication protocol with sin-

gle photons," *Chinese Science Bulletin,* vol. 58, no. 36, pp. 4571–4576, 2013.

[5] M. Fujiwara, T. Domeki, and et al., "Highly secure network switches with quantum key distribution systems," *International Journal Network Security,* vol. 17, no. 1, pp. 34–39, 2015.

[6] L. Gyongyosi and S. Imre, "Quantum informational divergence in quantum channel security analysis," *International Journal of Network Security*, vol. 13, no. 1, pp. 1–12, 2011.

[7] R. J. Hughes, G. Luther, and et al., "Quantum cryptography over underground optical fibers," in *Annual International Cryptology Conference,* pp. 329–342, Springer, 1996.

[8] H. Inamori, L. Rallan, and V. Vedral, "Security of EPR-based quantum cryptography against incoherent symmetric attacks," *Journal of Physics A: Mathematical and General,* vol. 34, no. 35, pp. 6913, 2001.

[9] Y. Kanamori, S. M. Yoo, and et al., "Authentication protocol using quantum superposition states," *International Journal Network Security,* vol. 9, no. 2, pp. 101–108, 2009.

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

[11] 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.

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

[13] N. Lütkenhaus, and M. Jahma, "Quantum key distribution with realistic states: photon-number statistics in the photon-number splitting attack," *New Journal of Physics,* vol. 4, no. 1, pp. 44, 2002.

[14] C. Marcos, and D. J. Santos, "Quantum authentication of classical messages," *Physical Review A,* vol. 64, no. 6, pp. 062309, 2001.

[15] MathWorks, *MATLAB and Statistics Toolbox Release 2014a,* The MathWorks, Inc., Natick, Massachusetts, United States.

[16] D. Mayers, "Unconditional security in quantum cryptography," *Journal of the ACM,* vol. 48, no. 3, pp. 351–406, 2001.

[17] D. McMahon, *Quantum Computing Explained*, John Wiley & Sons, 2007.

[18] A. Muller, H. Zbinden, and N. Gisin, "Underwater quantum coding," *Nature,* vol. 378, no. 6556, pp. 449–449, 1995.

[19] M. A. Nielsen, and I. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, 10th Anniversary Edition, 2002.

[20] OpenSSL, Mar. 31, 2017. (`https://www.openssl.org/source/`)

[21] C. Z. Peng, J. Zhang, and et al., "Experimental long-distance decoy-state quantum key distribution based on polarization encoding," *Physical Review Letters,* vol. 98, no. 1, pp. 010505, 2007.

[22] D. Rosenberg, J. W. Harrington and P. R. Rice, et al., "Long-distance decoy-state quantum key distribution in optical fiber," *Physical Review Letters,* vol. 98, no. 1, pp. 010503, 2007.

[23] W. Stallings, *Cryptography and Network Security: Principles and Practices*, Pearson Education India, 2006.

[24] S. F. Tzeng, M. S. Hwang, "Digital signature with message recovery and its variants based on elliptic curve discrete logarithm problem", *Computer Standards & Interfaces*, vol. 26, no. 2, pp. 61–71, Mar. 2004.

[25] D. Wiedemann, "Quantum cryptography," *ACM Sigact News,* vol. 18, no. 2, pp. 48–51, 1986.

[26] S. Wiesner, "Conjugate coding," *ACM Sigact News,* vol. 15, no. 1, pp. 78–88, 1983.

[27] W. K. Wootters, and W. H. Zurek, "A single quantum cannot be cloned," *Nature,* vol. 299, no. 5886, pp. 802–803, 1982.

[28] C. A. Yen, S. J. Horng, and et al., "Quantum direct communication with mutual authentication," *arXiv preprint arXiv:0903.3444,* 2009.

[29] K. I. Yoshino, T. Ochi, and et al., "Maintenance-free operation of WDM quantum key distribution system through a field fiber over 30 days," *Optics Express,* vol. 21, no. 25, pp. 31395–31401, 2013.

[30] Y. Zhao, B. Qi, and et al., "Simulation and implementation of decoy state quantum key distribution over 60km telecom fiber," *2006 IEEE International Symposium on Information Theory*, pp. 2094–2098, 2006.

Biography

Vurubindi Padmavathi is working as an Associate Professor at Sreenidhi Institute of Science and Technology affiliated to Jawarharlal Nehru Technological University, Hyderabad (JNTUH). She has done her Bachelor of Engineering (CSE), M. Tech. (CSE) and is pursuing Ph.D. in Computer Science and Engineering from JNTUH. Having 14 years of teaching experience, her focus is on the research areas like Cryptography, Information Security and Software Engineering. She has conducted and participated in several workshops, seminars and bridge courses. She has presented and published papers in the International Conferences and in Journals. Mrs. Padmavathi has organized a National conference.

Dr. Bulusu Vishnu Vardhan is working as a Professor in CSE at JNTUH College of Engineering, Nachupally, Karimnagar, Telangana, Inida. He was the Head of the Department of IT, JNTUH College of Engineering, Nachupally from the year 2010 to 2014. He has completed

his M. Tech. from Birla Institute of Technology, Mesra, Ranchi in the year 2001 and completed his Ph.D. from JNTUH in the year 2008. Dr. Vishnu Vardhan has 19 years of teaching experience, presently he is guiding 16 research scholars in the area of Cryptography and Information Security, Information Retrieval, Linguistic processing, Data mining and other elite areas. Three scholars are awarded with Ph.D., one from JNTUK and two from JNTUH. He is the member of Board of Studies for Sathavahana University, Karimnagar. He was an active member in Free Software Movement in India. He has completed Government of Andhra Pradesh funded project worth Rs. 5 lakhs from Ministry of IT on localization activity. As a co-investigator completed another UGC funded project worth Rs. 9 lacks. He has evaluated 4 theses for universities like Osmania, Acharya Nagarjuna University. He visited Singapore and presented paper in International Conference ICAEE in 2014. He has more than 35 papers International Journals and Conferences.

Dr. Addepalli V. N Krishna is a Professor in the department of CSE at Christ University, Bengaluru, India. He has completed M. Tech. and Ph.D. in Computer Science and Engineering discipline. He is in the field of teaching and research since 25 years. A. V. N Krishna has participated in various National and International Conferences. He has many publications to his credit. His areas of interests are Cryptography, Mathematical Modeling and Data Mining.

Achieving Collaborative Cloud Data Storage by Key-Escrow-Free Multi-Authority CP-ABE Scheme with Dual-Revocation

Nyamsuren Vaanchig, Hu Xiong, Wei Chen, and Zhiguang Qin

(Corresponding author: Nyamsuren Vaanchig)

School of Information and Software Engineering & University of Electronic Science and Technology of China
No. 4, North Jianshe Road, Chenghua District, Chengdu, Sichuan 610054, China
(Email: nyamsuren.v@gmail.com)

(Received Oct. 4, 2016; revised and accepted Feb. 20, 2017)

Abstract

Nowadays, more and more users store their data in cloud storage servers for great convenience and real benefits offered by the service, so cloud data storage becomes one of the desirable services provided by cloud service providers. Multi-Authority Ciphertext-Policy Attribute-Based Encryption (MA-CP-ABE) is an emerging cryptographic solution to data access control for large-scale collaborative cloud storage service, which allows any data owner to outsource the data to cloud data storage in order to enable users from collaborating domains or organizations to access the outsourced data. However, the existing MA-CP-ABE schemes cannot be directly applied to collaborative cloud storage services as data access control due to the key escrow problem and the absence of dual revocation mechanism (user revocation and attribute revocation). By addressing these issues, this paper presents a Key-Escrow-Free Multi-Authority Ciphertext-Policy Attribute-Based Encryption Scheme with Dual-Revocation by introducing "the essential attribute" and making use of a certificate authority apart from attribute authorities. Compared with the existing MA-CP-ABE schemes, the proposed scheme is the most suitable one to enable data access control for collaborative cloud storage systems. Furthermore, the security and performance analysis indicates that our scheme is more secure and reasonably efficient to be applied to practical scenarios as collaborative cloud storage systems.

Keywords: Access Control; Attribute Revocation; Collaborative Data Storage; Key Escrow; Multi-authority Ciphertext-policy Attribute-based Encryption; User Revocation

1 Introduction

Cloud storage is one of the popular services offered by cloud service providers, which allows data owners to store their data in third party storage servers for great convenience and real benefits offered by the service. However, this service introduces a great challenge to data access control which addresses how data owners ensure that their data stored in the third party storage servers are accessed by only authorized users [1, 9, 23, 30]. Since data storage servers are not in the same domain with data owners, they can not be fully trusted by data owners to be in charge of making data access decisions on behalf of the data owners.

Ciphertext-Policy Attribute-Based Encryption (CP-ABE) scheme is a great achievement to solve the challenge of data access control over data stored in cloud storage; it is a public key encryption technique designed for one-to-many communications, where data owners hold direct control on their access policies and the access policies are enforced cryptographically [2, 17, 25, 39]. In CP-ABE, each user is entitled to a set of attributes which are associated with the user's secret key. The data owner chooses an access policy over a set of attributes and encrypts the data under the access policy. A user is able to decrypt the ciphertext as long as the set of attributes associated with the user's secret key satisfies the access policy of the ciphertext. This desirable property of CP-ABE makes it suitable for data access control for cloud storage. Since its first introduction [4], there have been extensive research [7, 11, 15, 21, 33] regarding its various aspects. In all these schemes, one trusted central authority is required to handle all attributes in the system and issue a secret key for each user in the system. These schemes thus can be utilized as a data access control for cloud storage where data is encrypted under an access policy over attributes issued by a domain or an organization. However, in large-scale collaborative cloud storage systems, where data owners want to share their data according to the access policies described over attributes issued across different domains and organizations, the trusted central authority may become the bottleneck in the system per-

formance. Furthermore, it is not proper from the point of security due to the single point of trust.

Multi-Authority CP-ABE (MA-CP-ABE) scheme is introduced to overcome the issues mentioned above. In MA-CP-ABE schemes, users may hold attributes issued by multiple attribute authorities and data owners may also encrypt their data under access policies defined over attributes from different attribute authorities. Since, in MA-CP-ABE schemes, attributes are independently managed by different attribute authorities, both the workload and trust are distributed over multiple attribute authority instead of being centralized on a single authority. Since the introduction of MA-ABE [6] , researchers have done wide-ranging research on multi-authority CP-ABE schemes [8, 10, 18, 19, 20, 24, 26, 35, 36] by considering various challenges facing to it.

However, the existing MA-CP-ABE schemes cannot be directly applied to a collaborative cloud storage as data access control due to the key escrow problem and the absence of dual revocation mechanism (user revocation and attribute revocation). Although the existing MA-CP-ABE schemes overcome the issue of the single point of trust, they still hold the key escrow problem in the scope of any attribute authority. Consequently, any attribute authority could decrypt any data which encrypted under an access policy over a set of attributes from only the attribute authority, and the failure or corruption of any attribute authority also raises security issue for the data that encrypted under an access policy described by a set of attributes from one attribute authority. The consequences of these issues are not acceptable for collaborative cloud storage where data owners would like to make their data only accessible to designated users. Moreover, the existing MA-CP-ABE schemes lack dual revocation mechanism dealing with the invalidation of users' access privilege at either attribute level and system level. The mechanism invalidating a user's access privilege at the system level is called user revocation, which addresses the problem of revoking a user's full access privilege from the system when the user is detected as malicious, or the user leaves the system. Whereas the mechanism invalidating a user's access privilege at the attribute level is called attribute revocation, which concerns the issue of revoking a user's partial access privilege when the user's attribute set can be changed dynamically due to the user's role change in the system. Several works [16, 22, 29, 32, 34] have been proposed by solely concerning user revocation issue but in single-authority CP-ABE since the first discussion on user revocation in [4]. However, they are not directly applicable to collaborative cloud data storage systems due to the following drawbacks: vulnerable against collusion attack, scalability problem, and security degradation. Only considering attribute revocation issue, CP-ABE schemes with immediate attribute revocation mechanisms [13, 14, 31, 37, 38] and CP-ABE schemes with time-based attribute revocation mechanisms [27, 31] have been proposed in single-authority CP-ABE settings. Nevertheless, the time-based approaches cause the security degradation problem in terms of forward secrecy until the next expiration time, and the immediate revocation approaches except for the scheme in [37] make the system impurely CP-ABE-based such that it requires the server to be fully trusted. Moreover, the scheme in [37] cannot resist collusion attack from the server and non-revoked users since the updated ciphertexts are transformable to their old versions if the server misuses the update key, and also this scheme results in the stateless user problem. Later on, the similar attribute revocation solutions as those in [13, 31, 37] have been used in the MA-CP-ABE schemes [10, 20, 36], respectively, but these schemes inherit the drawbacks from the corresponding attribute revocation solutions. To the best of our knowledge, there is no MA-CP-ABE scheme which overcomes the key escrow problem and supports secure and scalable dual revocation mechanism (user revocation and attribute revocation).

In this paper, we propose a Key-Escrow-Free Multi-Authority CP-ABE Scheme with Dual-Revocation to enable a secure and scalable data access control for collaborative cloud storage systems. To accomplish the goal of this study, we encounter the following major challenges: (a) how to tie a user's secret keys together to prevent the collusion attack, (b) how to overcome the key escrow problem, (c) how to achieve dual revocation mechanism in which the trust of the cloud server must be reduced. In the proposed scheme, we thus overcome these challenges by introducing a dummy attribute called "the essential attribute" and user-central-key. In addition, our scheme makes use of a certificate authority apart from attribute authorities. The user-central-key is used to tie user's secret keys issued by different authorities; which is issued by the certificate authority. The essential attribute, which is handled by the certificate authority, is used in our construction for the dual-purpose of overcoming the key escrow problem and achieving user revocation mechanism.

For these purposes, in our scheme, each user holds an additional user-attribute-key associated with the essential attribute, and each data is encrypted under an access policy formed from a boolean formula that is expressed as a binary access tree which has two subtrees connected to an AND root gate. While one subtree consists of only leaf node associated with the essential attribute, the other subtree is formed from an access policy defined by the data owner over a set of meaningful attributes. That is to say, in our scheme, an access policy is defined over a set of attributes from at least two authorities: one must be the certificate authority and the other can be any attribute authority from which some attributes are involved in the access policy. Therefore, a data owner uses the public keys of those authorities together as a message encapsulation key, which results in that our scheme overcomes the key escrow problem. To achieve dual revocation mechanism, we assign an attribute secret for each attribute in the system including the essential attribute. Note that certificate authority is the revocation controller in user revocation mechanism, while attribute authorities are the charge of the revocation controller in attribute revoca-

tion mechanism. When either of user revocation or attribute revocation event takes place, the corresponding revocation controller redefines the attribute secret for the involved attribute. Accordingly, the attribute secret for the corresponding component in the public parameters, ciphertexts, and non-revoked users' secret keys must be updated in order to guarantee backward secrecy and forward secrecy.

To do so, we must consider the following practical problems that exist in the existing revocable CP-ABE schemes: the frequent update of public parameters, presence of a fully trusted cloud server, collusion attack that might be triggered by partially revoked or fully revoked user, and stateless user problem. The frequent update of the public parameters that exists in the revocation mechanisms of the previous works is not practical because it requires data owners to be aware of each revocation event and get the updated public parameters to avoid generating an invalid ciphertext no one could decrypt. To solve this impracticality, we introduce ill-formed ciphertext and ciphertext-heal-key components in our construction. The fully trusted cloud server that is considered in revocation mechanism of the previous works makes the system not purely CP-ABE based as well as vulnerable to the collusion attack from the cloud server and any revoked users. Although in our scheme, the cloud server is delegated to update ciphertexts upon any revocation event as similar as in previous works, we must prevent any unauthorized decryption by any fully revoked or partially revoked user with colluding with cloud server which is provided with a ciphertext-update key. Our solution for this challenge is that the ciphertext-update key generated by the revocation controller cannot be misused to transform the updated ciphertext to its old version. Moreover, the revocation controller generates a unique key-update-key for each non-revoked user, which cannot be used by any other user, such that our scheme can resist the collusion attack from any revoked user and non-revoked user. However, key-update-key generation for each user and its distribution to each non-revoked user bring more overhead on the revocation controller upon each revocation event.

In addition, it might result in stateless user problem such that authorized users can not decrypt the ciphertexts to which they authorized to access if the non-revoked users miss some key-update-keys due to some reasons. By simultaneously considering the overhead and the stateless user problem, our scheme uses a system-wide version number for each attribute in the system including the essential attribute to indicate the evolution of the attribute secret of the attribute. As a result, the revocation controller generates a unique key-update-key with the latest version number for the non-revoked user on demand whenever the user requests. The user request is triggered by the detection of the fact that any attribute component of his/her secret key is stateless by checking the corresponding version values in the both ciphertext and user-attribute-key upon a decryption process.

The main contributions of this work are summarized as follows. We provide the first construction of Key-Escrow-Free Multi-Authority CP-ABE Scheme with Dual-Revocation and summarize the comparisons between the proposed scheme and the existing MA-CP-ABE schemes regarding their functionalities, as in Table 1. Dual revocation mechanism of the proposed scheme guarantees both backward and forward secrecy, resists the potential collusion attacks and deals with practical problems that exist in the existing revocable CP-ABE schemes such as frequent update of the public parameters, the presence of a fully trusted cloud server, and the stateless user problem. Furthermore, we prove that the proposed scheme is selectively secure in the standard model under the q-parallel BDHE assumption as well as provide security and performance analysis compared with the existing works. The security and performance analysis show that our scheme is more secure and reasonably efficient to be applied to practical scenarios as collaborative cloud storage systems.

We organize the rest of the paper as follows. In Section 2, we provide some preliminary knowledge used in this work. The system model, algorithm definitions, security model and security properties are described in Section 3. The proposed Key-Escrow-Free Multi-Authority CP-ABE Scheme with Dual-Revocation is presented in Section 4 as a data access control for collaborative cloud storage. The security and performance analysis of the proposed scheme is provided in Section 5. Finally, a conclusion is given in Section 6.

2 Preliminaries

2.1 Access Structures

Definition 1. *(Access Structure [3]) Let* $\{P_1, P_2, ..., P_n\}$ *be a set of parties. A collection* $\mathbb{A} \subseteq 2^{\{P_1,P_2,...,P_n\}}$ *is monotone if* $\mathcal{B} \in \mathbb{A}$ *and* $\mathcal{B} \subseteq \mathcal{C}$ *implies* $\mathcal{C} \in \mathbb{A}$*. An access structure is a monotone collection* $\mathbb{A}$ *of non-empty subsets of* $\{P_1, P_2, ..., P_n\}$*, i.e.,* $\mathbb{A} \subseteq 2^{\{P_1,P_2,...,P_n\}} \setminus \{\emptyset\}$*. The sets in* $\mathbb{A}$ *are called the authorized sets, and the sets not in* $\mathbb{A}$ *are called the unauthorized sets.*

In our context, the role of the parties is taken by the attributes. Thus, the access structure $\mathbb{A}$ will contain the authorized set of attributes. We only consider monotone access structures.

2.2 Linear Secret Sharing Scheme

Definition 2. *(Linear Secret Sharing Scheme (LSSS) [3]). A secret-sharing scheme* $\prod$ *over a set of parties* $\mathcal{P}$ *is called linear over* $\mathbb{Z}_p$ *if*

1) *The shares of each party form a vector over* $\mathbb{Z}_p$*.*

2) *There exists a matrix an* M *with* l *rows and* n *columns called the secret-generating matrix for* $\prod$*. For all* $i = 1, ..., l$*, the* i*-th row of* M *is labeled by a party* $\rho(i)$*, where* ρ *is a function that maps a row to a*

Table 1: Comparison of multi-authority CP-ABE schemes

Properties	**[26]**	**[18]**	**[24]**	**[36]**	**[28]**	**[20]**	**[10]**	**Our**
Master Authority Free	No	Yes	No	Yes	Yes	Yes	Yes	Yes
Scalability	Yes	Yes	No	No	Yes	No	Yes	Yes
Key Escrow Free	No	No	Yes	No	No	No	No	Yes
Attribute Revocation	No	No	No	Yes	No	Yes	Yes	Yes
User Revocation	No	No	No	No	No	No	No	Yes
Bilinear Group Order	Prime	Composite	Composite	Prime	Prime	Composite	Composite	Prime
Security Model	GGM	ROM	ROM	ROM	ROM	SM	ROM	SM

* GGM: Generic Group Model, ROM: Random Oracle Model, SM: Standard Model

party for labeling. When we consider the column vector $\vec{v} = (s, r_2, ..., r_n)$, where $s \in \mathbb{Z}_p$ is the secret to be shared and $r_2, ..., r_n \in \mathbb{Z}_p$ are randomly chosen, then $M\vec{v}$ is the vector of l shares of the secret s according to $\prod$. The share $(M\vec{v})_i$ belongs to party $\rho(i)$.

According to [3], each linear secret sharing scheme meets *linear reconstruction* property defined as follows: Suppose that $\prod$ is a LSSS for the access structure $\mathbb{A}$. Let $\mathcal{S} \in \mathbb{A}$ be any authorized set, and let $I \subset \{1, 2, ..., l\}$ be the set of rows whose labels are in $\mathcal{S}$, i.e. $I = \{i : \rho(i) \in \mathcal{S}\}$. Then there exists constants $\{w_i \in \mathbb{Z}_p\}_{i \in I}$ such that if $\{\lambda_i\}$ are valid shares of any secret s according to $\prod$, then $\sum_{i \in I} w_i \lambda_i = s$. Additionally, it is shown in [3] that the constants $\{w_i\}_{i \in I}$ can be found in time polynomial in the size of the secret-generating matrix M.

2.3 Bilinear Map

Let $\mathbb{G}_0$ and $\mathbb{G}_1$ be two multiplicative cyclic groups of prime order p. Let g be a generator of $\mathbb{G}_0$ and e be a bilinear map, $e : \mathbb{G}_0 \times \mathbb{G}_0 \to \mathbb{G}_1$ with the following properties:

1) *Bilinearity*: for all $u, v \in \mathbb{G}_0$ and $a, b \in \mathbb{Z}_p$, we have $e(u^a, v^b) = e(u, v)^{ab}$.

2) *Non-degeneracy*: $e(g, g) \neq 1$.

3) *Computability*: There is an efficient algorithm to compute $e(u, v)$ for $\forall u, v \in \mathbb{G}_0$.

2.4 Decisional q-Parallel Bilinear Diffie-Hellman Exponent Assumption

Choose a group $\mathbb{G}_0$ of prime order p according to the security parameter. Let $a, s, b_1, ..., b_q \in \mathbb{Z}_p$ be chosen at random and g be a generator of $\mathbb{G}_0$. If an adversary is given $\vec{y} =$

$$g, g^s, g^a, ..., g^{(a^q)}, , g^{(a^{q+2})}, ..., g^{(a^{2q})}$$
$$\forall_{1 \leq j \leq q} g^{s \cdot b_j}, ..., g^{a/b_j}, ..., g^{(a^q/b_j)}, , g^{(a^{q+2}/b_j)}, ..., g^{(a^{2q}/b_j)}$$
$$\forall_{1 \leq j,k \leq q, k \neq j} g^{a \cdot s \cdot b_k / b_j}, ..., g^{(a^q \cdot s \cdot b_k / b_j)}$$

it must remain hard to distinguish $e(g, g)^{a^{q+1}s} \in \mathbb{G}_1$ from a random element in $\mathbb{G}_1$.

An algorithm $\mathcal{B}$ that outputs $z \in \{0, 1\}$ has advantage ϵ in solving decisional q-parallel BDHE in $\mathbb{G}_0$ if

$$\left| Pr[\mathcal{B}(\vec{y}, T = e(g, g)^{a^{q+1}s}) = 0] - Pr[\mathcal{B}(\vec{y}, T = R) = 0] \right| \geq \epsilon.$$

Definition 3. *The decisional q-parallel BDHE assumption holds if no polynomial time algorithm has a non-negligible advantage in solving the q-parallel BDHE problem.*

3 System and Security Models

3.1 System Model

We consider a collaborative cloud storage system as illustrated in Figure 1, and the system consists of the following entities:

Certificate Authority (CA): It sets up the system. In addition, CA is in charge of authorizing and revoking users' access privileges to and from the system. For this purpose, it handles the essential attribute such that it acts as an attribute authority handling only one attribute.

Attribute Authorities (AA): It sets up its own domain. Each AA is responsible for entitling and revoking attributes for and from users according to their roles in its domain. In our system, every attribute belongs to a single AA, but each AA manages an arbitrary number of attributes.

Cloud Server (CS): It stores data owners' data in encrypted form and provides data storage service to data users. The CS is not involved in access control enforcement or data decryption process. It means that, in our system, data access decisions are made cryptographically such that only authorized users can obtain data without depending any decision or process done by CS. We assume that CS is minimally trusted such that it converts ill-formed ciphertexts to well-formed ciphertexts upon data publication phase and updates well-formed ciphertexts to their latest version upon any revocation event.

Data Owners (DO): A data owner defines an access policy over a set of attributes from the relevant AAs

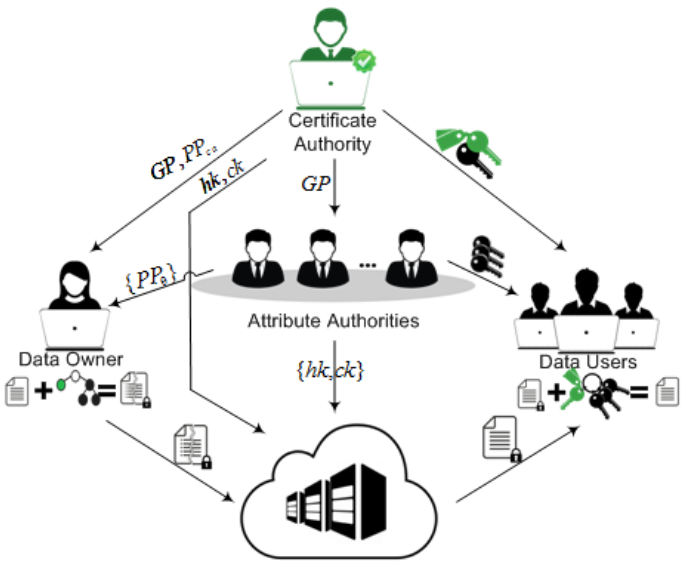


Figure 1: System architecture of collaborative cloud data storage

including CA and encrypts his/her data under the access policy. Then the DO publishes the encrypted data in ill-formed way to CS. The data access decision to this data is enforced cryptographically according to the access policy in the ciphertext without any dependence from CS. It means that any authorized user who has sufficient attributes can decrypt the ciphertext to obtain the data.

Data Users (DU): Each DU is assigned with a global identity from CA and entitled to a set of attributes from multiple AAs including the essential attribute from CA. Any DU can freely get any ciphertext from CS and can decrypt the ciphertext if and only if the DU's attribute set satisfies the access policy of the ciphertext. However, any DU's attribute set may dynamically change due to his/her role change in the different domains in the system, and this case will be handled by the attribute revocation mechanism of our system. Moreover, any DU can leave the system or detected as a malicious DU, and this case will be handled by the user revocation mechanism of our system.

3.2 Algorithms

The proposed scheme consists of the following algorithms:

- $GSetup(\lambda) \rightarrow GP, SK, VK$. This algorithm, run by CA, takes the security parameter λ as input and outputs the global parameters GP for the system. In addition, CA generates a pair of sign and verification keys, SK, VK.
- $ASetup(GP) \rightarrow MK_\theta, PP_\theta, \{hk_{x_\theta}^{h(ver_{x_\theta})}\}$. This algorithm, run by an AA, takes the global parameters GP as input, and it outputs the AA's master key MK_θ and public parameters PP_θ. It also outputs a set of ciphertext-heal keys $\{hk_{x_\theta}^{h(ver_{x_\theta})}\}$ for all the attributes $\{x_\theta\}$ that the AA handles, where $h(ver_{x_\theta})$ is the version value of the attribute x_θ.
- $GKeyGen(GP, gid) \rightarrow G_{gid}$. This algorithm, run by CA, takes the global parameters GP and a global identifier gid of a DU as input, and it outputs a user-central-key G_{gid} for the DU.
- $AKeyGen(GP, MK_\theta, G_{gid}, \mathcal{S}_{gid,\theta}) \rightarrow AK_{gid,\theta}$. This algorithm, run by an AA, takes the global parameters GP, the AA's master key MK_θ and a DU's user-central-key G_{gid} and an attribute set $\mathcal{S}_{gid,\theta}$ that describes the DU's role in the domain as input. It then outputs an user-attribute-key $AK_{gid,\theta}$ for the DU.
- $Encrypt(m, (M, \rho), GP, \{PP_\theta\}_{\theta \in F}) \rightarrow \tilde{ct}$. This algorithm, run by a DO, takes data m, an LSSS access structure (M, ρ) , the global parameters GP, and a set of public parameters $\{PP_\theta\}$ of the relevant AAs including CA as input. Here, θ is index of the relevant authorities including both CA and the AAs. It outputs an ill-formed ciphertext $\tilde{ct}$ that no one can decrypt.
- $CTConvert(\tilde{ct}, \{hk_{x_\theta}\}_{x_\theta=\rho(i),\theta \in F}^{h(ver_{x_\theta})}) \rightarrow CT$. This algorithm, run by CS, takes an ill-formed ciphertext $\tilde{ct}$ and a set of ciphertext-heal keys $\{hk_{x_\theta}\}_{x_\theta=\rho(i),\theta \in F}^{h(ver_{x_\theta})}$ for all the attributes included in the access structure of the ciphertext as input. It then outputs a well-formed ciphertext CT.
- $Decrypt(CT, G_{gid}, \{AK_{gid,\theta}\}_{\theta \in F}) \rightarrow m|\perp$. This algorithm, run by a DU, takes a well-formed ciphertext CT, a DU's user-central-key G_{gid} and a set of user-attribute-keys $\{AK_{gid,\theta}\}$ of the DU as input. It then outputs the data m when the DU's attribute set $\mathcal{S}_{gid}$ satisfies the access structure (M, ρ) of the CT. Otherwise the decryption fails and returns $\perp$.
- $CKeyGen(MK_\theta, y_\theta) \rightarrow hk_{y_\theta}^{h(ver'_{y_\theta})}, ck_{y_\theta}^{h(ver'_{y_\theta})}, MK'_\theta$. This algorithm, run by an AA, takes the AA's master key MK_θ, an attribute y_θ involved in the revocation event as input. It then outputs a ciphertext-heal-key $hk_{y_\theta}^{h(ver'_{y_\theta})}$, a ciphertext-update-key $ck_{y_\theta}^{h(ver'_{y_\theta})}$, and an updated master key MK'_θ.
- $CTUpdate(CT, ck_{y_\theta}^{h(ver'_{y_\theta})}) \rightarrow CT'$. This algorithm, run by CS, takes a ciphertext CT and a ciphertext-update key $ck_{y_\theta}^{h(ver'_{y_\theta})}$ as input, and it outputs an updated ciphertext CT'.
- $UKeyGen(MK_\theta, y_\theta, G_{gid}) \rightarrow uk_{gid,y_\theta}^{h(ver'_{y_\theta})}$. This algorithm, run by an AA, takes the AA's master key

MK_θ, an attribute y_θ, and an user-central-key G_{gid} as input. It then outputs a key-update-key $uk_{gid,y_\theta}^{h(ver'_{y_\theta})}$ associated with the attribute y_θ that a DU with the user-central-key G_{gid} is requested.

- $AKUpdate(AK_{gid,\theta}, uk_{gid,y_\theta}^{h(ver'_{y_\theta})}) \rightarrow AK'_{gid,\theta}$. This algorithm, run by a DU, takes the DU's user-attribute-key $AK_{gid,\theta}$ and a key-update-key $uk_{gid,y_\theta}^{h(ver'_{y_\theta})}$ for the DU as input, and it outputs an updated user-attribute-key $AK'_{gid,\theta}$ for the DU.

Definition 4. *The Key-Escrow-Free MA-CP-ABE Scheme with Dual-Revocation is correct if for any GP generated by GSetup(λ) algorithm, for any set of $\{PP_\theta, MK_\theta, hk_{x_\theta}\}$ generated by ASetup(GP) algorithm, for any CT encrypted by Encrypt$(m, (M, \rho), GP, \{PP_\theta\}_{\theta \in F})$ algorithm on data m and converted by CTConvert$(\tilde{ct}, \{hk_{x_\theta}\}_{x_\theta = \rho(i), \theta \in F}^{h(ver_{x_\theta})})$ algorithm, for any secret key $(G_{gid}, \{AK_{gid,\theta}\})$ generated by GKeyGen(GP, gid) and AKeyGen$(GP, MK_\theta, G_{gid}, \mathcal{S}_{gid,\theta})$ algorithms, it is true that Decrypt$(CT, G_{gid}, \{AK_{gid,\theta}\}_{\theta \in F}) = m$ if and only if the attribute set $\mathcal{S}_{gid} = \cup_{\theta \in F} \mathcal{S}_{gid,\theta}$ satisfies (M, ρ) and the version values of the corresponding components in the both of $\{AK_{gid,\theta}\}$ and CT are match.*

3.3 Security Model and Security Requirements

We consider potential attackers in the collaborative cloud storage system as follows: 1) The CA and each AA are assumed to be honest such that each of them does not collude with any other entity. However, CA or any AA can be corrupted by attackers, and also it should be prevented from decrypting any ciphertexts individually. 2) The CS is assumed to be minimally trusted. It might attempt to obtain the content of the encrypted data although it correctly performs the tasks assigned by legitimate entities. 3) Each DU is assumed to be dishonest and malicious, and he/she might attempt to obtain access to data beyond his/her access privilege. To simplicity, we classify dishonest and malicious DUs in the system into three categories: (a) unauthorized DU is a user who does not have sufficient attributes satisfying the access policy of the encrypted data. (b) partially revoked DU is a user whose attribute set no longer satisfies the access policy of the encrypted data. (c) fully revoked DU is a user whose access privilege is no longer valid in the system.

Now, we present a security model for the proposed Key-Escrow-Free Multi-Authority CP-ABE Scheme with Dual-Revocation. The security model is described by a game between a challenger $\mathcal{B}$ and an adversary $\mathcal{A}$. The phases of the game are following:

Global Setup. The $\mathcal{B}$ runs *GSetup* algorithm and sends the global parameters to the $\mathcal{A}$.

Init. In this phase, The $\mathcal{A}$ specifies a set of corrupted authorities $\mathcal{C} \subseteq \mathcal{N}$, where $\mathcal{N}$ is the set of all authorities. We assume that the $\mathcal{A}$ can corrupt at most $|\mathcal{N}| - 1$ number of authorities. In addition, the $\mathcal{A}$ declares a challenge access structure (M^*, ρ^*) along with the target version numbers $\{ver^*_{x_\theta}\}_{x_\theta = \rho^*(i)}$ for each attribute included in the challenge access structure, which he will try to attack.

Authority Setup. For non-corrupted authorities $\mathcal{N} - \mathcal{C}$, the $\mathcal{B}$ obtains the public parameters and master key by running *ASetup* algorithm and gives the public parameters to the $\mathcal{A}$.

Phase 1. The $\mathcal{A}$ makes secret key and key-update-key queries adaptively as follows:

- Secret key queries: A secret key consists of a user-central-key and a set of user-attribute-keys for an attribute set $\mathcal{S}_{gid}$. The $\mathcal{A}$ makes secret key queries by submitting $\mathcal{S}_{gid} = \bigcup_{\theta \in \mathcal{N} - \mathcal{C}} S_{gid,\theta}$, where $\mathcal{S}_{gid} \cup \bigcup_{\theta \in \mathcal{C}} U_\theta$ does not satisfy the challenge access structure. The $\mathcal{B}$ responds to each query by returning a user-central-key G_{gid} along with set of user-attribute-keys $\{AK_{gid,\theta}\}_{\theta \in \mathcal{N} - \mathcal{C}}$.
- Key-update-key queries: The $\mathcal{A}$ makes key-update-key queries by submitting (x_θ, G_{gid}) along with a version number ver_{x_θ} for the attribute x_θ, where x_θ must be in $S_{gid} = \bigcup_{\theta \in \mathcal{N} - \mathcal{C}} S_{gid,\theta}$ and $2 \leq ver_{x_\theta} \leq ver^*_{x_\theta}$. The $\mathcal{B}$ responds by returning the corresponding key-update-key $uk_{gid,x_\theta}^{h(ver_{x_\theta})}$ to the $\mathcal{A}$.

Challenge. The $\mathcal{A}$ submits two equal length messages m_0 and m_1 to the $\mathcal{B}$. The $\mathcal{B}$ flips a random coin $b \in \{0, 1\}$, encrypts m_b under the access structure (M^*, ρ^*) and sends the ciphertext CT^* to the $\mathcal{A}$, where the version numbers of each attribute in the access structure are equal to the corresponding version number $ver^*_{p(\cdot)}$ given with the challenge access structure.

Phase 2. Phase 1 is repeated.

Guess. The $\mathcal{A}$ makes a guess b' for b and it wins if $b = b'$.

The advantage of an adversary $\mathcal{A}$ in this game is defined as $Pr[b = b'] - 1/2$.

Definition 5. *Key-Escrow-Free Multi-Authority CP-ABE Scheme with Dual-Revocation is selectively secure (against static corruption of authorities) if all polynomial time adversaries have at most a negligible advantage in this security game.*

Our scheme also guarantees the following security requirements which are basic requirements for revocation:

Collusion Resistance: Any fully revoked/partially revoked DU should be prevented from decrypting any

ciphertext, to which he/she is not authorized to access, by colluding with CS (type-A) or any non-revoked DU (type-B).

Forward Secrecy: Any partially revoked DU should be prevented from decrypting any ciphertext which requires the attributes which are revoked from the DU to decrypt. Any fully revoked DU should be prevented from decrypting any ciphertext.

Backward Secrecy: Any new/non-revoked DU should be able to decrypt any ciphertext as long as the DU has sufficient attribute set satisfying the access structure of the ciphertext.

4 Data Access Control by Key-Escrow Free Multi-Authority CP-ABE with Dual Revocation

4.1 Overview

We now present our Key-Escrow-Free Multi-Authority CP-ABE Scheme with Dual-Revocation that can enable a secure and scalable data access control for collaborative cloud storage systems. As previously mentioned, our scheme introduces a dummy attribute called the essential attribute and makes use of a certificate authority (CA) apart from attribute authorities (AAs). Although our scheme separates authorities into a CA and multiple AAs, CA acts as an AA by handling the essential attribute besides setting up the system. The essential attribute is used in our scheme for dual-purpose: to achieve user revocation mechanism and key escrow free property. The essential attribute, therefore, must be included in the access structure of each ciphertext. Moreover, each user is required to possess an additional user-attribute-key which is associated with the essential attribute besides the DU's user-attribute-keys that are associated with a different set of attributes from different domains. Note that CA will be included in all the AA notation throughout this paper since it is involved in two different roles in the system.

In our construction, an access structure is formed from a boolean formula that expressed as a binary access tree which has two subtrees connected to an AND root gate. While one subtree consists of only leaf node associated with the essential attribute, the other subtree is formed from an access policy defined by the DO over a set of meaningful attributes, where interior nodes are AND or OR gates and the leaf nodes correspond to the attributes. Using the LSSS generation algorithm [18], one can convert any access trees into their equivalent LSSS matrices. Since the access structure is defined over a set of attributes from different AAs including CA, a DO encrypts his/her data under the access structure by using the public parameters of the relevant AAs including CA; it results in an ill-formed ciphertext. Before publishing the ill-formed ciphertext, CS converts it into a well-formed ciphertext by using a set of the corresponding ciphertext-heal-keys . To decrypt any ciphertext, a DU must have a user-central-key issued by CA and a set of user-attribute-keys satisfying the access structure of the ciphertext, in which the user-attribute-key associated with the essential attribute must be valid.

Our scheme supports both user revocation and attribute revocation, and we call it dual revocation. We achieve both user revocation and attribute revocation mechanism by applying the same technique, but CA is in charge of user revocation controller by handling the essential attribute and any AA is in charge of attribute revocation controller by handling the corresponding attribute. In our revocation mechanism, the revocation controller defines an attribute secret and a system-wide version number for each attribute in the system at the domain setup phase. The version number of the attribute indicates the evolution of the attribute's attribute secret, and all the version numbers are initially set to 1. Moreover, each version number is hashed to a version value which will be embedded into the corresponding attribute component in ciphertexts and user-attribute-keys. Whenever any revocation event takes place, the attribute secret for the involved attribute is replaced with a new one, and its version number is increased by 1. Since the attribute secret for the involved attribute is redefined, the corresponding attribute component in the public parameters, ciphertexts and non-revoked DUs' attribute-secret-keys must be updated in order to guarantee backward secrecy. In order to avoid the impracticality of the frequent update of the public parameters, we introduce ill-formed ciphertext and ciphertext-heal key components in our scheme, as mentioned above. The ciphertext-heal keys are generated by the revocation controller for the purpose of the subsequent data publication data instead of updating the corresponding component in the public parameters. The revocation controller also generates a ciphertext-update key for updating the corresponding attribute component in ciphertexts. Once CS receives the ciphertext-update key, it updates the corresponding attribute component in the ciphertexts by using proxy re-encryption technique [5] but in a unidirectional way. For the purpose of updating the corresponding attribute component in non-revoked DUs' user-attribute-keys, the revocation controller generates a unique key-update-key for each non-revoked DU on demand whenever the DU requests. The DU's request is triggered by the detection of whether any attribute component of his/her user-attribute-key is stateless. This detection is made upon the decryption process by matching the corresponding version values in both the ciphertext and the user-attribute-key. Upon receiving the key-update-key, the DU updates his/her key to its latest version. The increased version values are embedded within ciphertext-heal-keys, ciphertext-update-keys and key-update-keys; With these keys, the version values of the corresponding attribute components in ciphertexts and non-revoked DUs' user-attribute-keys are updated to their latest version as well.

4.2 Construction

Let $\mathcal{N}$ denote the set of all the authorities including CA and A_θ be an authority with the index θ. In addition, let U_θ denote a set of attributes managed by an A_θ and let $F \subseteq \mathcal{N}$ denote a set of the relevant authorities including CA from them some attributes are involved in the access structure for the encryption.

The construction of our scheme is presented as follows:

System Initialization: The CA sets up the system by running the *GSetup* algorithm and also sets up its own domain by running the *ASetup* algorithm. After setting up its domain, CA sends a ciphertext-heal-key related to the essential attribute to CS. In the Collaborative Cloud Storage System, any entity can simply act as an AA by setting up its own domain with the *ASetup* algorithm. After setting up its domain, it sends a set of ciphertext-heal-keys, each is related to an attribute of the domain, to CS.

- $GSetup(\lambda)$. It first chooses a bilinear group $\mathbb{G}_0$ of prime order p with generator g and a bilinear map $e : \mathbb{G}_0 \times \mathbb{G}_0 \to \mathbb{G}_1$. In addition, CA generates a pair of sign SK and verification VK keys. The global parameters are published as: $GP = (e, g, \mathbb{G}_0, \mathbb{G}_1, VK)$.
- $ASetup(GP)$. It first chooses randomly $\alpha_\theta, \beta_\theta, \delta_\theta, a_\theta \in \mathbb{Z}_p$ as its master key and $\{t_{x_\theta}, v_{x_\theta}\}_{x_\theta \in U_\theta} \in \mathbb{Z}_p$ as attribute secrets for the attributes in the domain. Next, it sets each attribute's version number ver_{x_θ} as 1 and selects a non-cryptographic hash function $h(\cdot)$ that maps a version number to a hash value. It then outputs the authority's master key MK_θ and public parameters PP_θ as follows:

$$\begin{aligned} MK_\theta &= (\alpha_\theta, \beta_\theta, \delta_\theta, a_\theta, \\ &\quad \{ver_{x_\theta} = 1, t_{x_\theta}, v_{x_\theta}\}_{x_\theta \in U_\theta}); \\ PP_\theta &= (e(g,g)^{\alpha_\theta}, g^{a_\theta}, g^{\delta_\theta \beta_\theta}, \\ &\quad \{h(ver_{x_\theta}), T_{x_\theta} = g^{t_{x_\theta}\beta_\theta}\}_{x_\theta \in U_\theta}). \end{aligned}$$

  Finally, it generates a set of ciphertext-heal-keys for all attributes in the domain as follows:

$$\{hk_{x_\theta}^{h(ver_{x_\theta})} = \frac{v_{x_\theta} - t_{x_\theta}}{\delta_\theta}\}_{x_\theta \in U_\theta}$$

User and Attribute Authorization: When a new DU joins the system, CA assigns a globally unique identifier *gid* to the DU and issues a user-central-key by running the *GKeyGen*, on which a signature of CA. Also, CA generates a user-attribute-key associated with the essential attribute for the DU by running the *AKeyGen* algorithm. Then, CA sends them to the DU through a secure channel. To get a user-attribute-key, any DU submits its user-central-key to an AA. Each AA authenticates any DU by verifying the signature on the DU's user-central-key with the verification key VK issued by CA. If the DU is a legal DU, then the AA entitles a set of attributes $\mathcal{S}_{gid,\theta}$ to the DU according to his/her role in the domain and then generates a user-attribute-key for the DU by running the *AKeyGen* algorithm. Then, the AA sends the user-attribute-key to the DU through a secure channel.

- $GKeygen(GP, gid)$. It first chooses a random $u_{id} \in \mathbb{Z}_p$ and generates a user-central-key for the DU as follows:

$$G_{gid} = g^{u_{id}}.$$

  Then, it signs on it with the sign-key SK to ensure its integrity.
- $AKeyGen(GP, MK_\theta, G_{gid}, \mathcal{S}_{gid,\theta})$. It generates a user-attribute-key which is associated with the DU's attribute set in the domain as follows:

$$\begin{aligned} AK_{gid,\theta} &= \Big(D_{1_\theta} = g^{\alpha_\theta}(G_{gid})^{a_\theta}, \forall x_\theta \in \mathcal{S}_{gid,\theta} : \\ &\quad \{h(ver_{x_\theta}), D_{x_\theta} = (G_{gid})^{v_{x_\theta}\beta_\theta}\}\Big) \end{aligned}$$

Data Publication: Before outsourcing data m, a DO defines an access structure (M, ρ). Then, the DO encrypts the data m under the (M, ρ) by running the *Encrypt* algorithm and submits an ill-formed ciphertext as the output to CS. After receiving the ill-formed ciphertext, CS converts it to a well-formed ciphertext by running the *CTConvert* algorithm. Finally, CS publishes the well-formed ciphertext to the collaborative cloud storage.

- $Encrypt(m, (M, \rho), GP, \{PP_\theta\}_{\theta \in F})$. Let M be an $l \times n$ matrix. The function ρ associates rows of M to attributes. Then, it chooses a random vector $\vec{v} = (s, y_2, ..., y_n) \in \mathbb{Z}_p^n$, where $y_2, ..., y_n$ will be used to share the encryption exponent s. For $i = 1$ to l, it calculates $\lambda_i = \vec{v} \cdot M_i$, where M_i is the vector corresponding to the i^{th} row of M. In addition, it randomly chooses $r_1, r_2, ..., r_l \in \mathbb{Z}_p^l$ and computes an ill-formed ciphertext as:

$$\begin{aligned} \tilde{ct} &= \Big(C_0 = m\Big(\prod_{\theta \in F} e(g,g)^{\alpha_\theta}\Big)^s, C_1 = g^s, \\ &\quad \{h(ver_{\rho(i)}), C_{i,1} = g^{r_i}, C_{i,2} = g^{-\delta_\theta \beta_\theta r_i}, \\ &\quad \tilde{C}_{i,3} = (\prod_{\theta \in F} g^{a_\theta})^{\lambda_i} T_{\rho(i)}^{-r_i}\}_{i=1...l}\Big). \end{aligned}$$

  Here, $h(ver_{\rho(i)})$ is set by the version value of the corresponding T_{x_θ} in the PP_θ, where $x_\theta = \rho(i)$.
- $CTConvert(\tilde{ct}, \{hk_{x_\theta}\}_{x_\theta = \rho(i), \theta \in F}^{h(ver_{x_\theta})})$. By using a set of ciphertext-heal keys $\{hk_{x_\theta}\}_{x_\theta = \rho(i), \theta \in F}$, it converts an ill-formed ciphertext to a well-formed ciphertext as:

$$\begin{aligned} CT &= \Big(C_0, C_1, \{h(ver_{\rho(i)}), C_{i,1}, C_{i,2}, \\ &\quad C_{i,3} = \tilde{C}_{i,3} \cdot (C_{i,2})^{hk_{\rho(i)}^{h(ver_{x_\theta})}}\}_{i=1...l}\Big). \end{aligned}$$

Here, $h(ver_{\rho(i)})$ is set by the version value of the corresponding ciphertext-heal-key $hk_{x_\theta}^{h(ver_{x_\theta})}$, where $x_\theta = \rho(i)$.

Data Retrieval: Any DU can freely get any ciphertext CT from CS and obtain the data by running the *Decrypt* algorithm if and only if the DU possesses sufficient attributes $\mathcal{S}_{gid} = \bigcup_{\theta \in F} \mathcal{S}_{gid,\theta}$ that satisfy the access structure (M, ρ) of the CT. Note that, in decryption process, the version values of the attribute components in the user-attribute-keys are matched with the version values of the corresponding components in the ciphertext for the purpose of detecting whether there is any component stateless in the DU's user-attribute-keys. If there is any mismatch, then it will trigger the DU to request a key-update-key for the attribute on which the version number mismatch occurs from CA or the corresponding AA.

- $Decrypt(CT, G_{gid}, \{AK_{gid,\theta}\}_{\theta \in F})$. Suppose that $\mathcal{S}_{gid} = \bigcup_{\theta \in F} \mathcal{S}_{gid,\theta}$ satisfies the access structure (M, ρ) and let $I \subseteq 1, 2, ..., l$ be defined as $I = \{i : \rho(i) \in \mathcal{S}_{gid}\}$. Then, it chooses a set of constants $\{w_i \in \mathbb{Z}_p\}_{i \in I}$ and reconstructs the secret s as $s = \sum_{i \in I} w_i \lambda_i$ if λ_i are valid shares of the secret s according to M. The decryption algorithm first computes:

$$\begin{aligned} K &= \frac{e(C_1, \prod_{\theta \in F} D_{1_\theta})}{\prod_{i \in I} \left(e(C_{i,1}, D_{\rho(i)}) e(C_{i,3}, G_{gid})\right)^{w_i}} \\ &= e(g,g)^{s \sum_{\theta \in F} \alpha_\theta}. \end{aligned}$$

Finally, it recovers the data by computing:

$$m = C_0 / K.$$

Attribute and User Revocation: When an attribute revocation event takes place, the AA that manages the involved attribute runs the $CKeyGen$ algorithm to generate a pair of ciphertext-heal-key and ciphertext-update-key for CS, where the keys are related to the involved attribute. In the case of user revocation event, the involved attribute will be the essential attribute and CA runs the $CKeyGen$ algorithm to generate a pair of ciphertext-heal-key and ciphertext-update-key related the essential attribute. Upon receiving the pair of keys, CS updates all the ciphertexts which include the involved attribute in their access structures by running the $CTUpdate$ algorithm. In the case of user revocation, all ciphertexts are updated since all ciphertexts include the essential attribute in its access structure. Note that CS keeps the new ciphertext-heal-key for subsequent data publications instead of the old one. The AA (or CA) generates a unique key-update-key for the non-revoked DU on demand by running the $UKeyGen$ algorithm. That is to say that the key-update-key generation is triggered by the request from any non-revoked DU. Upon receiving the unique key-update-key from the AA, the DU updates his/her corresponding user-attribute-key by running the $AKUpdate$ algorithm. To simplicity, we suppose that y_θ is the attribute involved in the revocation event.

- $CKeyGen(MK_\theta, y_\theta)$. It first randomly chooses $v'_{y_\theta} \in \mathbb{Z}_p$ as new attribute secret for y_θ, which is different from the previous secret $v'_{y_\theta} \neq v_{y_\theta}$ and increases the attribute's version number by 1 as $ver'_{y_\theta} = ver_{y_\theta} + 1$. Then, it generates a ciphertext-heal-key as follows:

$$hk_{y_\theta}^{h(ver'_{y_\theta})} = \frac{v'_{y_\theta} - t_{y_\theta}}{\delta_\theta}.$$

Next, it generates a ciphertext-update-key as follows:

$$ck_{y_\theta}^{h(ver'_{y_\theta})} = (ck' = \frac{v'_{y_\theta} - \delta_\theta}{\delta_\theta}, ck'' = v_{y_\theta}\beta_\theta - \delta_\theta\beta_\theta).$$

Finally, it updates the master key MK_θ into MK'_θ by replacing the old attribute secret v_{y_θ} for y_θ with new one v'_{y_θ} with the increased version number ver'_{y_θ}.

- $CTUpdate(CT, ck_{y_\theta}^{h(ver'_{y_\theta})})$. It updates only the corresponding component associated with y_θ and sets its version value as the version value associated with $ck_{y_\theta}^{h(ver'_{y_\theta})}$ as:

$$\begin{aligned} CT' = (C_0, C_1, \forall i : \{ &\text{ if } \rho(i) \neq y_\theta : \\ &h(ver(\rho(i)), C_{i,1}, C_{i,2}, C_{i,3}, \\ &\text{if } \rho(i) = y_\theta : h(ver'(\rho(i)), C_{i,1}, C_{i,2}, \\ &C'_{i,3} = C_{i,3} \cdot (C_{i,1})^{ck'} \cdot (C_{i,2})^{ck''} \}). \end{aligned}$$

- $UKeyGen(MK_\theta, y_\theta, G_{gid})$. It generates a unique key-update-key for the non-revoked DU as follows:

$$uk_{gid,y_\theta}^{h(ver'_{y_\theta})} = (G_{gid})^{v'_{y_\theta}\beta_\theta}.$$

- $AKUpdate(SK_{gid,\theta}, uk_{gid,y_\theta}^{h(ver'_{y_\theta})})$. It only replaces the corresponding component associated with the involved attribute and its version value with the corresponding values in $uk_{gid,y_\theta}^{h(ver'_{y_\theta})}$ as follows:

$$\begin{aligned} SK'_{gid,\theta} = (&D_{1_\theta}, y_\theta \in \mathcal{S}_{gid,\theta} : h(ver'_{y_\theta}), \\ &D'_{y_\theta} = uk_{gid,y_\theta}, \\ &\forall x_\theta \in \mathcal{S}_{gid,\theta} \setminus y_\theta : \{h(ver_{x_\theta}), D_{x_\theta}\}). \end{aligned}$$

4.3 Correctness

If the attribute set $\mathcal{S}$ satisfies the access structure, we have that $\sum_{i \in I} \lambda_i w_i = s$. Therefore:

$$\begin{aligned} K &= \frac{e(C_1, \prod_{\theta \in F} D_{1_\theta})}{\prod_{i \in I} \left(e(C_{i,1}, D_{\rho(i)}) e(C_{i,3}, G_{gid})\right)^{w_i}} \\ &= \frac{e(g^s, \prod_{\theta \in F} g^{\alpha_\theta} g^{u_{id}\alpha_\theta})}{\prod_{i \in I} \left(e(g^{r_i}, g^{u_{id} v_{x_\theta} \beta_\theta}) e((\prod_{\theta \in F} (g^{a_\theta})^{\lambda_i} g^{-v_{x_\theta}\beta_\theta r_i}, g^{u_{id}})\right)^{w_i}} \end{aligned}$$

Table 2: Security property comparison

Properties	[36]	[20]	[10]	Our
Collusion Resistance (type-A)	No	No	Yes	Yes
Collusion Resistance (type-B)	Yes	No	Yes	Yes
Forward Secrecy	Yes	No	No	Yes
Backward Secrecy	Yes	No	No	Yes

$$= \frac{e(g^s, g^{\sum_{\theta\in F}\alpha_\theta} g^{u_{id}\sum_{\theta\in F} a_\theta})}{\prod_{i\in I}\left(e(g^{r_i}, g^{u_{id}v_{x_\theta}\beta_\theta}) e((g^{\lambda_i\sum_{\theta\in F} a_\theta} g^{-v_{x_\theta}\beta_\theta r_i}, g^{u_{id}})\right)^{w_i}}$$
$$= \frac{e(g,g)^{s\sum_{\theta\in F}\alpha_\theta} e(g,g)^{u_{id}s\sum_{\theta\in F} a_\theta}}{\prod_{i\in I}\left(e(g,g)^{r_i u_{id} v_{x_\theta}\beta_\theta} e(g,g)^{\lambda_i u_{id}\sum_{\theta\in F} a_\theta} e(g,g)^{-v_{x_\theta}\beta_\theta r_i u_{id})}\right)^{w_i}}$$
$$= \frac{e(g,g)^{s\sum_{\theta\in F}\alpha_\theta} e(g,g)^{u_{id}s\sum_{\theta\in F} a_\theta}}{\prod_{i\in I}\left(e(g,g)^{\lambda_i u_{id}\sum_{\theta\in F} a_\theta}\right)^{w_i}}$$
$$= \frac{e(g,g)^{s\sum_{\theta\in F}\alpha_\theta} e(g,g)^{u_{id}s\sum_{\theta\in F} a_\theta}}{e(g,g)^{u_{id}\sum_{\theta\in F} a_\theta \sum_{i\in I}\lambda_i w_i}}$$
$$= \frac{e(g,g)^{s\sum_{\theta\in F}\alpha_\theta} e(g,g)^{u_{id}s\sum_{\theta\in F} a_\theta}}{e(g,g)^{u_{id}s\sum_{\theta\in F} a_\theta}}$$
$$= e(g,g)^{s\sum_{\theta\in F}\alpha_\theta}$$

Then, the data m is recovered as:

$$m = \frac{C_0}{K} = \frac{m\prod_{\theta\in F}(e(g,g)^{\alpha_\theta})^s}{e(g,g)^{s\sum_{\theta\in F_A}\alpha_\theta}} = \frac{me(g,g)^{s\sum_{\theta\in F}\alpha_\theta}}{e(g,g)^{s\sum_{\theta\in F}\alpha_\theta}}$$

5 Analysis of Our System

5.1 Security Analysis

As can be seen in Table 1, our scheme is the only scheme that supports both user revocation and attribute revocation. Therefore, we first give a comparison between the proposed scheme and the existing attribute-revocable MA-CP-ABE schemes [10, 20, 36] in terms of the security properties defined in Section 3.3.

We will prove the following theorem regarding the security of our Key-Escrow-Free Multi-Authority CP-ABE Scheme with Dual-Revocation.

Theorem 1. *If the decisional q-parallel BDHE assumption holds then all PPT adversary with challenge matrix of size $l^* \times n^*$, where $l^*, n^* \leq q$, have negligible advantage in selectively breaking our system.*

Proof. To prove the theorem, we will assume that there exists a PPT attacker $\mathcal{A}$ which has a non-negligible advantage $Adv_{\mathcal{A}}$ in selectively breaking our system. Moreover, we suppose the $\mathcal{A}$ chooses a challenge matrix (M^*, ρ^*) where both dimensions are at most q. Using this attacker, we will build a PPT simulator $\mathcal{B}$ that plays the decisional q-parallel BDHE problem with non-negligible advantage. Since we assume that the attacker can corrupt at most $|\mathcal{N}-1|$ authorities, the attacker cannot know the the master key of only non-corrupted authority. It implies that the security proof of our scheme can be proved under q-parallel BDHE assumption. Thus, in security proof, we will use only the terms of the q-parallel BDHE assumption instead of all the different g^{a_θ} terms of the relevant authorities.

Global Setup. The $\mathcal{B}$ sends the global parameters GP of the system to the $\mathcal{A}$.

Init. The $\mathcal{B}$ takes the q-parallel BDHE challenge $\vec{y}, T$. It receives a set of corrupted authorities $\mathcal{C} \subseteq \mathcal{N}$ and a challenge access structure (M^*, ρ^*) along with the target version numbers $\{ver^*_{x_\theta}\}_{x_\theta=\rho^*(i)}$ from the $\mathcal{A}$. We have that M^* is a $l^* \times n^*$ matrix, where $l^*, n^* < q$.

Authority Setup. For each non-corrupted authority $A_\theta \in \mathcal{N} - \mathcal{C}$, the $\mathcal{B}$ picks random exponents $\alpha'_\theta, \beta_\theta, \delta_\theta \in \mathbb{Z}_p$ and implicitly sets the master key of the authority to be $\alpha_\theta = \alpha'_\theta + a^{q+1}$ by letting

$$e(g,g)^{\alpha_\theta} = e(g^a, g^{a^q}) e(g,g)^{\alpha'_\theta}.$$

We describe how the simulator programs the public attribute keys $\{T_{x_\theta}\}_{x_\theta\in U}$, where $U = \bigcup_{\theta\in\mathcal{N}_A-\mathcal{C}_A} U_\theta$. For each $x_\theta \in U$, it chooses random values $z_{x_\theta}, v_{x_\theta} \in \mathbb{Z}_p$. Let X denote the set of indices i, such that $\rho^*(i) = x_\theta$. The simulator programs T_{x_θ} as:

$$T_{x_\theta} = \left(g^{z_{x_\theta}} \prod_{i\in X} g^{aM^*_{i,1}/b_i} \cdot g^{a^2 M^*_{i,2}/b_i} \cdots g^{a^{n^*} M^*_{i,n^*}/b_i}\right)^{\beta_\theta}$$

Note that if $X = \emptyset$ then we have $T_{x_\theta} = g^{z_{x_\theta}\beta_\theta}$. Also note that the public attribute keys are distributes randomly due to the $g^{z_{x_\theta}}$ and β_θ values.

Therefore, the public parameters of each non-corrupted authority A_θ are:

$$PP_\theta = \left(e(g,g)^{\alpha'_\theta}, g^{\delta_\theta\beta_\theta}, \{h(1), T_{x_\theta}\}_{x_\theta\in U_\theta}\right).$$

Moreover, the simulator selects randomly $ver^*_{x_\theta} - 1$ number of values $\{v'_{x_\theta}\}_{ver=2,\ldots,ver^*_{x_\theta}} \in \mathbb{Z}_p$ for each $x_\theta \in U_A$ and keeps them. Then, the simulator can calculate a set of ciphertext-heal-keys for all attributes that appear in the challenge access structure, where each ciphertext-heal-key can be calculated on its target version number by using v'_{x_θ} for version number $ver^*_{x_\theta}$.

$$hk_{x_\theta}^{h(ver^*_{x_\theta})} = \frac{v'_{x_\theta} - z_{x_\theta}}{\delta_\theta}.$$

Phase 1. In this phase, the simulator answers secret key and key-update key queries from the adversary. Without loss of generality, suppose the simulator is given a secret key query for $\mathcal{S}_{gid}$ for the version numbers of all the attributes in $\mathcal{S}_{gid}$ is 1, where $\mathcal{S}_{gid} = \bigcup_{\theta\in\mathcal{N}-\mathcal{C}} \mathcal{S}_{gid,\theta}$ is a set of attributes belonging to several non-corrupted authorities. Suppose $\mathcal{S}_{gid}$ does not satisfy M^* in combination with any keys that can be obtained from corrupted authorities.

Since $\mathcal{S}_{gid}$ does not satisfy $(M^*, \rho^*(i))$, there exist a vector $\vec{w} = (w_1, ..., w_{n^*}) \in \mathbb{Z}_p^{n^*}$ such that $w_1 = -1$ and $\vec{w} \cdot M^*_i = 0$ for all $i \in I = i | i \in [l] \wedge \rho^*(i) \in \mathcal{S}_{gid}$.

Then, the simulator chooses a random $u'_{id} \in \mathbb{Z}_p$ and implicitly sets u_{id} as

$$\begin{aligned} u_{id} &= u'_{id} + w_1 a^q + w_2 a^{q-1} + \cdots + w_{n^*} a^{q-n^*+1} \\ &= u'_{id} + \sum_{i=1\ldots n^*} w_i a^{q+1-i} \end{aligned}$$

This is properly distributed due to the u'_{id}. Then it calculates:

$$G_{gid} = g^{u_i d} = g^{u'_{id}} \prod_{i=1\ldots n^*} \left(g^{a^{q+1-i}}\right)^{w_i}$$

Now we describe how to create an user-attribute-key for every $\mathcal{S}_{gid,\theta} \subseteq \mathcal{S}_{gid}$. It has to create the user-attribute-keys separately since each subset of attributes $\mathcal{S}_{gid,\theta} \subseteq \mathcal{S}_{gid}$ belongs to different authority A_θ. Each user-attribute key consists of $D_{1_\theta}, \{D_{x_\theta}\}_{x_\theta \in \mathcal{S}_{gid,\theta}}$ components. Note that all attribute components in the user-attribute-keys are created on the version number 1. Then using the suitable terms from the assumption, the simulator calculates:

$$\begin{aligned} D_{1_\theta} &= g^{\alpha_\theta} g^{a u_{id}} = g^{a^{q+1}} g^{\alpha'_\theta} g^{a u'_{id}} \prod_{i=1\ldots n^*} g^{w_i a^{q+2-1}} \\ &= g^{\alpha'_\theta} (g^a)^{u'_{id}} \prod_{i=2\ldots n^*} \left(g^{a^{q+2-i}}\right)^{w_i} \end{aligned}$$

Now, for all $x_\theta \in \mathcal{S}_{gid,\theta}$, the simulator has to compute D_{x_θ} component. If $x_\theta \in \mathcal{S}_{gid,\theta}$ is not used in the access structure, for which there is no i such that $\rho^*(i) = x_\theta$, the the simulator can simply calculate:

$$D_{x_\theta} = \left(G_{gid}\right)^{v_{x_\theta}\beta_\theta}$$

If $x_\theta \in \mathcal{S}_{gid,\theta}$ is used in the access structure, then the simulator computes D_{x_θ} as follows:

$$D_{x_\theta} = \left(G_{gid}\right)^{v_{x_\theta}\beta_\theta} \prod_{i\in X} \prod_{j=1,\ldots,n^*} \Bigg(g^{(a^j/b_i)r} \cdot \prod_{k=1,\ldots,n^*, k\neq j} (g^{a^{q+1+j-k}/b_i})^{w_k}\Bigg)^{M^*_{i,j}\beta_\theta}$$

Towards key-update key queries, the simulator returns a key-update-key for the given attribute and on the given version number. Suppose that the simulator is given a key-update-key for (x_θ, G_{gid}) along with a version number t_{x_θ}, where $x_\theta \in \mathcal{S}_{gid}$ and $2 \leq t_{x_\theta} \leq ver^*_{x_\theta}$. By using v'_{x_θ} for the version number t_{x_θ}, the simulator calculates a key-update key $uk^{h(t_{x_\theta})}_{gid,x_\theta}$ for each x_θ in the similar way of calculating D_{x_θ} depending on whether x_θ is in the challenge access structure or not. If x_θ is not used in the access structure, the key-update key can calculate as:

$$uk^{h(t_{x_\theta})}_{gid,x_\theta} = \left(G_{gid}\right)^{v'_{x_\theta}\beta_\theta}$$

If x_θ is used in the access structure, then the key-update key can calculate as:

$$uk^{h(t_{x_\theta})}_{gid,x_\theta} = \left(G_{gid}\right)^{v'_{x_\theta}\beta_\theta} \prod_{i\in X} \prod_{j=1,\ldots,n^*} \Bigg(g^{(a^j/b_i)r} \cdot \prod_{k=1,\ldots,n^*, k\neq j} (g^{a^{q+1+j-k}/b_i})^{w_k}\Bigg)^{\beta_\theta M*_{i,j}}$$

Challenge. In this phase, we build the challenge ciphertext. The adversary gives two messages m_0, m_1 with equal length to the simulator, The simulator flips a coin b and contructs

$$C_0 = m_b \cdot T \cdot \prod_{\theta\in F_A} e(g, g^s)^{\alpha'_\theta}$$

and $C_1 = g^s$, where T is the challenge term and g^s is the corresponding term of the assumption.

The tricky part is to simulate the $C_{i,3}$ values since this contains the terms that must be cancelled out. However, the simulator can choose the secret splitting, such that these can be cancelled out. The simulator sets

$$\vec{v} = (s, sa + y'_2, sa^2 + y'_3, ..., sa^{n*-1} + y'_{n^*}) \in \mathbb{Z}_p^{n^*},$$

where $y'_2, ..., y'_{n^*} \in \mathbb{Z}_p$. We see that the secret s and the vector $\vec{v}$ are properly distributed, since s is information theoretically hidden from $\mathcal{A}$ and y'_i's are picked uniformly at random. As as result, since $\lambda_i = \vec{v} \cdot M^*_i$, we can construct the share of the secret as:

$$\lambda_i = \sum_{j=1\ldots n^*} M^*_{i,j} s a^{j-1} + \sum_{j=2,\ldots,n^*} M^*_{i,j} y'_j.$$

For each row, the simulator chooses a random r'_i and implicitly sets $r_i = r'_i + sb_i$. For $i = 1, ..., n^*$, we define R_i as the set of all $k \neq i$ such that $\rho^*(i) = \rho^*(k)$. That is the set of all other row indices that have the same attributes as row i. Using the above, the $\mathcal{B}$ calculates

$$\begin{aligned} C_{i,1} &= g^{r_i} g^{s b_i} \\ C_{i,2} &= (g^{-r_i} \cdot g^{-s b_i})^{\delta_\theta \beta_\theta} \\ \tilde{C}_{i,3} &= T^{-r'_i}_{\rho^*(i)} \Bigg(\prod_{j=2,\ldots,n^*} (g^a)^{M^*_{i,j} y'_j}\Bigg) (g^{b_i s})^{-z_{\rho^*(i)\beta_\theta}} \cdot \\ &\quad \Bigg(\prod_{k\in R_i} \prod_{j=1,\ldots,n^*} (g^{a^j \cdot s \cdot (b_i/b_k)})^{M^*_{k,j}\beta_\theta}\Bigg). \end{aligned}$$

Since $\tilde{C}_{i,3}$ is simulated in a ill-formed way, the simulator also must convert $\tilde{C}_{i,3}$ to a well-formed $C_{i,3}$ by using a set of ciphertext-heal keys $\{hk^{ver^*_{x_\theta}}_{x_\theta}\}$, where $x_\theta = \rho^*(i)$.

$$C_{i,3} = \tilde{C}_{i,3} \cdot (C_{i,2})^{hk^{ver^*_{\rho^*(i)}}_{\rho^*(i)}}.$$

Therefore, the $\mathcal{B}$ hands over the ciphertext $CT = ((M^*, \rho^*), C_0, C_1, \{C_{i,1}, C_{i,2}, C_{i,3}\})$ to the $\mathcal{A}$.

Phase 2. The same as Phase 1.

Guess. The adversary eventually outputs a guess $b^{'}$ for the challenge bit. If $b^{'} = b$ the simulator outputs 0 to guess that $T = e(g,g)^{a^{q+1}s}$. Otherwise, it outputs 1 to indicate that it believes T is a random group element in $\mathbb{G}_1$.

When T is a tuple the simulator $\mathcal{B}$ gives a perfect simulation so we have that

$$Pr\left[\mathcal{B}(\vec{y}, T = e(g,g)^{a^{q+1}s}) = 0\right] = \frac{1}{2} + Adv_{\mathcal{A}}.$$

When T is a random group element the message m_b is completely hidden from the adversary and we have that

$$Pr\left[\mathcal{B}(\vec{y}, T = R) = 0\right] = \frac{1}{2}.$$

Therefore, the $\mathcal{B}$ can play the decisional q-parallel BDHE game with non-negligible advantage.

□

5.2 Performance Analysis

As can bee seen in Table 1, our scheme is the only scheme that supports both user revocation and attribute revocation, whereas MA-CP-ABE schemes in [10, 20, 36] only support attribute revocation. However, the schemes in [10, 20] utilize composite order groups. The group operations in composite order groups –group exponentiations and group pairings– are several orders of magnitude slower than those in prime order groups. More information on the comparison between prime and composite groups can be found [12, 28]. Thus, we exclude those schemes that are set in composite order groups from performance comparisons, and we give performance analysis of our scheme by comparing with Yang et al.'s scheme [36] that utilizes prime order groups as the same as our scheme does. The comparison is made in terms of storage overhead, communication cost, and computational efficiency. Now, we describe the notations used in the comparisons. Let $|G|$, $|G_T|$, and $|Z_p|$ denote the size of an element in the source group, the target group, and the field Z_p, respectively. In addition, let u and a be the total number of users and attributes in the system, respectively. Moreover, let s, l, and f denote the number of attributes a user possess, the number of attributes in an access structure and the number of matched attributes in a decryption, respectively. Besides, r, k, a_k, and n_k denotes the number of AAs involved in an encryption, the number of AAs in the system, the number of attributes handled by the AA and the number of AAs from that a user holds his/her secret key, respectively. Furthermore, let n_x and n_c be the number of users who possess an attribute x and the number of ciphertexts including an attribute in its access structure, respectively. Finally, let E, E_T, and P denote exponentiation in the source group, exponentiation in the target group, and pairing operations, respectively.

Storage overhead. From Table 3, we can see the storage overhead on each entity in the schemes, namely cloud server (CS), certificate authority (CA), attribute authority (AA), data owner (DO), and data user (DU). In both schemes, the main storage overhead on CS comes from ciphertexts, but our scheme brings additional overhead on CS due to ciphertext-heal-keys. However, it helps to reduce the communication cost between DOs and the AA on each revocation event; this communication cost exists in the scheme [36]. As for the overhead on CA, the scheme [36] brings a considerable amount of overhead than our scheme does because of a pair of global public and secret keys of each user in the system. In both schemes, the main storage overhead on each AA comes from the attribute secrets of all the attributes managed by the AA. Besides that, the scheme in [36] causes a great deal of extra overhead on each AA because this scheme requires each AA to keep a pair of global public and secret keys of each user in the system. While only the public parameters and public attribute keys of the AAs involved in the encryption contribute the main storage overhead on a DO in the scheme [36], in our scheme the public parameters and public attribute keys of the AAs involved in encryption including CA contribute the main storage overhead on a DO. However, the total overhead on a DO in our scheme is less than that in [36]. In both schemes, the storage overhead on a DU comes from a secret key that consists of the DU's global key and keys related to his/her attribute set from different AAs. However, in our scheme the components of secret key associated to the essential attribute brings slightly additional overhead on a DU than that in [36]. In Table 3, the size of a secret key is expressed by all components possessed by a DU who has s number of attributes in the system.

Communication cost. In Table 4, we discuss the communication cost between the entities in the schemes. To distinguish the communication cost at initialization phase from that at revocation phase, we use the notations **I:** and **R:**, respectively. In the scheme [36], CA sends a pair of global public and secret keys of each user in the system to each AA; it brings considerable higher communication cost between CA and an AA. In both schemes, sending user-central-key/global-key to a DU results in the communication cost between CA and the DU, and the cost in our scheme is slightly higher than that in [36]. Due to the user revocation mechanism –the scheme [36] does not support– our scheme brings the communication cost between CA and CS, DU regarding the essential attribute that CA handles. In both schemes, the communication cost between CS and a DO, a DU mainly comes from the transmission of a ciphertext, and it is higher in the scheme [36] than that in our scheme. The communication cost between an AA

Table 3: Comparison of storage overhead

Ents	The scheme in [36]	Our scheme
CS	$(2+4l)\|G\|+\|G_T\|$	$(4+3l)\|G\|+\|G_T\|; a\|Z_p\|$
CA	$u\|Z_p\|+u\|G\|$	$6\|Z_p\|$
AA	$(2+a_k+u)\|Z_p\|+u\|G\|$	$(4+2a_k)\|Z_p\|$
DO	$((2+2a_k)\|G\|+\|G_T\|)r$	$((5+a_k)\|G\|+2\|G_T\|)r$
DU	$\|Z_p\|+(1+s+2n_k)\|G\|$	$(4+s+2n_k)\|G\|$

and a DO comes from the public parameters and public attribute keys managed by the AA, which is lower in our scheme. In both schemes, a secret key associated with a set of attributes from the AA contributes communication cost between the AA and the DU, whereas key-update-key brings the cost upon each attribute revocation event. This cost is nearly close in both schemes.

Computational efficiency. We present a theoretical analysis of the computational efficiency of our scheme in terms of key generation, encryption, decryption, ciphertext update and update key generation, as shown in Table 5 by comparing with the scheme [36]. The most time-consuming operations in ABE settings are group exponentiation and pairing. Therefore, we present only those operations in the analysis. In both schemes, the cost of key generation depends on the number of attributes the DU possess from the AA. The cost of update key generation upon each attribute revocation event in the scheme [36] depends on the number of non-revoked users for the involved attribute. In our scheme, the computational cost coming from update key generation is less than that in [36] since a key-update-key for a non-revoked DU is generated on demand in our scheme. In both schemes, the computational cost on a DO is occurred by encryption, and the cost linearly increases with the number of attributes expressing the access structure in the ciphertext. In both schemes, the cost by decryption depends on the number of attributes satisfying the access structure in the ciphertext, and the computation cost occurred by ciphertext update upon each attribute revocation event increases linearly as the number of ciphertexts including the involved attribute in their access structures. As can be seen from Table 5, our scheme is more efficient than Yang et al.'s scheme [36].

Table 4: Comparison of communication cost

Comp's	The scheme in [36]	Our scheme
CA&AA	$u\|Z_p\|+u\|G\|$	-
CA&CS	-	**I:** $\|Z_p\|$; **R:** $3\|Z_p\|$
CA&DU	$\|Z_p\|+\|G\|$	$5\|G\|$
CA&DO	-	$2\|G\|+\|G_T\|$
AA&CS	**R:** $2\|Z_p\|$	**I:** $a\|Z_p\|$; **R:** $3\|Z_p\|$
AA&DU	**I:** $(2+s_k)\|G\|$;**R:** $\|G\|$	**I:** $(2+s_k)\|G\|$; **R:** $\|G\|$
AA&DO	$(2+2n_k)\|G\|+\|G_T\|$	$\|Z_p\|+(1+n_k)\|G\|+\|G_T\|$
DO&CS	$(1+4l)\|G\|+\|G_T\|$	$(5+3l)\|G\|+\|G_T\|$
CS&DU	$(1+4l)\|G\|+\|G_T\|$	$(5+3l)\|G\|+\|G_T\|$

Table 5: Comparison of computational efficiency

Components	The scheme in [36]	Our scheme
KeyGeneration	$(3+2s)E$	$(1+s)E$
Encryption	$(2+5l)E+E_T$	$(1+4l)E+ET$
Decryption	$fE_T+(4f+2k)P$	$fE_T+(2f+1)P$
CTUpdate	$2n_cE$	$2n_cE$
KeyUpdate	n_xE	$\ll n_xE$

6 Conclusion

The existing MA-CP-ABE schemes cannot be applied to collaborative cloud data storage as data access control due to the key escrow problem and the absence of dual revocation mechanism. In this paper, we proposed a Key-Escrow-Free MA-CP-ABE scheme with Dual Revocation to overcome the issues existing in the previous works. The proposed scheme is key escrow free due to the fact that any access structure is formed in a way that attributes from an individual authority cannot satisfy it. In addition, dual revocation mechanism in the proposed scheme guarantees both forward and backward secrecy, and it resists any potential collusion attacks as well. Moreover, our scheme is set in groups of prime order and proved in the standard model under a standard assumption. Based on the comparison of MA-CP-ABE schemes regarding their functionalities, the comparison of attribute-revocable MA-CP-ABE schemes regarding security requirements for revocation, and performance analysis, we demonstrate that our proposed scheme is more suitable to be applied to collaborative cloud data storage as a secure and scalable access control.

Acknowledgments

This study was supported by International (Regional) Joint Research Project of China National Science Foundation under Grant No.61520106007. The authors gratefully acknowledge the anonymous reviewers for their valuable comments.

References

[1] B. Balusamy, P. V. Krishna, G. S. T. Arasi, and V. Chang, "A secured access control technique for cloud computing environment using attribute based hierarchical structure and token granting system," *International Journal of Network Security*, vol. 19, no. 4, pp. 559–572, 2017.

[2] M. Bayat, M. Aref, "An attribute based key agreement protocol resilient to KCI attack," *International*

Journal of Electronics and Information Engineering, vol. 2, no. 1, pp. 10–20, 2015.

[3] A. Beimel, *Secure Schemes for Secret Sharing and Key Distribution*. Technion, Haifa, Isreal: Ph.D Thesis, Isreal Institute of Technology, 1996.

[4] J. Bethencourt, A. Sahai, and B. Waters, "Ciphertext-policy attribute-based encryption," in *Proceedings of 2007 IEEE Symposium on Security and Privacy (SP'07)*, pp. 321–334, Berkeley, CA, May 2007.

[5] M. Blaze, G. Bleumer, and M. Strauss, "Divertible protocols and atomic proxy cryptography," in *Proceedings of International Conference on the Theory and Application of Cryptographic Techniques*, pp. 127–144, Espoo, Finland, May-June 1998.

[6] M. Chase, "Multi-authority attribute based encryption," in *Proceedings of 4th Theory of Cryptography Conference (TCC'07)*, pp. 515–534, Amsterdam, The Netherlands, Feb. 2007.

[7] L. Cheung and C. Newport, "Provably secure ciphertext policy abe," in *Proceedings of the 14th ACM conference on Computer and communications security (CCS07)*, pp. 456–465, Alexandria, Virginia, USA, Oct. 2007.

[8] S. Chow, "A framework of multi-authority attribute-based encryption with outsourcing and revocation," in *Proceedings of the 21st ACM on Symposium on Access Control Models and Technologies (SACMAT'16)*, pp. 215–226, Shanghai, China, June 2016.

[9] P. S. Chung, C. W. Liu, and M. S. Hwang, "A study of attribute-based proxy re-encryption scheme in cloud environments," *International Journal of Network Security*, vol. 16, no. 1, pp. 1-13, 2014.

[10] H. Cui and R. H. Deng, "Revocable and decentralized attribute-based encryption," *The Computer Journal*, vol. 59, no. 8, pp. 1220–1235, 2016.

[11] V. Goyal, A. Jain, O. Pandey, and A. Sahai, "Bounded ciphertext policy attribute based encryption," in *Proceedings of the 35th International Colloquium, ICALP 2008*, pp. 579–591, Reykjavik, Iceland, July 2008.

[12] A. Guillevic, "Comparing the pairing efficiency over composite-order and prime-order elliptic curves," in *Proceedings of the 11th International Conference on Applied Cryptography and Network Security (ACNS'13)*, pp. 357–372, 2013.

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

[14] L. Ibraimi, M. Petkovic, S. Nikova, P. Hartel, and W. Jonker, "Mediated ciphertext-policy attribute-based encryption and its application," in *10th International Workshop on Information Security Applications (WISA'09)*, pp. 309–323, 2009.

[15] L. Ibraimi, Q. Tang, P. Hartel, and W. Jonker, "Efficient and provable secure ciphertext-policy attribute-based encryption schemes," in *5th International Conference on Information Security Practice and Experience (ISPEC'09)*, pp. 1–12, Xi'an, China, Apr. 2009.

[16] S. Jahid, P. Mittal, and N. Borisov, "Easier: Encryption-based access control in social networks with efficient revocation," in *Proceedings of the 6th ACM Symposium on Information, Computer and Communications Security (ASIACCS'11)*, pp. 411–415, Hong Kong, China, Mar. 2011.

[17] C. C. Lee, P. S. Chung, M. S. Hwang, "A survey on attribute-based encryption schemes of access control in cloud environments," *International Journal of Network Security*, vol. 15, no. 4, pp. 231–240, July 2013.

[18] A. Lewko and B. Waters, "Decentralizing attribute-based encryption," in *Proceedings of 30th Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pp. 568–588, Tallinn, Estonia, May 2011.

[19] K. Li and H. Ma, "Outsourcing decryption of multi-authority ABE ciphertexts," *International Journal of Network Security*, vol. 16, pp. 286–294, 2014.

[20] Q. Li, J. Ma, R. Li, X. Liu, J. Xiong, and D. Chen, "Secure, efficient and revocable multi-authority access control system in cloud storage," *Computers & Security*, vol. 59, pp. 45–59, 2016.

[21] X. Liang, Z. Cao, H. Lin, and D. Xing, "Provably secure and efficient bounded ciphertext policy attribute based encryption," in *Proceedings of the 4th International Symposium on Information, Computer, and Communications Security (ASIACCS'09)*, pp. 343–452, Sydney, NSW, Australia, Mar. 2009.

[22] X. Liang, X. Lin R. Lu, and X. Shen, *Ciphertext Policy Attribute Based Encryption with Efficient Revocation*, Waterloo, Ontario, Canada: Technique Report, University of Waterloo, 2011.

[23] C. W. Liu, W. F. Hsien, C. C. Yang, and M. S. Hwang, "A survey of attribute-based access control with user revocation in cloud data storage", *International Journal of Network Security*, vol. 18, no. 5, pp. 900–916, 2016.

[24] Z. Liu, Z. Cao, Q. Huang, D. S. Wong, and T. H. Yuen, "Fully secure multi-authority ciphertext-policy attribute-based encryption without random oracles," in *Proceedings of 16th European Symposium on Research in Computer Security*, pp. 278–297, Leuven, Belgium, Sept. 2011.

[25] H. Ma, T. Peng, Z. Liu, "Directly revocable and verifiable key-policy attribute-based encryption for large universe," *International Journal of Network Security*, vol. 19, no. 2, pp. 272–284, 2017.

[26] S. Muller, S. Katzenbeisser, and C. Eckert, "On multi-authority ciphertext-policy attribute-based encryption," *Bulletin of the Korean Mathematical Society*, vol. 46, no. 4, pp. 803–819, 2009.

[27] M. Pirretti, P. Traynor, and P. Mcdaniel, "Secure attribute-based systems," in *Proceedings of the 13th ACM conference on Computer and communications*

security (CCS06), pp. 99–112, Alexandria, Virginia, USA, Oct. 2006.

[28] Y. Rouselakis and B. Waters, "Efficient statically-secure large-universe multi-authority attribute-based encryption," in *Proceedings of 19th International Conference (FC'15)*, pp. 315–332, San Juan, Puerto Rico, Jan. 2015.

[29] A. Sahai, H. Seyalioglu, and B. Waters, "Dynamic credentials and ciphertext delegation for attribute-based encryption," in *Proceedings of the 32nd Annual Cryptology Conference on Advances in Cryptology (CRYPTO'12)*, pp. 199–217, 2012.

[30] Y. Tian, Y. Peng, G. Gao, X. Peng, "Role-based access control for body area networks using attribute-based encryption in cloud storage," *International Journal of Network Security*, vol. 19, no. 5, pp. 720–726, 2017.

[31] L. Touati and Y. Challal, "Batch-based cp-abe with attribute revocation mechanism for the internet of things," in *Proceedings of International Conference on Computing, Networking and Communications (ICNC'15)*, pp. 1044–1049, Garden Grove, CA, Feb. 2015.

[32] G. Wang, Q. Liu, J. Wu, and M. Guo, "Hierarchical attribute-based encryption and scalable user revocation for sharing data in cloud servers," *Computers & Security*, vol. 30, no. 5, pp. 320–331, 2011.

[33] B. Waters, "Ciphertext-policy attribute-based encryption: an expressive, efficient, and provably secure realization," in *Proceedings of the 14th International Conference on Practice and Theory in Public Key Cryptography Conference on Public Key Cryptography (PKC'11)*, pp. 53–70, Taormina, Italy, Mar. 2011.

[34] Z. Xu and K. Martin, "Dynamic user revocation and key refreshing for attribute-based encryption in cloud storage," in *Proceedings of the 11th IEEE International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom'12)*, pp. 844–849, Liverpool, UK, June 2012.

[35] K. Xue, Y. Xue, J. Hong, W. Li, H. Yue, D. S. Wei, and P. Hong, "RAAC: Robust and auditable access control with multiple attribute authorities for public cloud storage," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 4, pp. 953–967, 2017.

[36] K. Yang and X. Jia, "Expressive, efficient, and revocable data access control for multi-authority cloud storage," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 7, pp. 1735–1744, 2014.

[37] K. Yang, X. Jia, and K. Ren, "Attribute-based fine-grained access control with efficient revocation in cloud storage systems," in *Proceedings of the 8th ACM SIGSAC Symposium on Information, Computer and Communications Security (ASIACCS'13)*, pp. 523–528, Hanzhou, China, May 2013.

[38] S. Yu, C. Wang, K. Ren, and W. Lou, "Attribute based data sharing with attribute revocation," in *Proceedings of the 5th ACM Symposium on Information, Computer and Communications Security (ASIACCS'10)*, pp. 261–270, Beijing, China, Apr. 2010.

[39] Y. Zhao, P. Fan, H. Cai, Z. Qin and H. Xiong, "Attribute-based encryption with non-monotonic access structures supporting fine-grained attribute revocation in M-healthcare," *International Journal of Network Security*, vol. 19, no. 6, pp. 1044–1052, 2017.

Biography

Nyamsuren Vaanchig is a Ph.D. candidate at the School of Information and Software Engineering, University of Electronic Science and Technology of China (UESTC). She received her M.Sc. degree from the Mongolian State University of Education (MSUE) in 2007. Her research interests include information security, cryptography, and cloud computing.

Hu Xiong is an associate professor at the School of Information and Software Engineering, UESTC. He received his Ph.D. degree from the School of Computer Science and Engineering, UESTC in 2009. His research interests include cryptographic protocols and network security.

Wei Chen is an associate professor at the School of Information and Software Engineering, UESTC. He received his Ph.D. degree from UESTC in 2010. His research interests include information security and software assurance.

Zhiguang Qin is a professor and the dean of School of Information and Software Engineering, UESTC. He received his Ph.D. degree from UESTC in 1996. He has engaged in distributed computing, information security, and electronic commerce research.

A Secure and Efficient Image Encryption Algorithm Based on DNA Coding and Spatiotemporal Chaos

Xiaodong Li[1,2], Cailan Zhou[2], Ning Xu[1,2]
(Corresponding author: Ning Xu)

Hubei Key Laboratory of Transportation Internet of Things, China[1]
School of Computer Science and Technology, Wuhan University of Technology[2]
122, Luoshi Rd., Wuhan, Hubei 430070, P.R. China
(Email: xuning@whut.edu.cn)
(Received Mar. 19, 2017; revised and accepted June 25, 2017)

Abstract

In this paper, an image encryption algorithm based on DNA random coding and random operation combined with chaotic map is proposed. In order to produce sequences with more chaotic characteristics, a new spatiotemporal chaotic system is proposed by employing the Tent-Sine system (TSS) in the coupled map lattice (CML). SHA-256 hash of the plain image is used to generate secret keys. The Lorenz Map, Logistic Map and TSS are applied to generate all parameters the proposed algorithm needs. In order to get the high randomness and overcome the limitations of DNA computing rules, encode the every rows of original image and key image with DNA rules respectively, which are randomly selected from eight encoding rules. Then, apply encoded original image to execute DNA operations with encoded key image row by row to obtain the transitional image and the one of the four DNA operations of every row is determined by logistic map; Finally, randomly decode the transitional image to gain the eventual encrypted image. experimental results demonstrate that the proposed algorithm have ability to resist typical attacks.

Keywords: Coupled Map Lattice; DNA Coding; Lorenz Chaotic Map; SHA-256; Tent-Sine System

1 Introduction

With the rapid development of the Internet, the security of the image constantly attract people's attention. in order to protect the image information are not to be disclosed, many image encryption algorithms are proposed and implemented [4, 16, 22]. Because of bulky data capacity, high redundancy and strong correlations among adjacent pixels, typical image encryption algorithms such as RSA [14], DES [10, 40], AES [15, 21] are not competent to encrypt such digital images. Nowadays, many image encryption algorithm have been proposed, such as, image cryptosystem based on chaos [9, 11, 19, 20, 30], DNA computing [6, 33, 34, 35, 38], fractional fourier transform [1, 36], or cellular automata(CA) [7, 18]. Among those, chaos based image cryptosystem have attracted extensive concerns because of a natural and close connection between chaos and cryptography. Such as, sensitive dependence on initial conditions, pseudo-randomness,ergodicity and reproduction are primary features of chaos system, which meets the requirements of encryption. However, digital implementations of chaotic systems will become periodic eventually because of finite precision and temporal discretization can result with the security risks into chaos based cryptosystem. To overcome the short period issue existing in chaos, spatiotemporal chaotic system with longer period has been widely employed in image cryptography [2, 17, 31]. To improve the chaotic property of three maps Logistic, the Logistic-Tent, Logistic-Sine and Tent-Sine systems were developed [39]. In this paper, TSS combined with CML is used to obtain longer period and generate pseudo-random sequences.

Many of the excellent properties of DNA computing have recently been found, for example: large-scale computational parallelism, huge storage space and tiny energy loss. Therefore, the use of DNA complementary rules to encrypt information technology has made great progress. Zhen et al. [37] proposed an image encryption algorithm based on spatiotemporal chaotic system and DNA coding. In this research, logistic and spatiotemporal chaotic system are proposed, the mix DNA coding and eight DNA encode rules will guarantee the efficiency of image confusion and diffusion. Chai et al. [3] proposed an algorithm based on memristive hyperchaotic system, cellular automata (CA) and DNA sequence. In the research, a dynamic DNA encoding scheme is proposed. Two DNA rule matrices for encoding the plain image and

two-dimensional (2D) CA are generated from chaotic sequences, and they are decided by the plain image, hence we can obtain different DNA encoding rules for different plain image. Wang et al. [26] proposed a novel image encryption algorithm based on DNA computing and chaotic map. In this research, a kind of spatiotemporal chaos map, such as coupled map lattice is exploited to confuse the plain image. After encoded the confused image, permute its rows and columns to obtain the encoded cipher image. Hu et al. [8] proposed a novel image encryption scheme which used hyper dimensional chaotic systems and cycle operation for DNA sequence. In this research, the pseudo-random sequence is controlled by a chens hyper-chaos system, a cycle operation for DNA sequences is used to diffuse the pixel values of the image. After carrying out exclusive-OR operation for decoded matrices, and then the cipher image is generated.

However, the current DNA encoding image encryption algorithm still exist problems [13] including: DNA encoding rules are limited, and the limited DNA computation rules, key sensitivity is low, etc. In view of these problems in the proposed encryption algorithm, the algorithm proposed in this paper will combine the DNA computing with the DNA coding rules, Using the excellent characteristics of chaotic map such as randomness to randomly determine the DNA encode rules and DNA operations. We proposed a new DNA XNOR operation that can increase the choice of DNA computing space. SHA-256 hash of the plain image is used to generate secret keys, as long as the original image has a slight change, SHA-256 hash value will make a huge difference, which enhances the sensitivity of the cryptosystem.

The cryptosystem utilizes a 256 bit external secret key K,which is generated by exploited SHA-256 function to original image. We use K to generate the initial values of TSS system and one-dimensional logistic map. Hence the secret keys are extremely related to plain image. the cryptosystem can resist brute force attack, chosen-plaintext attack and chosen-ciphertext attack. TSS is applied to generate key image of the size of M*N*4, then use randomly encoded key image to conduct random DNA operations with encoded three matrices R, G and B components row by row to obtain three encoded DNA transitional images,and DNA operation and DNA encode rules are randomly decided by one-dimensional logistic map. The Lorenz system [12, 24] is used to generate three sequences. These sequences are used to permute three encoded DNA transitional images.

The main contributions of the proposed encryption algorithm are as follows:

1) Exploit the chaotic map randomly determine the DNA encode rules and DNA operations,then executing DNA coding and computing row by row to guarantee the image's encoding rules and operations of each row are randomly selected.

2) A new spatiotemporal chaotic system is constructed by employing the Tent-Sine system (TSS) in the coupled map lattice (CML).

The rest of this paper is organized in the following manners: Section 2 introduce the basic theory of the proposed algorithm. The proposed image encryption method is explained in Section 3. In Section 4, experimental results and security analysis are proposed. Finally the conclusions are drawn in Section 5.

2 Basic Theory

2.1 TSS-based CML

Algebraic implementation of any chaos map could be periodic, but the period of discrete dynamic system such as, the CML is adequately long to ensure cryptosystem security [26]. The CML defined as in Equation (1):

$$x_{j+1} = (1-e)F(x_{j+1}(i) + eF(x_j)), \tag{1}$$

where i=1,2,...,n is the time variable, j=1,2,...,l is the spatial variable, l is the lattice length (In the proposed image cryptosystem, l=3. $e \in (0,1)$ expresses the coupling factor, $x_j(i)$ expresses the variate for the jth lattice site at time i. Moreover, the periodic boundary of the CML is $x_1(i) = x_{l+1}(i)$. In order to generate extremely random sequences, TSS is adopted as the map $F(x)$:

$$F(x) = \begin{cases} u(1-x)/2 + (4-u)\sin(\pi x)/4 \bmod 1 & x \geq 0.5 \\ (ux/2 + (4-u)\sin(\pi x)/4) \bmod 1 & x < 0.5 \end{cases} \tag{2}$$

Where $x \in (0,1), u \in (0,4]$.

2.2 1-D Logistic Map

In this proposed algorithm, we use 1-D logistic map [29] to select particular category of DNA operations or DNA encoding rules. 1-D logistic map can be defined as in Equation (3).

$$f(x) = rx(1-x) \ x \in [0,1], \tag{3}$$

where $x \in (0,1), r \in (0,4]$. we can figure that when $r \in (3.9,4]$ the random-like sequence is in 0 and 1.

2.3 Lorenz System

As a continuous three-dimensional chaotic system, Lorenz system is defined by the Equation (4):

$$\begin{cases} \dot{x} = a(y-x) \\ \dot{y} = cx - y - xz \\ \dot{z} = xy - bz \end{cases} \tag{4}$$

Where a, b, c are system parameters, the system is in a chaotic state, while $a = 10$, $b = 28$, $c = 8/3$. It is essential to disperse the system by the fourth-order Runge-Kutta method for encrypting image.

2.4 DNA Encoding and Computing

DNA is composed of four deoxynucleotides A (adenine), G (guanine), C (cytosine), T (thymine), where G and C are complementary, so are A and T. Generally, 0 and 1 are complement to each other in binary system. Hence, 00, 11, 01, 10 could be encoded into the four bases. There are 24 kinds of DNA encoding methods according to combinatorics, but out of which only 8 coding combinations are effective because of the complementary relationship between the four, as listed in Table 1.

Table 1: Encoding and decoding rules

Rule	1	2	3	4	5	6	7	8
00	A	A	T	T	C	C	G	G
01	C	G	C	G	A	T	A	T
10	G	C	G	C	T	A	T	A
11	T	T	A	A	G	G	C	C

In image cryptosystem, the gray value of a pixel can be expressed as its corresponding binary sequence, and then encoded into a DNA sequence. on the contrary, a DNA sequence can be translated into a pixel value. For instance: The pixel value 196, its binary sequence 11000100 could be encoded into a DNA sequence GCAC adopting DNA encoding Rule 5. And so on 55 is gained by decoding the DNA sequence with Rule 7. Additionally, we apply different operations of DNA sequence to encrypt the image. The details of the addition, subtraction, XOR DNA operations rules are shown in the following tables, Table 2 to Table 4.

Table 2: XOR operation

$\oplus$	A	C	T	G
A	A	C	T	G
T	T	G	A	C
C	C	A	G	T
G	G	T	C	A

Table 3: Addition operation

+	A	C	T	G
A	C	A	G	T
T	G	T	C	A
C	A	C	T	G
G	T	G	A	C

Table 4: Subtraction operation

-	A	C	T	G
A	C	G	A	T
T	G	T	C	A
C	A	C	T	G
G	T	A	G	C

Table 5: XNOR operation

$\odot$	A	C	T	G
A	C	A	G	T
T	T	G	C	A
C	A	C	T	G
G	T	G	A	C

Inspired by the DNA Addition, Subtraction, and Exclusive OR operations, we proposed XNOR DNA operation that is shown in Table 5. From this table, we can detect the the value of each row or column is unique. That is, the outcome of XNOR DNA operation is distinctive. In this paper, we will apply these DNA operations rules to diffuse pixel gray values.

2.5 Hash-256

Hash functions are mainly used to provide the security service of integrity. Hash-256 is a widely used cryptographic hash function, which generates 256 bits hash value typically presented as a 64 digit hexadecimal number literally. Due to its good feature of security, even one-bit change can lead to a significant difference between two images. We divide the 256-bit secret key into 8-bit blocks(ki), so K can also be expressed as follows.

$$K = k_1, k_2, k_3 \cdots, k_{32}$$

The initial values can be derived as follows.

$$\begin{cases} x_1 = x'_1 + \frac{(k_1 \oplus k_2 \oplus k_3 \oplus \cdots \oplus k_{11})}{256} \\ x_2 = x'_2 + \frac{(k_{12} \oplus k_{13} \oplus k_{14} \oplus \cdots \oplus k_{22})}{256} \\ x_3 = x'_3 + \frac{(k_{23} \oplus k_{24} \oplus k_{25} \oplus \cdots \oplus k_{32})}{256} \end{cases} \quad (5)$$

$$x_{avg} = \frac{x_1 + x_2 + x_3}{3}, \quad (6)$$

where x'_1,x'_2 and x'_3 are the initial given values.

3 Proposed Cryptosystem

3.1 Key Image Generation

In the image encryption algorithm, the key image is generated by the following steps:

Step 1: Use Equation(5) to modify the initial conditions x'_1,x'_2 and x'_3.

Step 2: On the condition of the parameters e, u and the modified initial values x_1x_2 and x_3, TSS-based CML is executed for 500 times for avoiding the transient effect. Continue to execute the chaotic map for M+$4N$ times and three pseudo-random sequences CL_1, CL_2, CL_3 are obtained. CL_1, CL_2, CL_3 are converted into six sequences as follows:

$$\begin{cases} s'_1 = CL_1(1:2M) \\ s'_2 = CL_2(1:2M) \\ s'_3 = CL_3(1:2M) \\ s'_4 = CL_1(2M+1:4M+N) \\ s'_5 = CL_2(2M+1:4M+N) \\ s'_6 = CL_3(2M+1:4M+N) \end{cases} \quad (7)$$

$$s_j = floor((s'_j \times 10^6 - fix(s'_j \times 10^6)) \times 10^{10}) \mod 256, \quad (8)$$

where j=1, 2, ..., 6, *floor*(a) returns the nearest integer to a towards minus infinity, *fix*(a) rounds a to the nearest integer towards zero. After executing Equation (7) and Equation (8), the pseudo-random sequence of s_j (j=1, 2, ..., 6) is in 0 and 255.

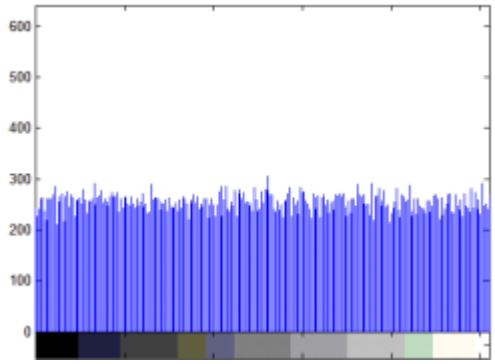


Figure 1: Histogram of the key image

Step 3: The six pseudo-random sequences are handled to generate tow sequences:I_1 of length M and I_2 of length N. Then a key image $KI(M,N)$ is constructed by I_1 and I_2:

$$\begin{aligned} I_1 &= s_1(1:M) \oplus s_2(M+1:2M) \oplus s_3(1:M) \\ I_2 &= s_4(N+1:2N) \oplus s_5(1:N) \oplus s_6(N+1:2N) \end{aligned}$$

$$KI = I_1 \times I_2 = \begin{bmatrix} I_{11} \\ I_{12} \\ \vdots \\ I_{1M} \end{bmatrix} \times \begin{bmatrix} I_{21} & I_{22} & \cdots & I_{2N} \end{bmatrix} = \begin{bmatrix} I_{11}I_{21} & I_{11}I_{22} & \cdots & I_{11}I_{2N} \\ I_{12}I_{21} & I_{12}I_{22} & \cdots & I_{12}I_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ I_{1M}I_{21} & I_{1M}I_{21} & \cdots & I_{1M}I_{2N} \end{bmatrix}$$

Figure 1 shows a sequence of random pixels value which are generated by the proposed method through histogram. The information entropy of the generated key image is 7.9979.

3.2 Encryption Algorithm

In the proposed encryption algorithm, particular DNA encoding rules and DNA operations are randomly decided by 1-D logistic map. Firstly, the SHA-256 is applied on the original image to produce the sequence K and then, the initial values of the TSS-based CML system can be calculated using K. Secondly, R,G and B components of original image and the key image are encoded, applying a randomly selected rule from Table 1, into four DNA sequence matrices. Thirdly, employ encoded key image to conduct random DNA operations with the encoded plain images to obtain a transitional image. Fourthly, the transitional image is permuted by using a Lorenz chaotic sequence. Finally, decode the permuted DNA matrix applying a randomly selected rule from Table 1 to gain the eventual cipher image. The details of the encryption algorithm is presented as follows:

Step 1: The input is a original image $P(M,N,3)$ which M and N express the width and height of the image, respectively.

Step 2: Produce the key sequence K and the initial values x'_1,x'_2 and x'_3 of the Lorenz system and the initial value x_avg of the 1-D Logistic map according to Section 2.5.

Step 3: the plain image is divided into three components, and we obtain three components, R, G and B, and convert the R, G, B to binary matrices R(M,N*8), G(M,N*8) and B(M,N*8), then encode R, G, B by rows with DNA rules that are decided by Equation (2) and Equation (9) and gain three DNA sequence matrices Pr(M,N*4), Pg(M,N*4) and Pb(M,N*4).

$$rule = \lfloor x \times 8 \rfloor + 1. \tag{9}$$

In Equation (9), *rule* is the selected type of DNA rule, which occupies an important position in the encoding stage. The initial value of Equation (2) is provided by Equation (5) and Equation (6). The details about DNA rules are shown in Table 1. Each pixel of a row is coding by particular DNA rule. After all pixels of image are encoded, the size of encoded images are 4*M*N.

Step 4: Generate key image according to Section 3.1, then encode KI by rows with DNA rules that are decided by Equation (2) and Equation (9) and obtain a encoded DNA sequence matrices $KI_e(M,N*4)$.

Step 5: Execute DNA operations between the encoded plain image (Pr, Pg and Pb) and the encoded key image (KI_e) row by row. The particular type of DNA operations is determined by Equation (2) and Equation (10). Details on DNA operations are listed in Table 2 to Table 5.

$$\begin{cases} op = \lfloor x \times 3 \rfloor + 1 \\ pr' = pr \ op \ KI_e \\ pg' = pg \ op \ KI_e \\ pb' = pb \ op \ KI_e \end{cases} \tag{10}$$

Where op is the selected type of DNA operation. Carry out the selected operation row by row After the encoded transitional images are generated, namely Pr', Pg' and Pb', in the process of this period, four kinds of DNA operations (XOR, XNOR +, -) are randomly executed. The size of encoded transitional images are 4*M*N.

Step 6: Generate three chaotic sequences according to the initial value x'_1,x'_2 and x'_3 of the Lorenz system. Executing Equation (4) with the fourth-order Runge-Kutta method for 1000 times to avoid the transient effect, where the step size of the Runge-Kutta method is 0.001. Continue to iterate Lorenz system, three pseudo-random sequences sx, sy and

sz are generated, whose length is $M*N*4$. Then the three sequences are handled by Equation (11).

$$\begin{cases} (lx, fx) = sort(sx) \\ (ly, fy) = sort(sy) \\ (lz, fz) = sort(sz) \end{cases} \quad (11)$$

Where $sort()$ is the sequencing index function, fx is the new sequence after ascending to sx, lx, ly and lz are the index value of fx, fy and fz, respectively.

Step 7: Convert the three binary matrices Pr', Pg' and Pb' to three vectors $Vr(M*N*4)$, $Vg(M*N*4)$ and $Vb(M*N*4)$, respectively. Confuse Vr, Vg and Vb according to:

$$\begin{cases} Vr'(i) = Vr(lx(i)) \\ Vg'(i) = Vg(ly(i)) \\ Vb'(i) = Vb(lz(i)) \end{cases}$$

Step 8: Convert Vr', Vg' and Vb' to three matrices $Re(M, N*4)$, $Ge(M, N*4)$ and $Be(M, N*4)$, respectively. Decode Re, Ge and Be exploiting a selected DNA encoding rule and generate three matrices Rb,Gb and Bb. The decoding rule is according to Equation (9). Randomly DNA decoding and DNA encoding enhance the performance of diffusion process of the proposed algorithm.

Step 9: Finally, merge Rb,Gb and Bb images and that is the ultimate cipher image.The cipher image is with size $M*N$.

3.3 Decryption Algorithm

Decryption algorithm is the inverse process of encryption. receivers should have already obtained the secret keys applied to encrypt the original images. Then we can decode cipher images by following steps:

Step 1: Using randomly selected DNA rules, encode the R, G and B components of the ciphered image. We obtain three matrices Re, Ge and Be, and we convert them to three vectors Vr', Vg' and Vb'. As it is mentioned in Step 3 of the encryption algorithm.

Step 2: Vr', Vg' and Vb' are confused vectors. In order to obtain the non-confused vectors Vr, Vg and Vb, we invert Step 7 which is mentioned in the encryption algorithm as follows:

$$\begin{cases} Vr(i) = Vr'(lx(i)) \\ Vg(i) = Vg'(ly(i)) \\ Vb(i) = Vb'(lz(i)) \end{cases}$$

Where lx,ly and lz are generated as it is mentioned in Step 6 of the encryption algorithm.

Step 3: Convert the three vectors Vr,Vg and Vb to three matrices Pr',Pg' and Pb'.

Step 4: Use encoded key image and encoded cipher image to generate the transitional encoded image. The particular DNA operation is illustrated in Step 5 of Encryption algorithm. After we invert the step 5 of the encryption algorithm to obtain Pr, Pg and Pb, and KI_e is obtained and as it is mentioned in Step 4 of the encryption algorithm.

Step 5: Decode the Pr, Pg and Pb to get the R, G and B components of the plain image. The particular rule is illustrated in Step 3 of the encryption algorithm.

Step 6: Finally, merge R, G and B images and that is the ultimate original image.

4 Simulation Result and Security Analysis

In this paper, we use the standard 256*256*3 color image of "Lena" as the input image. We utilize MATLAB 7.12 to simulate the encryption and decryption operations and set parameters e=0.01, u=1.4356, r=1.4356, x'_1=0.5346x'_2=0.4846, x'_3=0.6969.

4.1 Key Space

A key space larger than 2100 could guarantee high level of security from the crytography of view [27]. In the proposed cryptosystem, the keys are:

1) The given initial values of x'_1,x'_2 and x'_3.

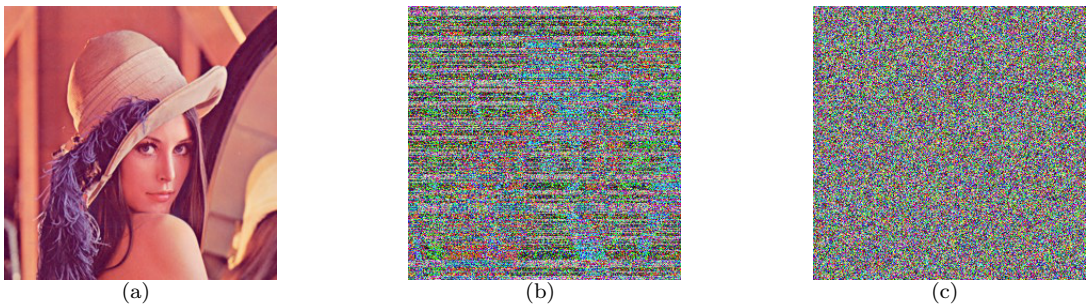
2) The 256-bit long hash value.

3) The parameters of e and u of TSS-based CML and r of the logistic map.

For the initial conditions x'_1,x'_2,x'_3, r, e, and u ,if the precision is 10^{14}, the key space size will be 10^{84}. Further, the key space of the security SHA-256 is 2^{128}, we can get the total key space $S = 2^{128}*10^{84} \approx 3.4*10^{122}$, which is enough to prevent the exhaustive attack. Thus, brute-force attacks on the key are impossible.

4.2 Key Sensitivity

A good encryption algorithm should be sensitive to the secret key; that is, a very tiny different in the secret key will cause a greatly significant change in the output. We conduct a secret key sensitivity test using a key that is just little different from the original key to encrypt Lena image. One of the keys x'_1, u is altered tinily and keep other keys parameters unchanged, then the encrypted image is decrypted by the changed keys. As it can be seen from Figure 2, no effective information is decrypted, which suggests that the proposed cryptosystem could resist the exhaustive attack.



(a) (b) (c)

Figure 2: Key sensitivity tests: decrypted images while (a) right keys, (b) $x'_1 + 10^{-16}$, (c) $u + 10^{-15}$

4.3 The Histogram Analysis

Image histogram is a significant characteristic in image analysis. An ideal cipher image should has a uniform frequency distribution. Figure 3 and Figure 4, illustrate the histograms of the plain and cipher images,it clearly shows that the histograms of the cipher image is uniform and random-like, which suggests that the proposed algorithm does not provide any useful statistic information in the cipher image.

4.4 Information Entropy

The information entropy calculated by Equation (12) is a significant features for measuring the randomness of the cipher image. The gray values distribute more uniformly, the entropy is more close to its ideal value:

$$H(m) = -\sum_{i=0}^{M-1} p(m_i)\log_2 p(m_i), \sum_{i=0}^{M-1} p(m_i) = 1, \quad (12)$$

where m_i (i = 0, 1,$\cdots$, M-1) represents the gray values, $p(m_i)$ (i = 0, 1,$\cdots$, M-1) represents the probability of the symbol s. The entropy should ideally be 8 for a cipher image with 256 gray levels, which indicates that the information is uncertainty. In the paper the average information entropy of the cipher image is 7.9985, close to the ideal value 8. Hence, we conclude that the proposed algorithm has high randomness. The entropy for cipher images using different encryption algorithm are calculated and listed in Table 6, the result of the proposed algorithm in this paper is larger than other algorithms.

Table 6: Results of information entropy

Algorithm	Entropy
Ours	7.9985
[27]	7.9971
[2]	7.9856
[23]	7.9965

4.5 Correlation of Two Adjacent Pixels

To analyze the correlation of the plain image and cipher image, we have randomly selected 5000 pairs of adjacent pixels from plain-image and cipher-image and have calculated the correlation coefficients as follows:

$$\begin{aligned} E &= \frac{1}{N}\sum_{i-1}^{N} x_i \\ D(x) &= \frac{1}{N}\sum_{i-1}^{N} (x_i - E(x))^2 \\ Cov(x,y) &= \frac{1}{N}\sum_{i-1}^{N} (x_i - E(x))(y_i - E(y)) \\ r_{xy} &= \frac{Cov(x,y)}{\sqrt{D(x)}\sqrt{D(y)}} \end{aligned} \quad (13)$$

The x and y represent gray-level values of two adjacent pixels. The correlation of R, G and B components of plain image and cipher image of Lena is shown in Figure 5. Table 7 shows that the dependence between adjacent pixels of the cipher image is much smaller than of plain-image. These results clearly show that the correlation coefficients of the plain image are close to 1 while those of the cipher image are nearly 0 and the distribution of adjacent pixels is fairly uniform. It indicates that the proposed algorithm has successfully eliminated the correlation of adjacent pixels in the plain image so that neighboring pixels in the cipher image virtually have no correlation. So the proposed algorithm can resist the statistic attacks.

4.6 Differential Attack

The number of pixels change rate (NPCR) and the unified average changing intensity (UACI) for the cipher images are generally applied to evaluate the number of pixels change rate. $C_1(i, j)$ and $C_2(i, j)$ stand for two cipher images which corresponding plain images are only one pixel value difference.

$$\begin{aligned} NPCR &= \frac{\sum_{i=0}^{M-1}\sum_{i=0}^{N-1} D(i,j)}{M \times N} \times 100\% \\ D(i,j) &= \begin{cases} 0, C_1(i,j) = C_2(i,j) \\ 1, C_1(i,j) \neq C_2(i,j) \end{cases} \\ UACI &= \frac{1}{M \times N}\left[\sum_{i=0}^{M-1}\sum_{i=0}^{N-1} \frac{|C_1(i,j) - C_2(i,j)|}{255}\right] \times 100\%, \end{aligned} \quad (14)$$

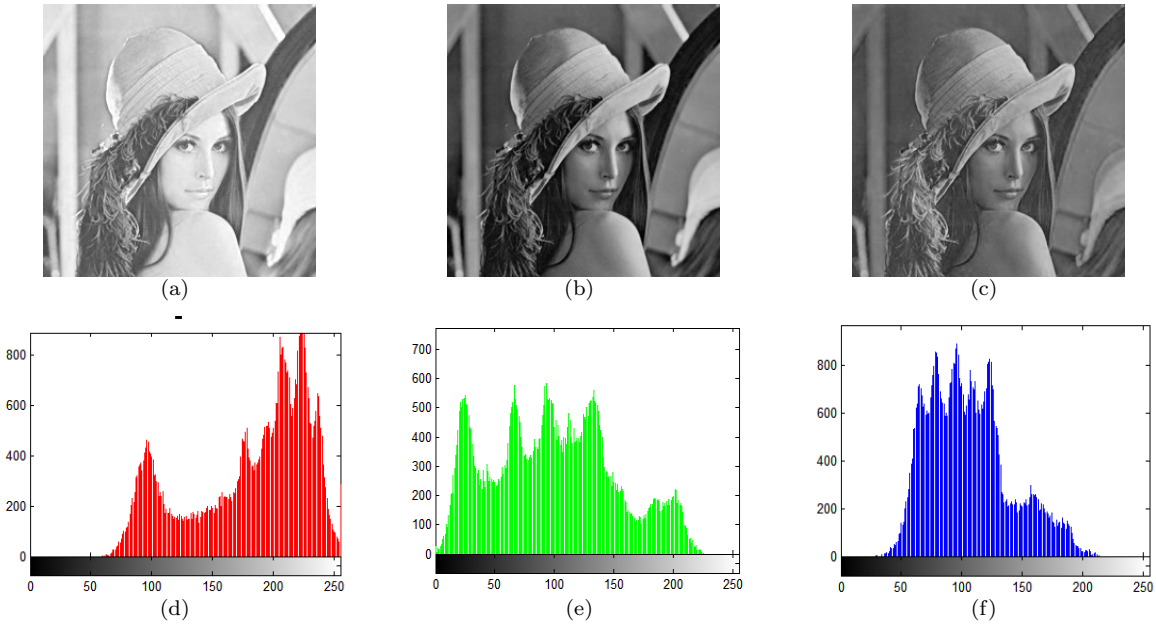


Figure 3: (a) plain image Lena-R, (b) plain image Lena-G, (c) plain image Lena-B, (d) the histogram of the plain image Lena-R, (e) the histogram of the plain image Lena-G, (f) the histogram of the plain image Lena-B

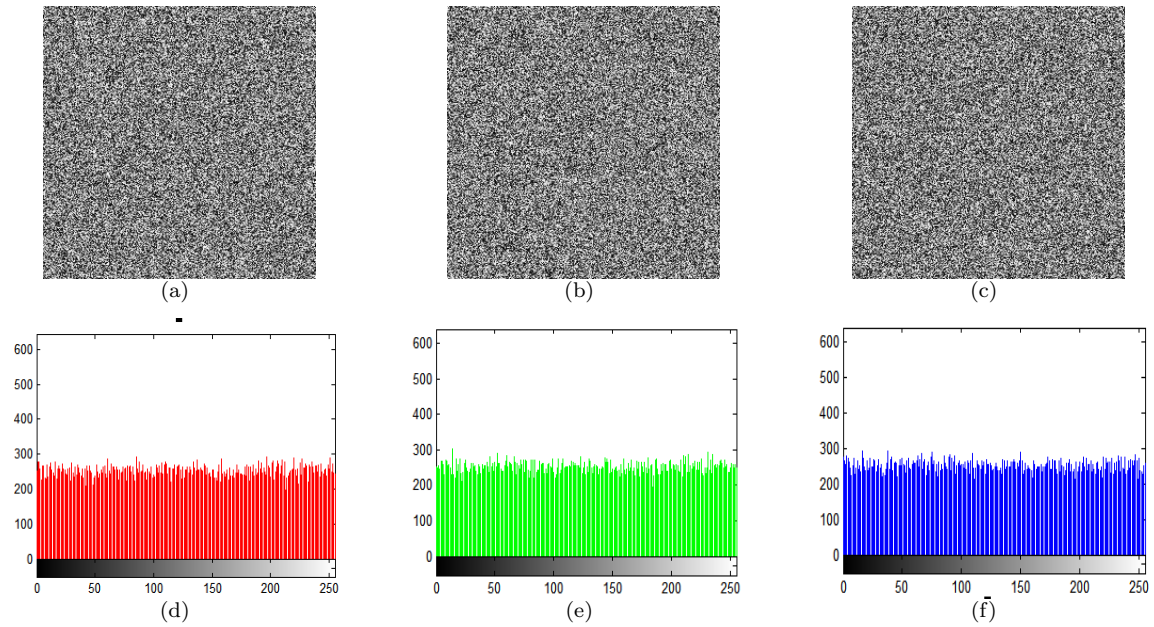


Figure 4: (a) The encrypted image Lena-R, (b) The encrypted image Lena-G, (c) The encrypted image Lena-B, (d) the histogram of the encrypted Lena-R, (d) the histogram of the encrypted Lena-G, (f) the histogram of the encrypted Lena-B.

Table 7: The related correlation coefficient between plain-image and cipher image

Scan direction	Lena					
	Plain image			Cipher image		
	R	G	B	R	G	B
Horizontal	0.9828	0.9725	0.9725	0.0095	0.0183	0.0034
Vertical	0.9689	0.9700	0.9486	-0.0026	0.0001	-0.0035
Diagonal	0.9704	0.9585	0.9585	0.0078	-00039	0.0052

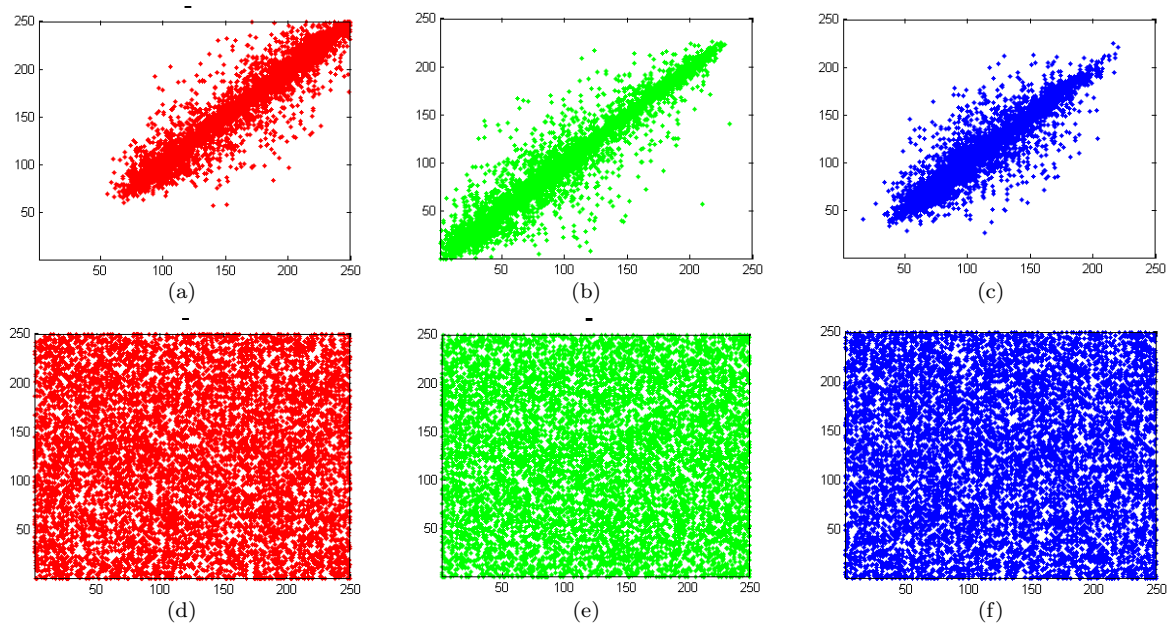


Figure 5: Distribution of two horizontally adjacent pixels in the plain image of Lena in the (a) red, (b) green, (c) blue components. The distribution of two horizontally adjacent pixels in the cipher-image of Lena in the (d) red, (e) green, (f) blue components.

where M and N are the width and height of the cipher image, respectively. In order to evaluate the plain image sensitivity of the proposed algorithm, one pixel randomly selected from the plain image is changed. Two cipher images are generated by encrypting the plain and the modified plain images using the proposed algorithm. The UPCR and UACI between the two cipher images are listed in Table 8 (The experiment is performed over 100 times.). So we can see that the NPCR and UACI are extremely close to the expected values, so the proposed algorithm has the ability to resist the differential attack.

4.7 Data Loss Attack

An ideal cryptosystem should be against data loss attack through transmission and storage. The size of 64*64, 128*128, 256*128 are deleted from the cipher image to evaluate the robustness of the proposed algorithm against the cropping attack, which are shown in Figure 6 (a)-(c). The corresponding decrypted images are shown in Figure 6 (d)-(f) can still be recognizable. So we prove that the proposed algorithm has the ability to resist the data loss attack.

5 Comparison

To demonstrate its superiority, the proposed cryptosystem is compared with the existing image encryption techniques towards some performance indicators, as in Tables 6 and 8. For key space analysis, it is sufficiently large to resist the exhaustive attack. The correlation coefficients of our cryptosystem are more close to 0 than the encryption methods [5, 25, 28, 32], which reveals that the cryptosystem withstands the statistical attack better. The information entropy in this paper is higher compared with those in [2, 23, 39]. Table 8 exhibits that the NPCR, UACI values of the proposed cryptosystem are close to the ideal values, meaning the image cryptosystem with the ability to against the known-plaintext and the chosen-plaintext attacks.

6 Conclusions

In proposed algorithm, a robust image encryption algorithm established on the spatiotemporal chaotic system and DNA operation is proposed. We proposed TSS combined with CML to make up a new spatiotemporal chaos for generating more random sequences. In the DNA operation process, adding the DNA XNOR operation, through this improvement, not only improve the randomness of the encryption,but also enhance the pixels diffusion effect. In order to guarantee the sensitivity of the cryptosystem, the algorithm randomly determine the DNA encode rules and DNA operation which is decided by one-dimensional logistic map. Through the experimental result and security analysis, we find that our algorithm has good encryption effect, larger secret key space and high sensitive to the secret key. Furthermore, the proposed algorithm also can resist most known attacks, such as statistical analysis and exhaustive attacks. All these features show that our algorithm is very suitable for digital image encryption.

Table 8: Results of NPCR and UACI

Image	NPCR			UACI		
	R	G	B	R	G	B
Our (Lena)	99.6395	99.6378	99.6564	33.6875	33.4883	33.4796
Our (Peppers)	99.6404	99.6299	99.6283	33.3998	33.5734	33.5387
[5]	99.60	99.60	99.59	33.52	33.49	33.38
[25]	99.6086	99.6086	99.6086	33.5000	33.5000	33.5000
[32]	99.61	99.61	99.61	33.38	33.38	33.38
[28]	99.5862	99.2172	98.8479	33.4834	33.4639	33.2689

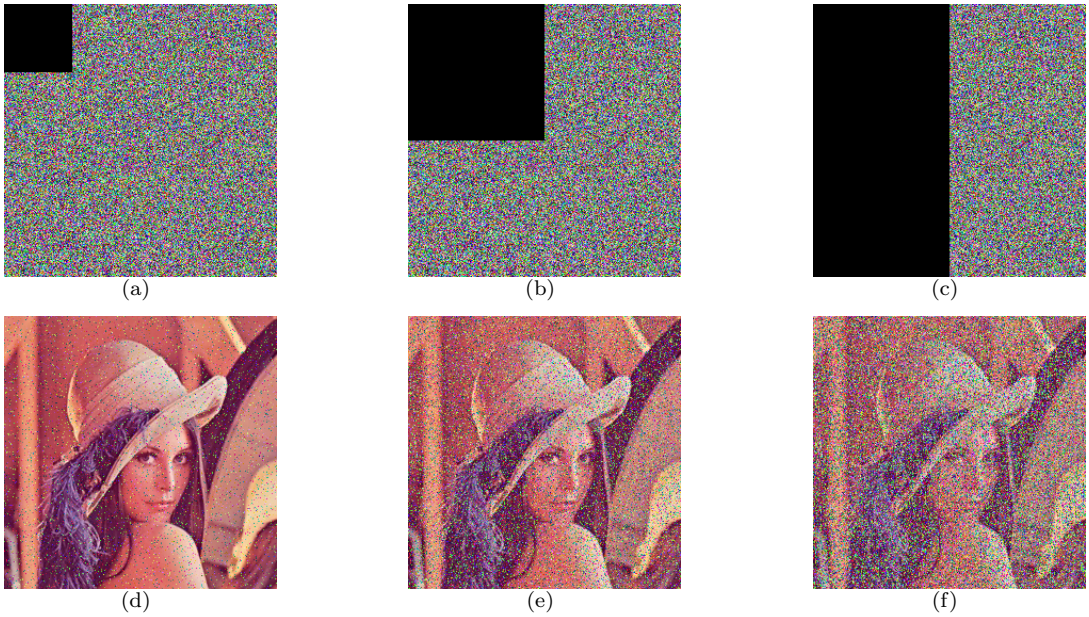


Figure 6: The encrypted Lena with (a) 1/16, (b) 1/4, (c) 1/2 data cropping; corresponding decrypted images (d)-(f) from (a)-(c).

References

[1] S. E. Azoug and S. Bouguezel, "A non-linear preprocessing for opto-digital image encryption using multiple-parameter discrete fractional fourier transform," *Optics Communications*, vol. 359, no. 1, pp. 85–94, 2016.

[2] R. Bechikh, H. Hermassi, A. A. A. El-Latif, R. Rhouma, and S. Belghith, "Breaking an image encryption scheme based on a spatiotemporal chaotic system," *Image Communication*, vol. 39, no. PA, pp. 151–158, 2015.

[3] X. Chai, Z. Gan, K. Yang, Y. Chen, and X. Liu, "An image encryption algorithm based on the memristive hyperchaotic system, cellular automata and DNA sequence operations," *Signal Processing Image Communication*, vol. 52, pp. 6–19, 2017.

[4] C. C. Chang, M. S. Hwang, T. S. Chen, "A new encryption algorithm for image cryptosystems", *Journal of Systems and Software*, vol. 58, no. 2, pp. 83–91, Sep. 2001.

[5] P. Devaraj and C. Kavitha, *A Coupled Chaos Based Image Encryption Scheme Using Bit Level Diffusion*, Springer International Publishing, 2015.

[6] R. Enayatifar, A. H. Abdullah, and I. F. Isnin, "Chaos-based image encryption using a hybrid genetic algorithm and a DNA sequence," *Optics & Lasers in Engineering*, vol. 56, no. 5, pp. 83–93, 2014.

[7] R. Enayatifar, H. J. Sadaei, A. H. Abdullah, M. Lee, and I. F. Isnin, "A novel chaotic based image encryption using a hybrid model of deoxyribonucleic acid and cellular automata," *Optics & Lasers in Engineering*, vol. 71, pp. 33–41, 2015.

[8] T. Hu, Y. Liu, L. H. Gong, and C. Y. Ouyang, "An image encryption scheme combining chaos with cycle operation for DNA sequences," *Nonlinear Dynamics*, pp. 1–16, 2016.

[9] I. A. Ismail, M. Amin, and H. Diab, "A digital image encryption algorithm based a composition of two chaotic logistic maps," *International Journal of Network Security*, vol. 11, no. 1, pp. 1–10, 2010.

[10] B. Jin, D. Wang, and G. Xu, "A new method of fast image encryption based on image characteristics

and DES," in *Wireless Communication and Sensor Network: Proceedings of the International Conference on Wireless Communication and Sensor Network (WCSN'15)*, pp. 520–525, 2016.

[11] C. Jin, H. Liu, "A color image encryption scheme based on Arnold scrambling and quantum chaotic," *International Journal of Network Security*, vol. 19, no. 3, pp. 347–357, 2017.

[12] J. Li, Y. Xing, C. Qu, and J. Zhang, "An image encryption method based on tent and lorenz chaotic systems," pp. 582–586, 2015.

[13] Y. Liu, "Cryptanalyzing a RGB image encryption algorithm based on DNA encoding and chaos map," *Optics and Laser Technology*, vol. 60, pp. 111–115, 2013.

[14] G. Lokeshwari, S. Susarla, and S. U. Kumar, "A modified technique for reliable image encryption method using merkle-hellman cryptosystem and RSA algorithm," *Journal of Discrete Mathematical Sciences & Cryptography*, vol. 18, no. 3, pp. 293–300, 2015.

[15] K. J. Nandu and R. G. Kumar, "Security enhanced image encryption using password based AES algorithm," *International Journal of Engineering & Technical Research*, vol. V4, no. 6, 2015.

[16] S. V. Sathyanarayana, M. A. Kumar, and K. N. H. Bhat, "Symmetric key image encryption scheme with key sequences derived from random sequence of cyclic elliptic curve points," *International Journal of Network Security*, vol. 12, no. 3, pp. 137–150, 2011.

[17] C. Song and Y. Qiao, "A novel image encryption algorithm based on DNA encoding and spatiotemporal chaos," *Entropy*, vol. 17, no. 10, pp. 6954–6968, 2015.

[18] A. Souyah and K. M. Faraoun, "An image encryption scheme combining chaos-memory cellular automata and weighted histogram," *Nonlinear Dynamics*, vol. 86, no. 1, pp. 1–15, 2016.

[19] W. Srichavengsup and W. San-Um, "Data encryption scheme based on rules of cellular automata and chaotic map function for information security," *International Journal of Network Security*, vol. 18, no. 6, pp. 1130–1142, 2016.

[20] Z. Tang, X. Zhang, and W. Lan, "Efficient image encryption with block shuffling and chaotic map," *Multimedia Tools & Applications*, vol. 74, no. 15, pp. 1–20, 2015.

[21] S. M. Wadi and N. Zainal, "High definition image encryption algorithm based on AES modification," *Wireless Personal Communications*, vol. 79, no. 2, pp. 811–829, 2014.

[22] O. Wahballa, A. Wahaballa, F. Li, I. Ibn Idris and C. Xu, "Medical image encryption scheme based on Arnold transformation and ID-AK protocol," *International Journal of Network Security*, vol. 19, no. 5, pp. 776–784, 2017.

[23] X. Wang, L. Teng, and X. Qin, "A novel colour image encryption algorithm based on chaos-memory," *Signal Processing*, vol. 92, no. 4, pp. 1101–1108, 2012.

[24] X. Y. Wang, P. Li, Y. Q. Zhang, L. Y. Liu, H. Zhang, and X. Wang, "A novel color image encryption scheme using DNA permutation based on the lorenz system," *Multimedia Tools & Applications*, pp. 1–23, 2017.

[25] X. Y. Wang, L. Yang, R. Liu, and A. Kadir, "A chaotic image encryption algorithm based on perceptron model," *Nonlinear Dynamics*, vol. 62, no. 3, pp. 615–621, 2010.

[26] X. Y. Wang, Y. Q. Zhang, and X. M. Bao, "A novel chaotic image encryption scheme using DNA sequence operations," *Optics and Lasers in Engineering*, vol. 73, no. 73, pp. 53–61, 2015.

[27] X. Y. Wang, Y. Q. Zhang, and Y. Y. Zhao, "A novel image encryption scheme based on 2-D logistic map and DNA sequence operations," *Nonlinear Dynamics*, vol. 82, no. 3, pp. 1269–1280, 2015.

[28] X. Wei, L. Guo, Q. Zhang, J. Zhang, and S. Lian, "A novel color image encryption algorithm based on DNA sequence operation and hyper-chaotic system," *Journal of Systems and Software*, vol. 85, no. 2, pp. 290–299, 2012.

[29] G. C. Wu and D. Baleanu, "Discrete fractional logistic map and its chaos," *Nonlinear Dynamics*, vol. 75, no. 1-2, pp. 283–287, 2014.

[30] W. S. Yap, C. W. Phan, W. C. Yau, and S. H. Heng, "Cryptanalysis of a new image alternate encryption algorithm based on chaotic map," *Nonlinear Dynamics*, vol. 80, no. 3, pp. 1483–1491, 2015.

[31] M. Zhang and X. Tong, "A new algorithm of image compression and encryption based on spatiotemporal cross chaotic system," *Multimedia Tools & Applications*, vol. 74, no. 24, pp. 11255–11279, 2015.

[32] Q. Zhang, L. Guo, and X. Wei, "Image encryption using DNA addition combining with chaotic maps," *Mathematical & Computer Modelling*, vol. 52, no. 1112, pp. 2028–2035, 2010.

[33] S. Zhang and T. Gao, "An image encryption scheme based on DNA coding and permutation of hyperimage," *Multimedia Tools & Applications*, vol. 75, no. 24, pp. 17157–17170, 2016.

[34] Y. Zhang, "Cryptanalysis of an image encryption algorithm based on chaotic modulation of arnold dual scrambling and DNA computing," *Advanced Science Focus*, vol. 2, no. 1, pp. 67–82(16), 2014.

[35] Y. Q. Zhang, X. Y. Wang, J. Liu, and Z. L. Chi, "An image encryption scheme based on the mlncml system using DNA sequences," *Optics & Lasers in Engineering*, vol. 82, pp. 95–103, 2016.

[36] T. Zhao, Q. Ran, L. Yuan, Y. Chi, and J. Ma, "Security of image encryption scheme based on multi-parameter fractional fourier transform," *Optics Communications*, vol. 376, pp. 47–51, 2016.

[37] P. Zhen, G. Zhao, L. Min, and X. Jin, "Chaos-based image encryption scheme combining DNA coding and entropy," *Multimedia Tools & Applications*, vol. 75, no. 11, pp. 6303–6319, 2016.

[38] S. Zhou, B. Wang, X. Zheng, and C. Zhou, "An image encryption scheme based on DNA computing and cellular automata," vol. 2016, no. 2, pp. 1–9, 2016.
[39] Y. Zhou, L. Bao, and C. L. P. Chen, "A new 1D chaotic system for image encryption," *Signal Processing*, vol. 97, no. 7, pp. 172–182, 2014.
[40] W. Y. Zibideh and M. M. Matalgah, "Modified data encryption standard encryption algorithm with improved error performance and enhanced security in wireless fading channels," *Security & Communication Networks*, vol. 8, no. 4, pp. 565–573, 2015.

Biography

Xiaodong Li is a M.S. candidate at school of Computer science and technology of Wuhan University of Technology. His research interests include: image encryption, information security.

Cailan Zhou received the M.S degrees in Computer science and technology from Wuhan University of Technology, China. She is an associate professor at the Computer science and technology from Wuhan University of Technology, China. Her main research interests include digital image processing,Machine learning and Deep learning, etc.

Ning Xu received his Ph.D. degree in electronic science and technology from the University of Electronic Science and Technology of China in 2003. Later, he was a postdoctoral fellow with Tsinghua University from 2003 to 2005. Currently, he is a professor at the Computer Science Department of Wuhan University of Technology. Dr. Xus research interests include computer-aided design of VLSI circuits and systems, computer architectures, data mining, and highly combinatorial optimization algorithms.

Cost Analysis for Classification-based Autonomous Response Systems

Yudha Purwanto[1,2], Kuspriyanto[1], Hendrawan[1], and Budi Rahardjo[1]
(Corresponding author: Yudha Purwanto)
School of Electrical Engineering and Informatics, Bandung Institute of Technology[1]
Jl. Ganesha, Tamansari, Bandung, Indonesia
Telkom University[2]
Jl. Telekomunikasi, Dayeuhkolot, Bandung, Indonesia
(Email: omyudha@telkomuniversity.ac.id)
(Received Apr. 18, 2017; revised and accepted Aug. 23, 2017)

Abstract

Recently, cost-based Autonomous Response System (ARS) proposals are based on intrusion detection analysis. However, the implementation of the analysis in multi-class classification-based ARS potentially leads to a wrong response action set decision. This is because the analysis may produces irrelevant response value, as it is not considering the false possibility in a true positive condition. In this paper, we introduce ARS based on cost analysis from a multi-class classification output. The analysis is not only considering the possibility of a right response, but also the possibility of a wrong response from false classification prediction. The response value and expected lost rate are introduced to quantitatively estimate the best response action set. Our simulation for Denial of Service (DoS) attack cases, confirmed the capability of response action set decision algorithm. Our proposed system provides more accurate estimation of response value which leads to lower expected lost rate.

Keywords: Autonomous Response System; Classification; Decision Analysis; Denial of Service; Intrusion Detection

1 Introduction

To stop the traffic flooding attack in Denial-of-Service (DoS) is one important task in network security. Intrusion Detection and Response System (IDRS) is one security mechanism which aims to mitigate the attack impacts on a victim while keeping the damage level to a minimum [8]. Autonomous response system (ARS) is a kind of IRS, which responses to the detected attack autonomously without human intervention. The system is not only required to accurately detect the attacks, but also adaptively determines the best response action set to stop the attacks. Actions such as traceback [28], Intrusion Detection System (IDS) based filtering [12, 16], rate-limiting [13, 20], artificial immunity [1, 2], are examples of DoS attack response proposals. And in [17], the defensive mechanism was categorized by capability focus of each approach.

The heart of an ARS is the decision analysis process, as it is responsible for the decision to be made. Research on cost-based decision analysis has their own characteristics and mechanism in determining the variables. Most approaches obliged to assess every single risk and cost embroiled in decision analysis, such in [3, 6, 10]. But, the major issue of cost-based decision analysis is the necessity to estimate many building factors which have to be defined first during implementation.

To overcome above limitations, several research has proposed the decision analysis, which only directly concerns with IDS effectiveness. In [25], they have proposed decision analysis which considers damage lost and response cost consequences by no intrusion condition. These approaches have not considered the lost or cost consequences by intrusion condition. And in [15], they have modified the previous analysis of cost and lost consequences according to IDS possible conditions which are no intrusion and intrusion. The research has proposed IDS value to quantitatively measure IDS effectiveness by considering most relevant costs of the decision process. Cost analysis also used in Intrusion Detection Network research [4, 5] which used to measure the effectiveness of detection feedbacks in collaboration selection process.

However, the cost-based decision analysis for intrusion detection cannot be directly applied to the multi-class intrusion classification cases. Furthermore, it potentially leads to less precise response action set and higher lost rate. This is because the output of intrusion classification provides more action set possibilities. Thus, the decision analysis must take into account all consequences possibility, including from wrong response possibility. This is because there is still the possibility of false in every classification algorithm. From previous intrusion classification

Table 1: Example of intrusion classification output case

Actual	Classification Prediction		
	Normal	Attack A	Attack B
Normal	TN	FP A	FP B
Attack A	FN A	TP AA	TP AB

algorithm research reports; such in [18, 19, 22]; still, no classifier has a perfect accuracy. For example, in Table 1, the true positive (TP) state in intrusion classification, may consist of several specific false predictions, such as TP_{AB} (false as attack A was predicted as attack B), etc. Each false prediction will take effect on cost and lost consequences in the decision analysis. When this situation is neglected, then the lack of proper decision may be ended in higher system lost.

Therefore, we propose false-aware cost-based decision algorithm for classification-based ARS. Our decision algorithm quantitatively estimates response value of each possible set of response action, and determine best response action set based on response value. We have upgraded the state-of-the-art cost-based decision analysis in [15], by considering the possibility of right or wrong response consequences in the analysis process. Thus, it leads to estimate the relevant best response action set, as it provides more precise response value estimation on all possible responses. We have validated and tested the decision algorithm by synthetic confusion matrices and by the used of three different classification algorithms using KDD Cup 1999 DoS/DDoS revised dataset in [24].

Our decision analysis can accommodate the necessity of cost-based decision analysis for classification-based ARS. To the extent of our knowledge, this is the first research that shows false-aware cost-based decision analysis on intrusion classification. Our decision analysis provides a quantitative estimation of response value and expected lost rate based on classification output. Our proposal is important due to the recent development of ARS, which does not only detect the existence of attack but also determines relevant response action set. In addition, our paper differs from the related study by providing a complete algorithm that covers an autonomous response capability.

This paper is orderly written as follows. In Section 2, discuss the state-of-the-art of response system. In Section 3, our novel decision analysis proposal is introduced with an example of an intrusion classification case. Section 4 shows the response system design in complete framework and algorithm, and also discuss experimental and performance evaluation procedure. Finally, in Section 5, the evaluation results are shown and analyzed to validate our proposal. Section 6 summarized our conclusions and an open problem for possible further research.

2 Related Work

In respect of response system, decision analysis has already been studied in previous research. Game theory is one widely used analysis in response system. In [29], they proposed automated response based on a Stackelberg stochastic game which is a two-player game-theoretic response and recovery strategy, named response and recovery engine (RRE). The multi-objective response action selection quantitatively ranks by fuzzy logic, and the optimal action is determined from game-theoretic optimization process.

The probabilistic method also occupied in decision analysis. In [14], a probability analysis based on stochastic Petri nets, consider detection result in a network which comprised of many nodes. By adjusting a minimum threshold, a dynamic response system was developed based on the detected attacker strength. Reinforcement learning was used in [13], which proposed autonomous response by distributing reinforcement learning of throttle agents. Those agents adaptively and autonomously response DoS attack by learning the scale of rate-limiting action during reinforcement learning.

Cost-benefit analysis is one promising method in response system [15]. However, it has limitation as all of the cost must be defined first and have to be updated periodically, otherwise it will be static cost analysis. Research in [11] has proposed risk analysis by damage cost, operational cost and response cost in a cost-sensitive analysis for IDS. This proposal determines the autonomous response, according to the cumulative cost matrix that combines the different cost features. While others work in the scope of technical approach, in [6] has proposed a cost-benefit analysis which has considered the technical and managerial aspects. The analysis estimates the Return on Investment (ROI) variables in determining best IDS system which provides better ROI.

Cost-based decision analysis based on intrusion detection in [5, 15, 25, 26] have gone beyond the static cost by dynamically calculate the cost based on IDS output. In [25, 26], they have proposed decision analysis based on cost per unit lost ratio, which considers damage Lost and response Cost by no intrusion condition. Then, [15] have upgraded the decision analysis by simplifying cost-benefit estimation. It proposed IDS Value to quantitatively measure IDS effectiveness by considering most relevant costs in an ARS (Believe Desire Intention (BDI) agent environment.

Research in [5] also consider cost analysis of detection feedbacks by the used of false positive and false negative feedbacks. However, those proposals still not pay attention to the possibility of wrong response in intrusion classification output case. We develop beyond those existing cost-based decision analysis by concentrating on cost-based analysis which considers the possibility of wrong response.

3 False-Aware Cost-based Decision Analysis

Research in [5, 15, 25, 26], are fundamentally constructed our false-aware cost-based decision analysis for classification-based response system. The analysis consists of response decision nodes and event nodes. A response decision node is possible response action taken by response system at an operating point. An event node is network condition uncertainty at an operating point. From the combination of possible taken response and condition uncertainty, decision analysis may end up in consequences which are cost or lost condition. In decision analysis, the first important step is to determine the environment and workflow of the system. It significantly affects the result of decision analysis. In this proposal, the analysis is developed according to the framework such in Figure 1.

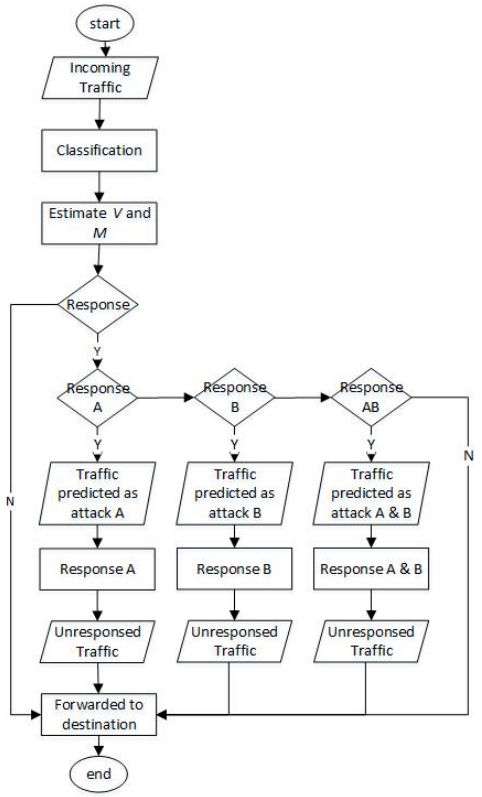


Figure 1: Traffic flow in our proposed ARS

Definition 1. *The Possible responses are all possible response actions that available for the ARS to react to any predicted attack.*

The possible responses at a given operating point are whether the system chose to respond or not respond to any predicted attack. By this situation, the possible responses are no-response and response to the predicted attacks. In classification case, the predicted attack may consist of several types of attack. Thus, the possible responses are no-response and power set of responses to predicted attacks. For example, from a classification report such in Table 1, the all possible responses are no-response, response to attack A, response to attack B, and response to both attacks A and B.

Definition 2. *Cost is a condition where the system takes any precautionary response action. The response took specific cost related to certain action set, based upon the predicted class of attack.*

Definition 3. *Lost is a condition where the system suffers any lost from the attack as system take no-response or wrong response when the attack occurred. The lost is related to damage lost from not responding to the predicted class of attack.*

We have upgraded the decision tree analysis proposed in [5, 15, 25, 26], by considering all possible consequences according to intrusion classification output. Our decision tree analysis not only considers lost consequence by the no-response decision but also lost consequence by wrong response decision. This approach is based on the real condition probability from possible responses. As system took response due to any certain type of attack, it might end up in cost consequence when the response was right, or in lost consequence as the response was wrong. From this process, we have optimized the decision tree analysis as depicted in Figure 2.

Definition 4. *The expected cost of response is the sum of product of expected consequence if the system takes any response. In intrusion detection case, the expected cost of response is related to false positive, false negative, and a true positive.*

In a cost-based decision analysis for intrusion classification, it estimates the expected cost and lost condition for each possible response. From classification output confusion matrix in Table 1, the system has information of hit rate ($H = TP/(TP + FN)$) and False alarm rate ($F = FP/(FP+TN)$). Given the prior probability of an intrusion is happening (p), the expected cost in each possible condition is then estimated by decision tree analysis in Figure 2.

Definition 5. *The expected cost of an operating point is the sum of the product of the expected cost of response from each possible condition. Thus, the expected cost per unit lost of operating point (M) is the cost of an operating point normalized by unit lost.*

The expected cost per unit lost of an operating point is dependent on the cost of an operating point in No-Alarm and in an Alarm condition. To estimate the cost per unit lost ratio (M), all cost per unit lost ratio of all possible responses need to be defined first. For example, in classification case such in Table 1, the all possible responses are summarized in Table 2. In No-Alarm condition, the cost per unit lost ratio given every possible response are

Table 2: Cost analysis based on possible responses in Table 1

Report	**Traffic**	**PossibleResponse**			
		No Response	**Response Attack A**	**Response Attack B**	**Response Attack A B**
No Alarm	Normal	0	C(TN)	C(TN)	C(TN)
	Attack A	$L(FN_A)$	$C(FN_A)$	$L(FN_A)$	$C(FN_A)$
Alarm	Normal	0	$C(FP_A)$	$C(FP_B)$	$C(FP_A + FP_B)$
	Attack A	$L(TP_{AA} + TP_{AB})$	$C(TP_{AA}) + L(TP_{AB})$	$C(TP_{AB} + L(TP_{AB} + TP_{AA})$	$L(TP_{AB}) + C(TP_{AA} + TP_{AB})$

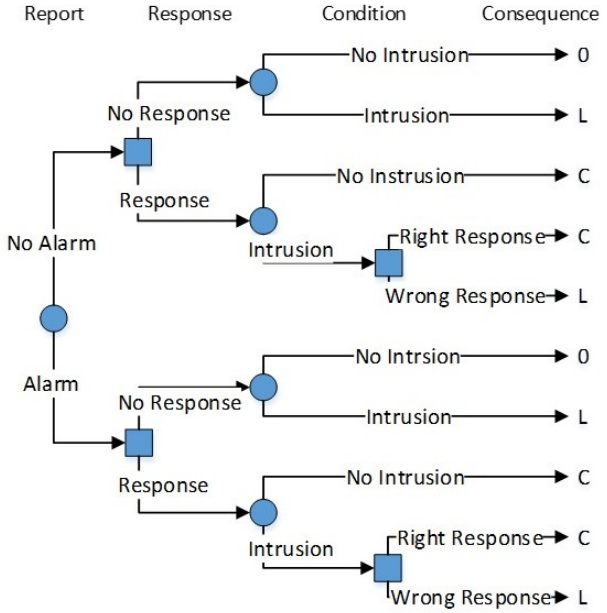


Figure 2: Cost-based decision tree for intrusion classification that considers right or wrong response

such in Equation (1) to Equation (4).

$$M(NoResponse|NoAlarm) = C_{(NoResponse|normal)} + C_{(NoResponse|Attack_A)} = p(1-H) \quad (1)$$

$$M(Response_A|NoAlarm) = C_{(Response_A|normal)} + C_{(Response_A|Attack_A)} = \frac{C}{L}((1-p)(1-F)) + \frac{C}{L}(p(1-H)) \quad (2)$$

$$M(Response_B|NoAlarm) = C_{(Response_B|normal)} + C_{(Response_B|Attack_A)} = \frac{C}{L}((1-p)(1-F)) + (p(1-H)) \quad (3)$$

$$M(Response_{A+B}|NoAlarm) = C_{Response_{(A+B)}|normal)} + C_{(Response_{(A+B)}|Attack_A)} = \frac{C}{L}((1-p)(1-F)) + \frac{C}{L}(p(1-H)) \quad (4)$$

From Equation (1) to Equation (4) system can estimates expected cost to lost ratio from No-Alarm condition which is such in Equation (5).

$$M_{NoAlarm} = min \left\{ \begin{array}{l} (p(1-H)), \\ \frac{C}{L}(((1-p)(1-F)) + (p(1-H)))), \\ ((\frac{C}{L}((1-p)(1-F))) + p(1-H)), \\ (\frac{C}{L}((1-p)(1-F)) + (p(1-H))) \end{array} \right\} \quad (5)$$

The same procedure is performed to estimate the cost per unit lost ratio in an Alarm condition. The cost per unit lost ratio is estimated as every possible response given Alarm condition, such in Equation (6) to Equation (9).

$$M(NoResponse|Alarm) = C_{(NoResponse|normal)} + C_{(NoResponse|Attack_A)} = pH(\frac{(TP_{AA} + TP_{AB})}{TP}) \quad (6)$$

$$M(Response_A|Alarm) = C_{(Response_A|normal)} + C_{(Response_A|Attack_A)} = \frac{C}{L}((1-p)F(\frac{FP_{AA}}{FP}) + pH(\frac{TP_{AA}}{TP})) + pH(\frac{TP_{AB}}{TP}) \quad (7)$$

$$M(Response_B|Alarm) = C_{(Response_B|normal)} + C_{(Response_B|Attack_A)} = \frac{C}{L}((1-p)F(\frac{FP_{BB}}{FP}) + pH(\frac{TP_{AB}}{TP})) + pH \quad (8)$$

$$M(Response_{A+B}|Alarm) = C_{(Response_{A+B}|normal)} + C_{(Response_{A+B}|Attack_A)} = \frac{C}{L}((1-p)F + pH)) + pH(\frac{TP_{AB}}{TP}) \quad (9)$$

Thus, the expected cost to lost ratio given Alarm con-

dition will be such in Equation (10).

$$M_{Alarm} = min \left\{ \begin{array}{l} pH(\frac{(TP_{AA}+TP_{AB})}{TP}), \\ \frac{C}{L}((1-p)F(\frac{FP_{AA}}{FP}) + pH(\frac{TP_{AA}}{TP})) \\ \quad +pH(\frac{TP_{AB}}{TP}), \\ \frac{C}{L}((1-p)F(\frac{FP_{BB}}{FP}) + pH(\frac{TP_{AB}}{TP})) + pH, \\ \frac{C}{L}((1-p)F + pH)) + pH(\frac{TP_{AB}}{TP}) \end{array} \right\} \quad (10)$$

Finally, the expected cost per unit lost(M) is the sum of the product of the expected cost per unit lost of detector's reports at an operation point which are No-Alarm and Alarm condition, which is $M = M_{NoAlarm} + M_{Alarm}$.

Definition 6. *The response value such in [15]; namely IDS value; is an estimated value of the possible responses at a given operating point.*

It was derived from the normalization between actual reduction of expected cost and maximum possible reduction of expected cost. Actual reduction of expected cost is the reduction of actual expected cost (M) over the expected cost which based only on the information of the probability of intrusion (M_{prop}). And maximum possible reduction is the reduction between actual expected cost per unit lost of operating point (M) over the expected cost of perfect classifier (M_{per}). M_{prop} is the expected cost that corresponds only to the information of the probability of intrusion (p), such in Equation (11). The expected cost of a perfect classifier (M_{per}) was achieved when expected cost per unit lost was applied in a perfect classifier which has $H = 1$ and $F = 0$, such in Equation (12).

$$M_{prop} = min(p, \frac{C}{L}) \quad (11)$$

$$M_{per} = min(p, \frac{C}{L}p). \quad (12)$$

The system estimates the response value (V) which is the same procedure as IDS value from [15] such in Equation (13). The difference is in response value, the system objective is to evaluate the value of every action response set. To extend the analysis, we provide the response value calculation algorithm in Section 4.

$$V = \frac{(M_{prop} - M)}{(M_{prop} - M_{per})} \quad (13)$$

Definition 7. *Best response action set (a), is the set of action determined from related to cost to lost ratio ($\frac{C}{L}$) when decision analysis reaches a maximum response value.*

For the ARS, the best response was automatically determined from minimal expected cost per unit lost. However, from this analysis, the best response is just the decision of whether to respond or not to all predicted attacks (chosen from possible responses). The actual action taken by ARS is response action set ($\{a\}$) at a determined best response. The best response action set then acquired from action set which is related to cost to lost ratio ($\frac{C}{L}$) from obtained maximum response value (V_{max}).

Definition 8. *Lost rate parameter (L) is the expected lost consequence at a given expected cost per unit lost, normalized by maximum lost consequence.*

To estimate lost rate (L), the system estimates the maximum cost consequence from confusion matrix input. It is the sum of product of true in true positive, and all False Negative. From the resulting expected cost consequence, the system can estimate expected lost consequence by a reduction between maximum lost consequence and expected cost consequence. The lost rate is then estimated by the expected lost consequence divided by the maximum lost consequence, such in Equation (14).

$$L = \frac{(TP + FN) - (\frac{M(trueTP+FN)}{M|V_{max}}}{TP + FN}. \quad (14)$$

4 System Design

4.1 Framework and Algorithm

Our proposed system may reside on any node in a network, including in near destination network as it provides more benefit in security system [22]. Figure 3 represents our framework. The input of our system is basically raw incoming traffic records, which is packet level data. The first stage of our framework is basic features generation process, which is to generate each traffic feature of each data traffic. Assumed, the output of traffic features generation process is a set of traffic features records $X = \{x_1, x_2, x_3, \ldots x_g\}$. Each data in x then enters the second stage; the classification system; to predict types of each individual data in record x. The system evaluates every classification output in confusion matrix and gets set of g confusion matrix records $Y = \{y_1, y_2, y_3, \ldots y_g\}$. Hit rate ($H$), false rate ($F$) and the probability of attack (p) are straightly calculated from each confusion matrix records y in Y. When the traffic data are predicted as normal, then the data enter the fifth stage which is traffic forwarding process. But, when the traffic data are predicted as an attack, then the system enters the third stage. At this stage, the system estimates the response value of each confusion matrix from the earlier step. The decision analysis applied to estimates a set of response value of each y; which produces $V = \{v_1, v_2, v_3, \ldots v_g\}$. In the final stage, the system determines the best responses $Z = \{z_1, z_2, z_3, \ldots z_g\}$ for every estimated response value (v). And finally determines best response action set $A = \{a_1, a_2, a_3, \ldots a_g\}$ from related cost to lost ratio ($\frac{C}{L}$) at given maximum response value.

In this research, we present the algorithm of autonomous response in Algorithm 1. The algorithm firstly determines all possible responses of given classification output (y_g), which is no-response and all subsets of the attack detected responses. Then, the system estimates the cost per unit lost value for every possible response in No-Alarm and Alarm condition. This was done by following decision-tree analysis in Figure 2. All possible

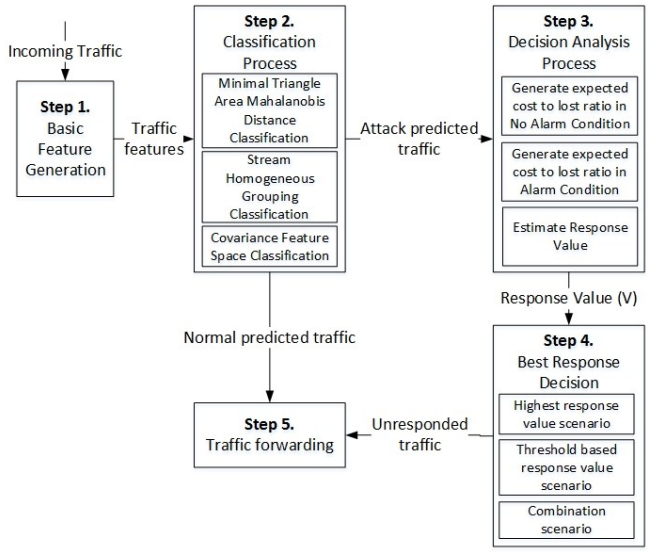


Figure 3: Framework of classification-based response system

responses are represented in a predicted set. The system then estimates the expected cost per unit lost (M_g) from the minimal value of $M_{Alarm} + M_{NoAlarm}$, for every possible action (represented by $col = \forall predicted$). From the expected cost per unit lost calculation (M_g), the system then estimates the best response (z_g), response value (v_g) using Equation (13) and best response action (a_g).

4.2 Test and Data Acquisition

In general, we evaluate our proposed ARS in two stages by doing a comparison between our decision analysis and state-of-the-art cost-based decision analysis in [15]. First is validation, using synthetic confusion matrix which best represents the possible condition of classification output. Second, we evaluate the system by the used of three different classification algorithms. We use KDDCup 99 revised dataset in [24] as input for three classification algorithm to evaluate our decision analysis. The classification was done for all DoS/DDoS data in KDDCup 99 revised, except for Pod, teardrop, and Land attack. This proposal is the extension of our previous research in the classification algorithm in [18]. First is Minimal Mahalanobis Distance Classification (MMDC) algorithm, which is our proposed classification algorithm. We have upgraded the algorithm by the used of minimal triangle area Mahalanobis distance to classify data from [22, 23]. Second is Stream Homogeneous Group Classification (SHGC) algorithm, which is our proposal in [18]. And the last is the covariance feature space classification (CFSC) algorithm in [7], which is stream group-based classification without homogeneous grouping. However, in this paper, we only present the comparison to reveal the effectiveness of false-aware decision analysis, instead of doing a comparison between detection or classification algorithms.

Algorithm 1 Response value (V) estimation

Begin
Initialize input : $ConfusionMatrixY = \{y_1, y_2, \cdots, y_g\}$; $p = \{predictedattackiny_g\}$; $q = \{actualtraffic inx_g\}$
while $y_g \neq \{\}$ **do**
  for $(\frac{C}{L})$ =0.01 to 0.99 **do**
    Calculate H, F, p
    $\{attack\} \leftarrow Powersetof\{p\}$
    $predicted \leftarrow \{normal\} \cup \{attack\}$
    $i \leftarrow size(predicted)$
    $\{traf\} \leftarrow Powersetof\{q\}$
    $actual \leftarrow \{normal\} \cup \{traf\}$
    $j \leftarrow size(actual)$
    GenerateCostTableNoAlarm
    GenerateCostTableAlarm
    for $col = 1$ to j **do**
      $CostNoAlarm_{orderX(col))} \leftarrow sum(CostTableNoAlarm_g(:, col, 1) * (\frac{C}{L}))$
      $LostNoAlarm_{(orderX(col)))} \leftarrow sum(CostTableNoAlarm_g(:, col, 2))$
      $CL_{NoAlarm(col)} \leftarrow CostNoAlarm_{orderX(col)} + CostNoAlarm_{orderX(col)}$
      $CostAlarm_{(orderX(col)))} \leftarrow sum(CostTableAlarm_g(:, col, 1) * (\frac{C}{L}))$
      $LostAlarm_{(orderX(col)))} \leftarrow sum(CostTableAlarm_g(:, col, 2))$
      $CL_{Alarm(col)} \leftarrow CostAlarm_{orderX(col)} + CostAlarm_{orderX(col)}$
    end for
    for $col = 1$ to j **do**
      $M_{(NoAlarm_{orderX(col)})} \leftarrow min(CostNoAlarm_{orderX(col)} + CostNoAlarm_{orderX(col)})$
      $z_{g(\frac{C}{L})NoAlarm} \leftarrow predicted|min(orderX(col))$
      $M_{(Alarm_{orderX(col)})} \leftarrow min(CostAlarm_{orderX(col)} + CostAlarm_{orderX(col)})$
      $z_{g(\frac{C}{L})Alarm} \leftarrow predicted|min(orderX(col))$
    end for
    $M_g \leftarrow M_{(NoAlarm_{orderX(col)})} + M_{(Alarm_{orderX(col)})}$
    $M_{Per} \leftarrow min(p, \frac{C}{L}p)$
    $M_{Prop} \leftarrow min(p, \frac{C}{L})$
    $V_{g(\frac{C}{L})} \leftarrow (M_{Prop} - M_g)/(M_{Prop} - M_{Per})$
  end for
  $z_g \leftarrow z_{g(\frac{C}{L})NoAlarm} \cup z_{g(\frac{C}{L})Alarm}$
  $v_g \leftarrow max(V_{g(\frac{C}{L})})$
  $a_g \leftarrow \exists a_g : (\frac{C}{L}|a_g) = (\frac{C}{L}|v_g)$
end while
End

5 Result and Analysis

5.1 Validation Using Synthetic Confusion Matrices

We do validation of our proposal by generating synthetic confusion matrix such in Table 3. Suppose, we have

Table 3: Synthetic confusion matrix cases

ConfMat1				
Actual	Predicted			
	Norm	Nept	Smu	Back
Norm	200	0	0	0
Nept	2	18	0	0
Smu	3	0	297	0
Back	5	0	0	395

ConfMat2				
Actual	Predicted			
	Norm	Nept	Smu	Back
Norm	192	1	2	5
Nept	0	20	0	0
Smu	0	0	299	0
Back	0	0	0	400

ConfMat3				
Actual	Predicted			
	Norm	Nept	Smu	Back
Norm	199	0	1	0
Nept	0	20	0	0
Smu	1	5	294	0
Back	0	20	0	380

ConfMat4				
Actual	Predicted			
	Norm	Nept	Smu	Back
Norm	199	0	1	0
Nept	0	20	0	0
Smu	1	5	294	0
Back	0	200	0	200

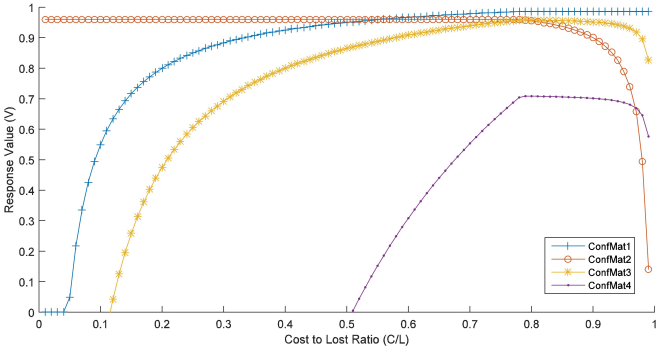


Figure 4: Response value computed over four synthetic confusion matrix cases

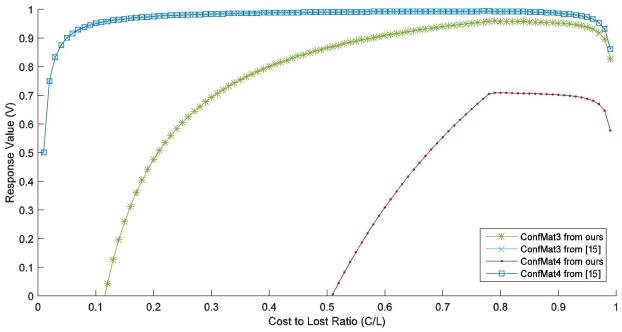


Figure 5: Response value comparison from the validation process

ConfMat1 which represents a high accuracy with some False Negative, ConfMat2 which represents a high accuracy with some false positive, ConfMat3 which represents a low accuracy with low false in true positive, and ConfMat4 which represents a low accuracy with high false in true positive. From these examples, the system determines possible responses set.

The curves of response value (V) toward the different cost to lost ratio ($\frac{C}{L}$), show the value of best response action set for given best response set which is $bestresponse = responsetoNeptuneSmurfBack$. It is obtained as the response provides minimal cost per unit lost among all possible responses set elements. Figure 4 shows that ConfMat2 is the best classification algorithm among these four. The curve from ConfMat2 shows high response value in $\frac{C}{L} < 0,56$ which means the damage cost is almost two times higher than response cost. However, when $\frac{C}{L} > 0,56$ then the best algorithm is ConfMat1, which means the system may afford the higher cost to reach a higher response value. In ConfMat2, higher false positive affects the higher cost but no lost consequences. It means when the $\frac{C}{L} > p$, the higher cost has no benefit as the lost consequence of false positive is none. Even when the cost of action gets higher, the damage lost is none. As for ConfMat1 with higher false negative, the higher cost takes effect on higher response value. It means more response need to be taken to lower the lost consequence from undetected attack in a false negative.

The used of response value in our proposal can accurately estimate the performance of ARS at the corresponding response action set. The lower response value represents the lower accuracy of classification output, which mostly influences by higher false prediction in true positive (TP). The differences are depicted in Figure 5. In the case of ConfMat3 and ConfMat4, our analysis can differentiate the quality of response estimated from classification output. The lower accuracy of ConfMat4 can be estimated by lower response value. However, the analysis in [15] can not differentiate the quality between them. Even when the accuracy of ConfMat4 is getting worse (increasing false in true positive), the IDS value remains high as long as the H, F and p values are the same. For ConfMat1 and ConfMat2 input, which have no false prediction in true positive, both decisions generate the same IDS values.

Wrong estimation of response value (V) potentially leads to poor response action set $\{a\}$. It is important as the action set is determined from the relation between estimated response value (V) and cost to lost ratio ($\frac{C}{L}$). If the response value is wrong, so does the action set determined. In this research, best response action set $\{a\}$ is determined from the maximum response value. For example, in ConfMat3 case. By the used of our proposed analysis, the best response action set was action set related to $\frac{C}{L} = 0.79$ at $V_{max} = 0.956$. By this condition, estimated best response action set is an action set $\{a_{(\frac{C}{L})} : \frac{C}{L} = 0.79\}$ and the expected cost was estimated at $M = 0.625$. Thus, the system potentially experiences a maximum unresolved attack (expected lost rate (L)) of maximum expected lost divided by maximum actual lost, which is 3,47%. However, when the system occupies decision analysis from [15], the best response value is achieved at $V = 0.956$. Then the response action set is estimated at $\{a_{(\frac{C}{L})} : \frac{C}{L} = 0.77\}$ and expected cost $M = 0.610$. Thus, it is worse than our proposal as it potentially raises maximum unresolved attack to 5.83%. The raising ex-

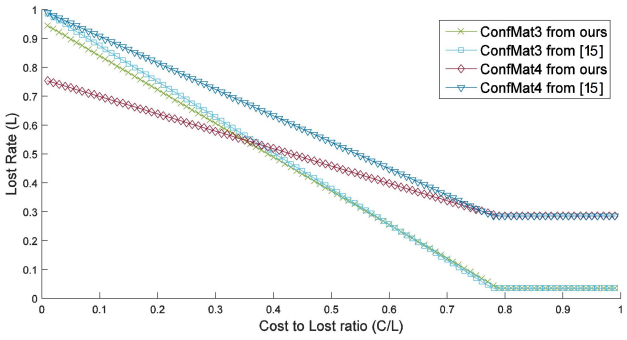


Figure 6: Lost rate exploration from the validation process

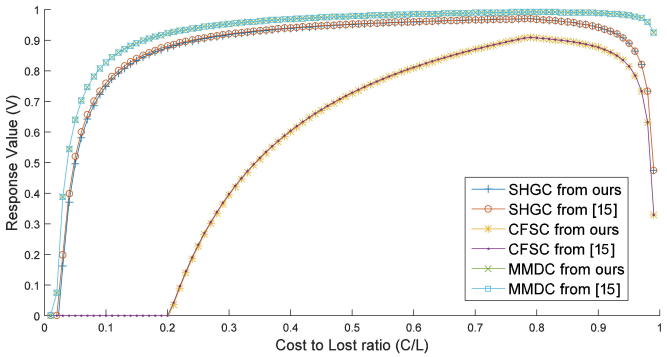


Figure 7: Response value comparison computed over KDD'99 revised dataset

pected lost rate (L) is also shown in ConfMat4 condition, which is the L increases from 28.47% to 29.72%.

From the response system point of view, the expected lost rate is expected lost when the system takes any action on estimated response value. Thus, the difference lost rates are revealed from the difference expected cost per unit lost (M) provided by both decision analysis. The expected lost rate also can be used to estimates response action set. It is used when the system has any expected toleration of lost rate. Figure 6 shows the comparison of expected lost rate(L) between our proposal and [15] for ConfMat3 and ConfMat4. The lower curves show better expected lost rate result. As for ConMat1 and ConfMat2, the lost rates are exactly the same as there are no false in the true positive in both cases.

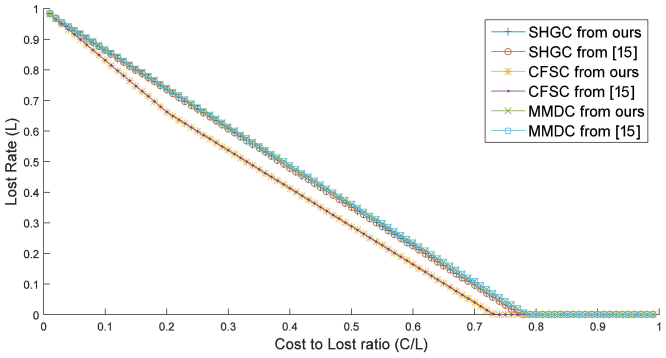


Figure 8: Lost rate exploration computed over KDD'99 revised dataset

5.2 Simulation Using Classification Algorithms

From classification outputs, the ARS autonomously decide whether to respond or not with $1 + 2^n$ possible response subsets (1 is for no-response, and n is the number of predicted attacks). From our simulation using KDD-Cup 99 revised dataset, decision analysis adaptively decides whether to respond or not according to a certain operating point. Best response from minimal expected cost per unit lost is $\{bestresponse\} = \{responsetoNeptune \cap Smurf \cap Back\}$, which is the same for all classification outputs. However, the minimal expected cost per unit lost value of each analysis and case is different. It makes the response value and best response action set are different among these three. The MMDC algorithm which is single by single data analysis provides best response value as the accuracy is relatively higher than SHGC, which is 99.48% compared to 99.05% at an SHGC group size of 50. The higher IDS value from [15] analysis does not always represent better classification accuracy, which leads to wrong response action set ($\{a\}$). The response value curves of these three algorithms are shown in Figure 7 for a group size of 50.

The expected lost rate of an SHGC algorithm at a group size of 50 is the lowest among test cases, which has a minimum expected lost rate at 0% for $\frac{C}{L} >= 0.78$. It is because when the system occupies best response action set at $\frac{C}{L} > 0.78$ then the minimum lost is only influenced by the value of false in false positive which is zero. However, as the false negative rate of MMDC is slightly higher than SHGC, then the expected lost rate is slightly higher for $\frac{C}{L} < 0.79$. And from the expected lost rate curves, it visually seems that CFSC has lower expected lost rate for $\frac{C}{L} < 0.79$. The lost rate of classification outputs is depicted in Figure 8. From all result of expected lost rate, our proposal has shown better expected lost rate in every test case. However, as the number of false in true positive is very small compared to overall data, then the lost consequence is remained unnoticed in the scale of 10^{-3}%.

In this paper, we only describe the autonomous response action set as a common process. But still, this proposal has not managed to decide which is possible response action set specified to specific cost to lost ratio ($\frac{C}{L}$). It remains an open problem in this report. This is because each response action in a set has a different response cost in a different environment. And up until now, there is still no proposal to describe the specific response action related to specific response cost to lost ratio. In [6], specified cost related to dollar cost in investment and op-

erational process are operated in a cost-based analysis. In [21], the response actions optimal strategy was identified by decision weight and decision sequence in the analytical hierarchy process. This approach has developed optimal strategy selection analysis, but still, has not mentioned the appropriate lost if the strategy was not deployed. For example, in report and alert process, it certainly takes response cost but has no effect on the targeted attack. Research in [9] has proposed a taxonomy of response actions for a specific case in a relational database. The response action set was divided into three categories which are conservative, fain-grained and aggressive. In [27], the time processing costs of request packets were analyzed by implementing DoS rate limiting process in Linux Click router.

6 Conclusions

This paper has proposed ARS based on cost-based decision analysis for multi-class DoS classification. Our cost-based decision analysis takes beneficial of classification output, which leads to related consequences of every possible response. The false-aware analysis is done by considering the possibility of wrong response in decision analysis. Our proposed system provides a quantitative calculation of response value which is used to estimate the best response action set autonomously. In low accuracy of classification output, our false-aware decision analysis provides more precise estimate of response value and expected lost rate than traditional cost-based analysis. It can accurately differentiate the classification output quality with the existence of false in true positive. Result regarding response value and expected lost rate have validated using synthetic test-case, and tested by the used of a well-establish KDD Cup 1999 DoS/DDoS attack dataset.

In this study, ARS is developed based on the classification algorithm output. Later, classification results can be exchanged between ARS and forming collaborative multi-agent system. It looks promising because, by information exchange between agents, ARSs can form collaborative ARSs that can classify, do decision analysis, evaluate, and ultimately overcome the attacks on the network. The exploration of the cost of the different response action also a part of our future research as it will beneficial for the cost-based response system.

References

[1] R. H. Dong, D. F. Wu, and Q. Y. Zhang, "The integrated artificial immune intrusion detection model based on decision-theoretic rough set," *International Journal of Network Security*, vol. 19, no. 6, pp. 880–888, 2017.

[2] Y. Farhoui, "Design and implementation of an intrusion prevention system," *International Journal of Network Security*, vol. 19, no. 5, pp. 675–683, 2017.

[3] A. Fawaz, R. Berthier, and W. H. Sanders, "A response cost model for advanced metering infrastructures," *IEEE Transactions on Smart Grid*, vol. 7, no. 2, pp. 543–553, 2016.

[4] C. Fung and R. Boutaba, *Intrusion Detection Network: A Key to Collaborative Security*, Boca Raton, Florida: CRC Press, 2014.

[5] C. Fung and Q. Zhu, "Facid: A trust-based collaborative decision framework for intrusion detection networks," *Elsevier Ad Hoc Networks Journal*, vol. 53, pp. 17–31, 2016.

[6] C. Iheagwara, A. Blyth, and M. Singhal, "Cost effective management frameworks for intrusion detection system," *Journal of Computer Security*, vol. 12, no. 5, pp. 777–798, 2004.

[7] S. Jin, D. S. Yeung, and X. Wang, "Network intrusion detection in covariance feature space," *Pattern Recognition*, vol. 40, no. 8, pp. 2185–2197, 2007.

[8] P. Kabiri and A. A. Ghorbani, "Research on intrusion detection and response: A survey," *International Journal of Network Security*, vol. 1, no. 2, pp. 84–102, 2005.

[9] A. Kamra and E. Bertino, "Design and implementation of an intrusion response system for relational databases," *IEEE Transactions on Knowledge and Data Engineering*, vol. 23, no. 6, pp. 875–888, 2011.

[10] Y. B. Leau and S. Manickam, "A cost-sensitive entropy-based network security situation assessment model," *Advanced Science Letters*, vol. 22, no. 10, pp. 2865–2870, 2016.

[11] W. Lee, W. Fan, M. Miller, S. Stolfo, and E. Zadok, "Toward cost-sensitive modeling for intrusion detection and response," *Journal of Computer Security*, vol. 10, no. 1, pp. 5–22, 2002.

[12] Z. Ling, J. Luo, K. Wu, W. Yu, and X. Fu, "Torward : Discovery, blocking, and traceback of malicious traffic over tor," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 12, pp. 2515–2530, 2015.

[13] K. Malialis, S. Devlin, and D. Kudenko, "Distributed reinforcement learning for adaptive and robust network intrusion response," *Connection Science*, vol. 27, no. 3, pp. 234–252, 2015.

[14] R. Mitchell and I.-R. Chen, "Effect of intrusion detection and response on reliability of cyber physical systems," *IEEE Transactions on Reliability*, vol. 62, no. 1, pp. 199–210, 2013.

[15] A. Orfila, J. Carbo, and A. Ribagorda, "Autonomous decision on intrusion detection with trained BDI agents," *Computer Communications*, vol. 31, pp. 1803–1813, 2008.

[16] E. Popoola, A. O. Adewumi, "Efficient feature selection technique for network intrusion detection system using discrete differential evolution and decision," *International Journal of Network Security*, vol. 19, no. 5, pp. 660–669, 2017.

[17] Y. Purwanto, Kuspriyanto, Hendrawan, and B. Rahardjo, "Traffic anomaly detection in DDoS flood-

ing attack," in *Proceeding of 8th International Conference on Telecommunication Systems Services and Applications (TSSA'14)*, pp. 1–6, Oct. 2014.

[18] Y. Purwanto, Kuspriyanto, Hendrawan, and B. Rahardjo, "Multistage process to decrease processing time in intrusion prevention system," in *Proceeding of 3rd International Conference on Wireless and Telematics*, Palembang, Indonesia, July 2017.

[19] Q. S. Qassim, A. M. Zin, and M. J. A. Aziz, "Anomalies classification approach for network-based intrusion detection system," *International Journal of Network Security*, vol. 18, no. 6, pp. 1159–1172, 2016.

[20] V. M. Shah, A. K. Agarwal, "Reliable alert fusion of multiple intrusion detection systems," *International Journal of Network Security*, vol. 19, no. 2, pp. 182–192, 2017.

[21] M. Sun and Y. Guo, "The research on enhanced cost-based auto intrusion response decision," in *Proceeding of International Conference on Wireless Communications*, pp. 4550–4553, Sept. 2009.

[22] Z. Tan, A. Jamdagni, X. He, P. Nanda, and R. P. Liu, "A system for denial of service attack detection based on multivariate correlation analysis," *IEEE Transaction on Parallel and Distributed Systems*, vol. 25, no. 2, pp. 447–456, 2013.

[23] C. F. Tsai and C. Y. Lin, "A triangle area based nearest neighbors approach to intrusion detection," *Pattern Recognition*, vol. 43, pp. 222–229, 2010.

[24] UCI KDD, *KDD Cup 1999 Data*, Information and Computer Science University of California, Irvine, Oct. 28, 1999. (`http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html`)

[25] J. W. Ulvila and J. John E. Gaffney, "A decision analysis method for evaluating computer intrusion detection systems," *Decision Analysis*, vol. 1, no. 1, pp. 35–50, 2004.

[26] J. W. Ulvila and J. John E. Gaffney, "Evaluation of intrusion detection systems," *Journal of Research of NIST*, vol. 108, no. 6, pp. 453–473, 2003.

[27] X. Yang, D. Wetherall, and T. Anderson, "Tva : A dos-limiting network architecture," *IEEE/ACM Transactions on Networking*, vol. 16, no. 6, pp. 1267–1280, 2008.

[28] G. Yao, J. Bi, and A. V. Vasilakos, "Passive ip traceback: Disclosing the locations of ip spoofers from path backscatter," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 3, pp. 471–484, 2015.

[29] S. A. Zonouz, H. Khurana, W. H. Sanders, and T. M. Yardley, "Rre : A game-theoretic intrusion response and recovery engine," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 2, pp. 395–406, 2014.

Biography

Yudha Purwanto completed his undergraduate degree at STTTelkom, Bandung and master degree at Electrical Engineering, Institut Teknologi Bandung. He currently work as a lecturer at Telkom University in Bandung. His research interests is security system esspecially in network security and cryptography.

Kuspriyanto completed his undergraduate degree at Electrical Engineering, Institut Teknologi Bandung in 1974. He received his Master and Doctoral degree from Universit des Sciences et Techniques de Montpellier (USTL) France. He is currently a Full Professor at the Department of Electrical Engineering, Institut Teknologi Bandung, Indonesia. His current research interests include real time computing systems, computer architecture, and robotics. Contact at kuspriyanto@lskk.ee.itb.ac.id.

Hendrawan is an Associate Professor in School of Electrical Engineering and Informatics, Institut Teknologi Bandung, Indonesia. He completed undergraduate degree at Electrical Engineering, Institut Teknologi Bandung, Master and Doctoral degree in Telecommunications and Information Systems from University of Essex, UK. Contact at hend@stei.itb.ac.id.

Budi Rahardjo completed undergraduate degree at Electrical Engineering, Institut Teknologi Bandung. And received his Master and Doctoral degree from Manitoba University, Canada. His current research interests include network security, forensic and cryptography. Contact at rahard@lskk.itb.ac.id.

Mutual Information-based Intrusion Detection Model for Industrial Internet

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

School of Computer and Communication, Lanzhou University of Technology
No. 287, Lan-Gong-Ping Road, Lanzhou 730050, China
(Email: zhangqylz@163.com)
(Received May. 25, 2017; revised and accepted August 26, 2017)

Abstract

High dimension, redundancy attributes and high computing cost issues usually exist in the industrial Internet intrusion detection field. For solving these problems, the mutual information-based intrusion detection model for industrial Internet was proposed. Firstly, by using features selection method based on mutual information, the attributes set was reduced and traffic characteristics vector was established. Secondly, the normal and abnormal traffic characteristics maps were obtained via the traffic characteristics map technology based on multi correlation analysis. Finally, with the using of discrete cosine transform and nonnegative matrix factorization, we can produce normal and abnormal hash digest, which were used to produce intrusion detection rules. To verify the effectiveness of this model, we adopt NSL-KDD data as the experimental data. The experimental results show that, by using the features selection approach based on mutual information, the proposed model has good classification accuracy and gets good detection performance.

Keywords: Intrusion Detection; Mutual Information; NSL-KDD Data; Perceptual Hash; Traffic Characteristics Map

1 Introduction

With the appearances of Industrial Internet and Industrial 4.0, industrial control area attaches more and more attentions from researchers [25]. Nowadays, big data, cloud computing and Internet of things have been the kernel techniques for the national critical infrastructures [13]. Furthermore, industrial control network security gradually transformed into the Industrial Internet security [26]. Looking through these major security issues happened in USA, Israel, German and Poland in 2016, it is conclude that network intrusion in industrial control area is powerful destructive and complex, which can launch attacks widely [4].

Aiming at the security problems in Industrial Internet, the existing solutions mainly have two kinds. One is to build a passive defense line via firewall, information encryption and user authentication. Another is to establish an active defense line by the intrusion detection method or system. In the research of Industrial Internet intrusion detection problems, researchers adopt classification, clustering, information theories and mathematical statistics four classes of approaches to deal with intrusion detection problems [12, 19]. The recent approaches include Naive Bayes [6, 15], Bayes Tree [7], Hidden Bayes [7], SVM [16, 17], least square SVM [2], artificial neural network [10], neural network with random weight [3], accelerated deep neural networks [18], artificial immune [8].

Meanwhile, many researchers made great contributions in the major of standard experimental data set and preprocessing works. Many researches choose NSL-KDD [24] standard experimental data to validate the detection performance. What's more, features selection method, as the pre-processing operations, can be used in features selection and dimension reduction of data. By this way, the computing cost and time cost of intrusion detection can be reduced. And the detection performance can also be improved obviously. The recent features selection methods include correlation based on features selection method [6], linear discriminant analysis [15], fast correlation based filter [7], mutual information [2], rough set, decision-theoretic rough set [8], information gain [14], gain ratio [14].

In the Industrial Internet intrusion detection issues, there exist high dimension of data, redundancy attributes, high computing cost problems. Aiming at these above problems, the research works about intrusion detection in Industrial Internet and features selection methods were finished in this paper. Comparing the classification accuracy among information entropy, information gain, decision-theoretic rough set and mutual information four methods, the features selection method based on mutual information got the highest accuracy. In the perceptual hash intrusion detection for Industrial Internet, the dy-

namic feedback mechanism was added. And the NSL-KDD standard experimental data was used in the validation experiments.

The rest of the paper is organized as follows. In Section 2, the features selection method, standard data set and image perceptual hash features extraction approach three aspects are introduced in the related works. The problem statement and preliminaries parts including the research problems and related theories are illustrated in Section 3. In Section 4, we present the intrusion detection model based on mutual information for industrial Internet. The results and performances of the proposed model are analyzed in Section 5. Finally, we conclude our paper in Section 6.

2 Related Works

The related works are mainly about features selection method, standard experimental data and image perceptual hash features extraction approaches. The amount of experimental data is continually increasing, which directly affect the detection performance, efficiency, and accuracy of IDS. Considering these issues, features selection method is adopted in the pre-processing operations.

In [6], the supervised classification algorithm based on Naive Bayes was used to establish intrusion detection system. By using correlation based on features selection method (CFS), the correlation between different attributes and the relation between attributes and class were analyzed. Therefore, the redundant attributes and irrelevant attributes were removed. In [15], two-tier network intrusion detection model based on machine learning was introduced. Adopting Naive Bayes and K nearest neighbor method, the intrusion detection classifier was established. The linear discriminant analysis (LDA) was utilized to achieve features selection missions. This method got high detection rate for U2R and R2L two kinds of attacks. In [7], intrusion detection classifiers were established by machine learning algorithms and preprocessing operations. By using fast correlation based filter (FCBF), the vital attributes were selected. The research utilized Naive Bayes, Hidden Naive Bayes and Naive Bayes Tree to classify the normal and abnormal traffic records. The high computing payload of FCBF decreased the efficiency of intrusion detection. In [16], the intrusion detection method based on SVM was researched. The redundant attributes were deleted via filter method. And the vital attributes were selected to build attributes set, which obviously decreased the computing cost of IDS. However, the setting of threshold in filter and classification accuracy still needs optimization. In [2], the least square support vector machine (LSSVM) was used to establish intrusion detection system. The experimental results show that, this method obtained high detection rate and low computing cost. In [14], adopting embedded filters to build intrusion detection model in cloud platform, information gain, gain ratio, Chi-square and feature weighting algorithm were used to select vital attributes. Yet, the structure of features selection method is complex and with high computing payload. In [21], the correlation-based feature selection for intrusion detection system was presented. Introducing the correlation-based feature selection matrices and symmetrical uncertain matrices, the correlation between attributes and classes were analyzed. And different classification approaches were used to validate the accuracy of features selection method. In [11], the embedded SVM and non-linear projection techniques were used to achieve the classification and detection of abnormal intrusions. With the help of linear and non-linear dimension reduction methods, 5 kinds classifiers were produced. The research chose NSL-KDD data to test the detection performance of the proposed method.

The NSL-KDD data is an improvement of KDD Cup 99 data. In many recent researches, the NSL-KDD data is used to validate the performance of the proposed method or model. And it is widely used in the research of Industrial Internet intrusion detection problems.

In [10], the artificial neural network (ANN) was used to analyze the performance of NSL-KDD data. And the dimension reduction is obvious after the information entropy, information gain and correlation analysis features selection operations. In [3], the neural network with random weights (NNRw), a semi-supervised learning algorithm, was proposed. The non-iterative neural network model was trained by the randomize method. By using fuzzy variable, the unsigned record can be classified. This method had better learning ability and computing efficiency, but the classification accuracy need to be improved. In [18], the intrusion detection system was established via deep neural network (DNN). The DNN has forward transmission and back forward transmission features. By this way, large amount of data characteristics obtained from training data were used to establish intrusion detection model and classifiers. However, the proposed approach highly depended on the hardware in the platform and the computing cost is high. In [5], a varying chaos particle swarm optimization approach (TVCPSO) was presented. And the intrusion detection model based on SVM was proposed. The chaos particle, time varying inertia coefficient and time varying acceleration coefficient three conceptions were introduced. The local optimum solutions problems of particle swarm algorithm were effectively resolved. And, the weight object function was utilized to balance the max true positive rate and the min false positive rate. In [20], utilizing four frequently used classification methods in the NSL-KDD data, the imbalance problem of the data was analyzed. The experimental results show that the NSL-KDD data, as the standard test data, can be used to validate the detection performance of intrusion detection for Industrial Internet. In [22], an intrusion detection system based on neural network was proposed. Adopting the feed forward and reverse methods, combining with optimal technique, the computing payload of intrusion detection method was decreased. The NSL-KDD data was chosen as the exper-

imental data. In [1], using a colony optimization method, the effective and key features were selected, which improved the performance of intrusion detection system.

The intrusion detection method for Industrial Internet based on perceptual hash is a new and effective method from the point of image [8]. The recent image perceptual hash features extraction method includes discrete cosine transform (DCT), discrete wavelet transform (DWT), singular value decomposition (SVD), non-negative matrix factorization (NMF) and local binary pattern (LBP). The intrusion detection method based on perceptual hash has robustness and discrimination, which maintains the detection performance. It is proofed that the proposed method is short time consuming. The NMF method needs less storage and it sensitive to the local features in the traffic characteristics map. Therefore, by using DCT and NMF methods, hash digest are produced.

3 Problem Statement and Preliminaries

The proposed model includes three parts: features selection method based on mutual information, traffic characteristics map technique and improved image perceptual hash intrusion detection method.

3.1 Features Selection Method Based on Mutual Information

The attributes redundancy and attributes irrelevant issues existing in intrusion detection lead to the low classification accuracy, high computing cost and time consuming problems. Taking into account of these problems, the feature selection method based on mutual information [9] is adopted to select vital attributes set and reduce the high dimension of data. The produced attributes set is the input data to the traffic characteristics map technique. And the information entropy and decision-theoretic rough set are forward features selection approaches. These methods have low computing cost but without considering the relationships between each attribute. In the feature selection method based on mutual information, the correlation and redundancy concepts are used to describe the correlations between attributes.

Assume that the total number of the experimental data is N. Every traffic record includes m attributes described as $\{f_1, f_2, \ldots, f_m\}$. $P(f_t)$ is the corresponding probability when f_t get different values. And the information entropy can be defined as

$$H(f_t) = -\sum_{f_t} P(f_t) \log P(f_t),$$

where $H(f_t)$ is the information entropy of attribute f_t, $0 \leq P(f_t) < 1$. When the value of f_t is known, the uncertainty of attribute f_t can be described with condition information entropy, which can be defined as

$$H(\frac{f_t}{f_i}) = -\sum_{f_i} P(f_i) \sum_{f_t} P(\frac{f_t}{f_i}) \log P(\frac{f_t}{f_i}),$$

where $H(f_t/f_i)$ expresses the condition entropy of f_t with the condition of f_i. $P(f_t/f_i)$ is the condition probability of the corresponding attribute, $0 \leq P(f_t/f_i) < 1$. According to information entropy and condition entropy concepts, the mutual information can be defined as

$$I(f_t; f_i) = I(f_i; f_t) = H(f_t) - H(\frac{f_t}{f_i}),$$

where $I(f_t; f_i)$ is the mutual information value. And $I(f_t; f_i)$ equal to $I(f_i; f_t)$. The average mutual information is the mean value of mutual information between attribute f_i and every possible attribute f_t, $t \in [1, 41]$. The average mutual information can be defined as

$$ave\_MI(f_i) = \frac{1}{m} \sum_{t=1}^{m} I(f_i; f_t), \tag{1}$$

where $ave\_MI(f_i)$ is the average mutual information of attribute f_i.

Definition 1. *(Correlation Degree) The correlation degree of attribute f_i is the average mutual information of f_i. The correlation degree can be defined as*

$$Rel(f_i) = \frac{1}{m} \sum_{t=1}^{m} I(f_i; f_t),$$

where $Rel(f_i)$ is the correlation degree of f_i, it is average mutual information, $m = 41$.

Definition 2. *(Condition Correlation Degrees) In the condition of attribute f_i, condition correlation degrees of attribute f_t can be defined as*

$$Rel(\frac{f_t}{f_i}) = \frac{H(\frac{f_t}{f_i})}{H(f_t)} Rel(f_t),$$

where $Rel(f_t/f_i)$ is the condition correlation degrees of attribute f_t in condition of f_i. $Rel(f_t)$ is the correlation degree of attribute f_t.

Definition 3. *(Redundancy Degrees) The redundancy degrees between attribute f_i and f_t can be expressed as*

$$Red(f_i; f_t) = Rel(f_t) - Rel(\frac{f_t}{f_i}),$$

where $Red(f_i; f_t)$ is the redundancy degrees between attribute f_i and attribute f_t. $Rel(f_t)$ is the average mutual information of f_t, and $Rel(f_t/f_i)$ is the condition correlation degrees of f_t in condition of f_i.

In the features selection method based on mutual information, the significance of attribute can express the importance of the waiting selecting attributes in attributes set $\boldsymbol{U}$. Meanwhile, the most significant attribute can be

added into selected attributes set $\boldsymbol{S}$. The significance of attributes can be expressed as

$$UmRMR(f_i) = Rel(f_i) - \max_{f_t \in S_m - 1}\{Red(f_i; f_t)\}, \quad (2)$$

where $Rel(f_i)$ is the correlation degree of attribute f_i. $Red(f_i; f_t)$ is the redundancy degree between attribute f_i and f_t. And attributed f_t belongs to selected attributes set $\boldsymbol{S}$. Every time, the max significance of attribute f_i can be added into selected set $\boldsymbol{S}$.

Algorithm 1 Features Selection method based on Mutual Information

1: Input: Experimental data train-set and the number of the selecting features $\boldsymbol{K}$
2: Output: The selected features set $\boldsymbol{S}$
3: Initialize the features selected set $\boldsymbol{S} = \emptyset$, store selected attributes.
4: Initialize the features selecting set $\boldsymbol{U} = \{f_1, f_2, \ldots, f_m\}$, $m \in [1, 41]$.
5: According to Equation (1), computing average mutual information of every attribute.
6: Choose the max average mutual information of attribute f_i, add attribute f_i to the selected set $\boldsymbol{S}$, and delete attribute f_i in the selecting set $\boldsymbol{U}$.
7: **while** the number of selected features $< \boldsymbol{K}$ **do**
8: According to Equation (2), compute the significance of every selecting attribute f_i
9: Choose the most vital attribute f_i, add f_i to set $\boldsymbol{S}$ and delete f_i in the set $\boldsymbol{U}$
10: **end while**
11: Output the selected attributes $\boldsymbol{S}$

After the features selection method, the selected features set is the input for the traffic characteristics map technique.

3.2 Traffic Characteristics Map Technique

The traffic characteristics map technique is based on multi correlation analysis (MCA) [23]. By computing the triangle area mapping, the correlation information between attributes in normal and abnormal network traffics. By this way, the text network traffic records with $1 \times m$ vector format can be transformed into $m \times m$ network traffic matrices. And $m = 14$ is the feature selection results. The traffic characteristics map includes correlation between each attributes.

The experimental data is $X = \{x_1, x_2, \ldots, x_n\}$, and according to the selected features set, the $i-$th traffic record is $x_i = [f_1^i, f_2^i, \ldots, f_m^i]$, $(1 \le k \le m)$. The correlation between $j-$th attribute and $k-$th attribute can be computed by triangle area.

The vector x_i can map into the $(j-k)$ two-dimension Euclidean subspace, $y_{i,j,k} = [\varepsilon_j \varepsilon_k]^T = [f_j^i f_k^i]^T$, $(1 \le i \le n, 1 \le j \le m, 1 \le k \le m, j \ne k)$. Variable $\varepsilon_j = [\varepsilon_{j,1}, \varepsilon_{j,2}, \ldots, \varepsilon_{j,n},]^T$, where $e_{j,j} = 1$, $e_{k,k} = 1$, other elements is 0. $y_{i,j,k}$ is two-dimension column vector, which is the point $(f_j^i f_k^i)$ of $(j-k)$ two-dimension Euclidean subspace in Descartes coordinate system. Then, in Descartes coordinate system, connecting the origin with the point f_j^i mapping in j coordinate axis and point f_k^i mapping in k coordinate axis, the triangle area is obtained, named $\Delta f_j^i O f_k^i$. The triangle area is marked as $Tr_{j,k}^i$.

$$Tr_{j,k}^i = (\parallel (f_j^i, 0) - (0,0) \parallel \times \parallel (0, f_k^i) - (0,0) \parallel)/2,$$

where $1 \le i \le n$, $1 \le j \le m$, $1 \le k \le m$, and $j \ne k$. The complete triangle area mapping of a network traffic record includes the triangle area of every pairs of attributes. $Tr_{j,k}^i$ is $j-$th row and $k-$th column. When $j = k$, $Tr_{j,k}^i = 0$. The correlation between different attributes is the key point. The symmetric matrix $\boldsymbol{TAM}$ can be got. For example, the $4-$dimension $\boldsymbol{TAM}$ is shown.

$$TAM_x^i = \begin{bmatrix} 0 & Tr_{1,2}^i & Tr_{1,3}^i & Tr_{1,4}^i \\ Tr_{2,1}^i & 0 & Tr_{2,3}^i & Tr_{2,4}^i \\ Tr_{3,1}^i & Tr_{3,2}^i & 0 & Tr_{3,4}^i \\ Tr_{4,1}^i & Tr_{4,2}^i & Tr_{4,3}^i & 0 \end{bmatrix}$$

3.3 Image Perceptual Hash Intrusion Detection Method

The image perceptual hash features extraction based on DCT and NMF is utilized to produce normal and abnormal hash digests. Meanwhile, after the features selection operation, the selected features is $m = 14$. Therefore, the 14×14 traffic characteristics map is established.

In the preparing stage of perceptual hash features extraction method, DCT coefficient is computed.

$$F(u,v) = \frac{C(u)C(v)}{4} \sum_{x=0}^{m} \sum_{y=0}^{m} f(x,y) \bullet \cos(\frac{\pi(2x+1)u}{m^2}) \cos(\frac{\pi(2y+1)v}{m^2}) \quad (3)$$

where f is $m \times m$ image pixel point and F is $m \times m$ DCT coefficient matrix. $\boldsymbol{C}$ is cosine coefficient matrix.
Algorithm descriptions:

1) After the features selection method, 14×14 pixel traffic characteristics map is obtained by using traffic characteristics map technique.

2) According to Equation (3), 14×14 coefficient matrix can be computed by DCT.

3) In order to extract the local features in the map and get the discriminative perceptual hash digest, the low-frequency region of coefficient matrix is rebuilt by NMF. And the local saliency information of map is extracted. By NMF, the DCT coefficient matrix

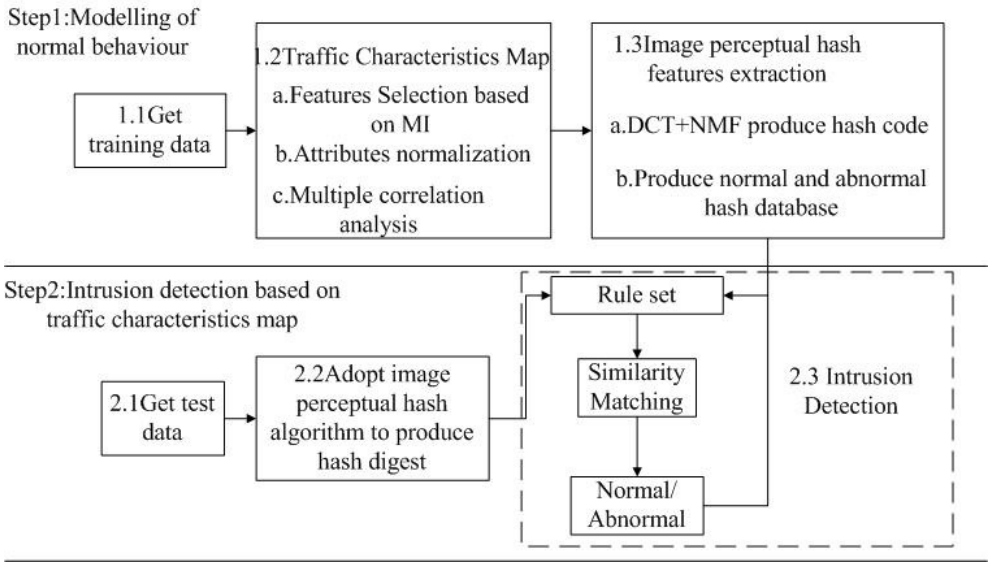


Figure 1: Intrusion Detection model based on mutual information for Industrial Internet

can factorize into basis matrix $\boldsymbol{W}$ and weights coefficient matrix $\boldsymbol{H}$.

$$DCT\_Coefficient = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{bmatrix} \begin{bmatrix} h_1 & h_2 & \cdots & h_m \end{bmatrix},$$

where $\boldsymbol{DCT\_Coefficient}$ is the DCT coefficient matrix. $\boldsymbol{W}$ is the basis matrix and $\boldsymbol{H}$ is the weights coefficient matrix. Each column vector sum in $\boldsymbol{W}$ is the column vector of DCT coefficient matrix. The matrix factorization vector $[W'H]$ is established and the mean value of $[W'H]$ is computed.

$$NMF\_matrix = [W'Hmean].$$

4) According to the hash rules, the hash digest can be produced. The hash rule is defined as

$$h(x_{i+1}) = \begin{cases} 1 & x_{i+1} > x_i \\ 0 & x_{i+1} \leq x_i \end{cases}$$

where the $(i+1)$th x_{i+1} is got. When $x_{i+1} > x_i$, the hash code of x_{i+1} is 1, or 0, $1 \leq i \leq 29$. The length of hash code is 28 bit with the binary form. The normal and abnormal hash digest base is established and the corresponding intrusion detection rules are extracted.

In the hash matching phase, the normalization Hamming distance is used to measure the similarity between different hash digest. The normalization Hamming distance is defined as

$$D_H(H_{s1}, H_{s2}) = \frac{1}{L} \sum_{w=1}^{L} \mid H_{s1}(w) - H_{s2}(w) \mid, \qquad (4)$$

where H_{s1} and H_{s2} are two hash digest, whose length are 28 bit. w is one bit in the hash digest. When the hash matching threshold is set, if the hash similarity is above threshold, it is abnormal or normal. The joint threshold is 0.15.

4 The Proposed Method

The flow works of the intrusion detection model based on mutual information for Industrial Internet (IDM-MI) is shown in Figure 1. The method is divided into preparing phase and intrusion detection phase based on Industrial Internet.

As the Figure 1 shown, in the preparing phase, the numerical operation is achieved. The features are reduced via features selection method based on mutual information. Then the normalization operation is finished. By using MCA, the traffic characteristics map is produced. With the using of DCT and NMF, the hash digests are extracted to establish normal and abnormal hash digest base. And the intrusion detection rule is also produced.

In the intrusion detection phase, the numerical operation is finished and the vital features set are extracted. By MCA, the traffic characteristics map of test record is produced. And the hash code is produced via DCT and NMF. Adopting normalization Hamming distance, the similarity between test hash and hash digest base is measured. If the similarity is lower than threshold, it's normal or abnormal.

The scale of the training data is N_1, and the scale of the test data is N_2. The original number of attributes is M. After features selection operation, the number features is M_1. The number abnormal hash digest is t_1 and the abnormal hash digest is t_2. The length of hash digest is L. According to the analysis of the algorithm flows, the

Algorithm 2 IDM-MI

```
Input: The standard NSL-KDD data train-data and test-data
Output: Intrusion detection result
Obtain the train-data
while the number of train-data > 0 do
    Adopt MCA method to produce traffic characteristics map
    Utilize DCT and NMF methods to extract hash digest of normal and abnormal network traffic
    Establish the intrusion detection rule set
    Number −−
end while
Obtain test-data
while the number of test-data > 0 do
    Adopt MCA method to produce traffic characteristics amp
    Using DCT and NMF method to produce hash code
    According to Equation (4), compute the similarity between hash code
    if similarity < threshold then
        The record is normal
    else
        The record is abnormal
    end if
    Number −−
end while
```

time complexity is $O((N_1 + N_2)(M_1^2 + 5M_1))$.

5 Experimental Results and Analysis

5.1 Preparing for the Simulation Experiment Environment

The experiments were carried out by a ThinkPad computer with 2.5 GHz quad-core i5-3210M and 8GB of RAM. The operate system is windows 7, 64bits and the simulation platform is Matlab R2013a.

The NSL-KDD [24] standard data set was chose to validate the performance of the proposed model, which is improved from the KDD Cup 99 data set. To keep the effective evaluation for the intrusion detection methods, the percentage of each kinds of data are same with the KDD Cup 99. Some intrusion detection methods only detect part of the repeated data effectively. Therefore, the redundant records in the training data set and the repeated records in the test data set were removed. In order to decrease the running payload, the number of the data is reasonable. So, the NSL-KDDTrain+_20Persent and NSL-KDDTest-21 were selected to take part in the experiments, which are named as Data1 and Data2 respectively. The details of the experimental data are shown as Table 1.

Firstly, the numerical operations of the training data and test data were finished. The protocol-type, service, flag and attack four attributes experienced numerical operations. Secondly, according to the label of the attack, the training data were classified into normal and abnormal. Label $\{1\}$ is normal and Label $\{2, 3, 4, 5\}$ are abnormal. Finally, the normalization of the training data and test data were achieved. The data is normalized into $[0, 255]$.

Table 1: The details of the experimental data

Name	Total	Normal	Dos	Probe	U2R	R2L
Data1	25192	13449	9234	2289	11	209
Data2	22544	9711	7458	2421	200	2754

$$f(x) = \begin{cases} 0 & x \in [0, \min) \\ \frac{255x}{\max - \min} & x \in [\min, \max] \\ 255 & x \in (\max, \infty) \end{cases},$$

where *max* is the max value and *min* is the min value. $f(x)$ is the normalized value.

5.2 Feature Selection Methods

According to the recent research results, the preprocessing selections of the experimental data were finished. Attributes $\{9, 20, 21\}$ take no effect on the classification. Attributes $\{15, 17, 19, 32, 40\}$ have little influence on the classification. The values of attributes $\{7, 8, 11, 14\}$ are mostly 0. The above mentioned attributes were removed. Then, by feature selection method based on mutual information, the dimension of the data was reduced.

According to the conceptions of the information entropy, condition entropy and mutual information, the average mutual information was computed. By correlation degree, condition correlation degrees and redundancy degree, the significance of the selecting attributes, named max correlation-min redundancy, was obtained. Considering the significance of the selecting attributes, the candidate attributes can be added into the selected attributes set $\boldsymbol{S}$. The number of the selected attributes is set $K = 14$. Therefore, the result of the feature selection is 14 vital attributes.

The result of the feature selection operation is $set = \{3, 5, 6, 8, 12, 23, 24, 32, 33, 34, 35, 37, 38, 39\}$. Table 2 displays several recent features selection approaches and compares their accuracy of the classification.

From Table 2, when the number of the attributes is 14, the classification accuracy of mutual information is 0.9940, which is the max value. So, the mutual information features selection method is used to reduce the dimension of data. However, the time consuming and computing cost of this method are still need to be optimal. It's difficult to balance the time cost and classification accuracy, which attach more attentions from the researchers. In the next work, this problem is still the research emphasis.

Table 2: The details of the experimental data

Method	Features selection results	Accuracy
Information Entropy	3,5,6,23,24,29,30,31,32,33,34,35,36,37	0.9587
Information Gain	3,4,5,6,12,23,25,29,30,33,34,35,38,39	0.9744
Decision-Theoretic Rough Set	2,3,4,5,6,8,12,17,27,28,30,31,36,37	0.9865
Mutual Information	3,5,6,8,12,23,24,32,33,34,35,37,38,39	0.9940

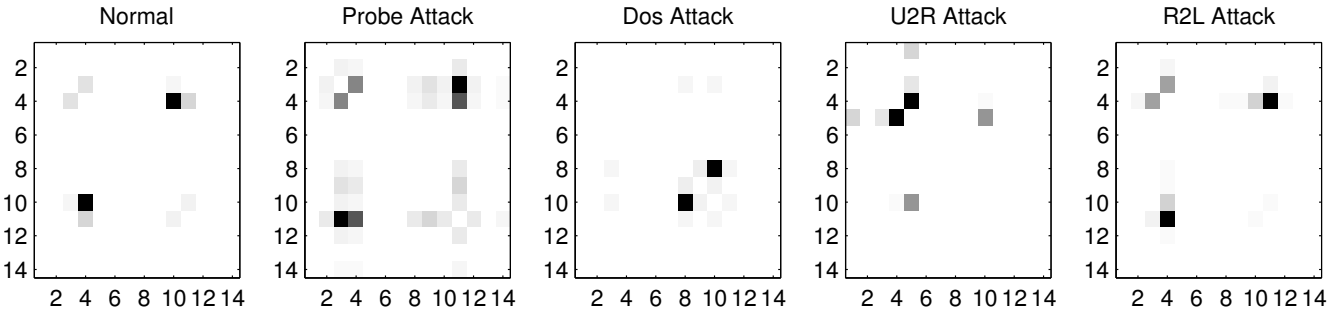


Figure 2: The traffic characteristics map of NSL-KDD training data set

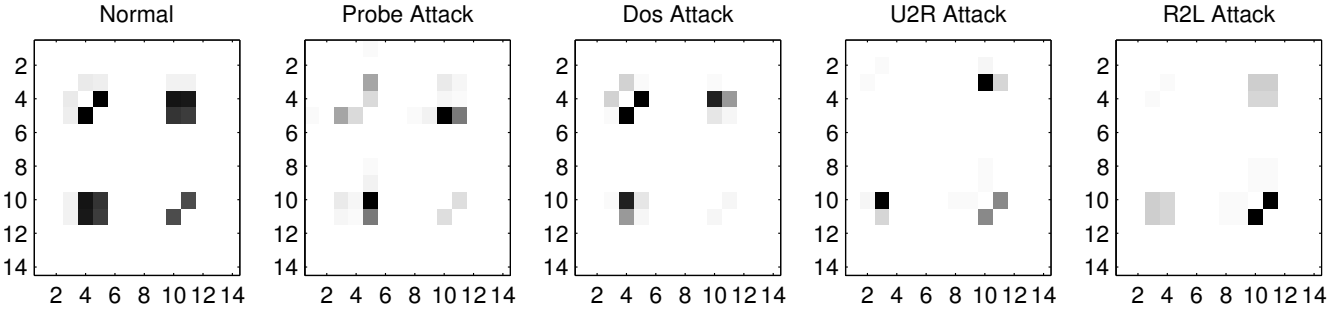


Figure 3: The traffic characteristics map of the NSL-KDD test data set

By utilizing image perceptual hash features extraction method, the hash code of the normal and abnormal network traffic records are produced. The length of the hash code is 28 bits. The number of the normal rule is 471 and the number of the abnormal rule is 535.

5.3 Traffic Characteristics Map

After the features selection operations, the features space is reduced. Then, the traffic characteristics map is produced via traffic characteristics technique, as the Figure 2 and 3 shown. The size of the selected features is 14, and the traffic characteristics map is 14×14 matrix.

In Figure 2, 5 kinds of records in the training data are shown. According to the results of the features selection method, the size of the map is 14×14. The difference between every map is obvious. These maps are the input data for the next operation.

As the Figure 3 shown, 5 kinds of record in test data are produced. Comparing Figure 2 with Figure 3, the features of the image is obvious. The discrimination is good. In the 14×14 area, the almost same place exist some pixel blocks which have different grey value.

5.4 Discriminative Experiments

The intrusion detection method based on mutual information for industrial Internet has robustness and discrimination which keep the performance and efficiency of the intrusion detection. The robustness ensures that the same normal and abnormal records produce same hash code. And the discrimination ensures that the different records produce different hash code. With the help of robustness and discrimination, the proposed model can detect the existing records or new records. Therefore, this model has adaptability. The false accept ratio (FAR) is selected in the discriminative experiments. The FAR can be defined as

$$FAR = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\tau} \exp[\frac{-(x-\mu)_2}{2\sigma^2}],$$

where μ is the exception mean of the normal distribution. σ is the standard deviation and τ is the matching threshold. Figure 4 is the normal distribution figure of this mode.

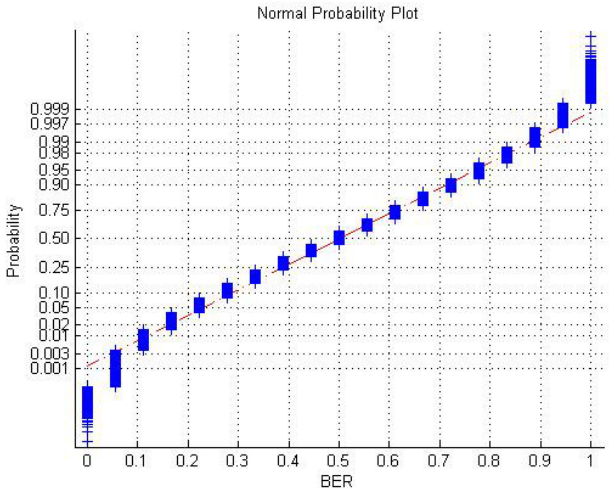


Figure 4: Normplot figure of the proposed model

The total number of the hash code in training data is 1006. So, 505515 bit error ratio (BER) can be obtained. Figure 4 is the normal distribution curve of the BER. In Figure 4, the curve is almost overlapping with the straight line of mean value. But the fluctuations also exist in the two sides of the curve. The mean value is 0.4951 and the standard deviation is 0.0945. The real standard deviation is 0.1724.

When $\tau = 0$, $FAR = 0.0020$. That is to say that, when the matching threshold is $\tau = 0$, in 1000 traffic records, there are 2 false detections, which meet the requirements of detection. The threshold of FAR is shown in Table 3.

Table 3: FAR comparing table

Threshold τ	0	0.005	0.01	0.015
FAR	0.0020	0.0022	0.0024	0.0027

As the Figure 5 shown, the hash matching of normalization Hamming distance obey to Gaussian distribution, the mean value $\mu = 0.5$, standard deviation $\mu = 0.5/\sqrt{N}$. And N is the length of hash code, $N = 28$. Figure 5 is the BER histogram obtained from the discriminative experiments. The center of the histogram is 0.4951 which is close to 0.5. The standard deviation of distribution is 0.1724.

5.5 Experiments Results and Analysis

We choose TP, FP, TN, FN and Acc as the evaluation indexes. The TP is the percentage of abnormal records correct classification in all abnormal records. The number

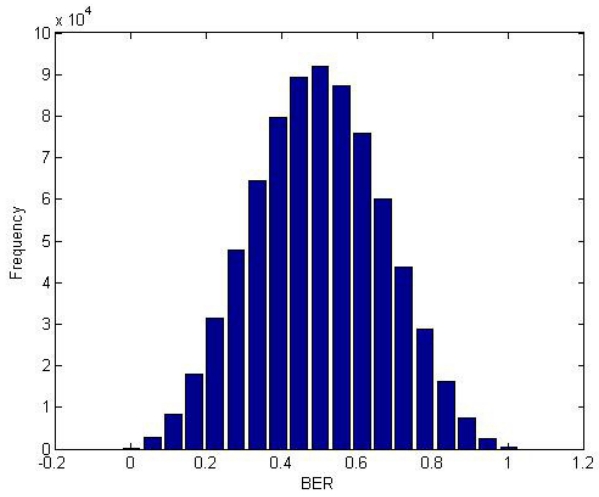


Figure 5: The BER predict histogram of image perceptual hash features extraction method

of abnormal record correct classification is $num1$, and the total number of abnormal records is N. The TP can be defined as

$$TP = \frac{num1}{N},$$

where the FN is the percentage of abnormal records false classification in all abnormal records. The number of abnormal record false classification is $num2$. The FN can be expressed as

$$FN = \frac{num2}{N},$$

where $TP + FN = 1$. The FP is the percentage of normal records false classification in all normal records. The number of normal record false classification is $num3$, and the total number of normal records is M. The FP can be defined as

$$FP = \frac{num3}{M},$$

where the TN is the percentage of normal records correct classification in all normal records. The number of normal record correct classification is $num4$. The TN can be expressed as

$$TN = \frac{num4}{M},$$

where $FP + TN = 1$. The Acc express the average detection ratio of normal and abnormal records. It is also the ratio of correct predicted records to the entire records. The Acc can be defined as

$$Acc = \frac{TP + TN}{TP + FN + FP + TN}.$$

Table 4 displays the difference with different methods in these indexes.

As Table 4 shown, the *Acc* of the proposed method is 0.9940, which is less than 0.9985 of NB Tree method. And the *FP* of the proposed method is 0.0012 which also less than 0.0020 of NB Tree approach. The *TP* is 0.9893 which is less than 0.9990 of NB Tree and 0.9916 of ANN. But, the *FP* of our method is less than 0.0036 of ANN. The *Acc* of our method is equal to ANN. The features selection method based on mutual information has better performance and good results. The perceptual hash method has a better detection efficiency and short time consuming. Therefore, the IDM-MI has a better detection performance.

Table 4: The details of the experimental data

Method	Reference	TP	FP	Acc
LSSVM-IDS	Ref. [2]	0.9893	0.0028	0.9932
ANN	Ref. [22]	0.9916	0.0036	0.9940
TVCPSO-SVM	Ref. [5]	0.9703	0.0087	0.9808
NB Tree	Ref. [6]	0.9990	0.0020	0.9985
Naive Bayes	Ref. [6]	0.9360	0.1340	0.9010
AD Tree	Ref. [6]	0.9890	0.0190	0.9850
FCBF	Ref. [7]	-	-	0.8704
SVC	Ref. [11]	0.9340	0.1400	0.8970
SVM	Ref. [16]	0.8200	0.1500	0.8350
IDM-MI	Our method	0.9893	0.0012	0.9940

6 Conclusions

An intrusion detection model based on mutual information for industrial Internet was presented. Adopting features selection method based on mutual information, the issues of high dimension of data, attributes redundancy and high computing cost were effectively resolved. The dynamic feedback mechanism was added into the intrusion detection model based on perceptual hash. When the new normal or abnormal records appearances, the new hash digest was added into the hash digest base. And the corresponding intrusion detection rule was updated. The adaptability of the proposed method was enhanced. The NSL-KDD data set was utilized to validate the efficiency and accuracy of detection. As the experimental results show that the *TP* is 0.9893, the *FP* is 0.0012 and the *Acc* is 0.9940, which proof the good performance of detection. In the future, the algorithm optimization work of our method is vital.

Acknowledgments

This work is supported by the National Natural Science Foundation of China ($No.61363078$), the Natural Science Foundation of Gansu Province of China ($No.1310RJYA004$), the Open Project Program of the National Laboratory of Pattern Recognition (NLPR) ($No.201700005$). The authors would like to thank the anonymous reviewers for their helpful comments and suggestions.

References

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

[2] M. A. Ambusaidi, X. He, P. Nanda, and Z. Tan, "Building an intrusion detection system using a filter-based feature selection algorithm," *IEEE Transactions on computers*, vol. 65, no. 10, pp. 2986–2998, 2016.

[3] R. A. R. Ashfaq, X. Z. Wang, and J. Z. Huang, "Fuzziness based semi-supervised learning approach for intrusion detection system," *Information Sciences*, vol. 378, pp. 484–497, 2017.

[4] B. Babu, T. Ijyas, P. Muneer, and J. Varghese, "Security issues in scada based industrial control systems," in *2nd International Conference on Anti-Cyber Crimes (ICACC'17)*, pp. 47–51, Abha, Saudi Arabia, Mar. 2017.

[5] S. M. H. Bamakan, H. Wang, Y. Tian, and Y. Shi, "An effective intrusion detection framework based on mclpsvm optimized by time-varying chaos particle swarm optimization," *Neurocomputing*, vol. 199, pp. 90–102, 2016.

[6] D. H. Deshmukh, T. Ghorpade, and P. Padiya, "Intrusion detection system by improved preprocessing methods and naive bayes classifier using NSL-KDD 99 dataset," in *International Conference on Electronics and Communication Systems (ICECS'14)*, pp. 1–7, Coimbatore, India, Feb. 2014.

[7] D. H. Deshmukh, T. Ghorpade, and P. Padiya, "Improving classification using preprocessing and machine learning algorithms on NSL-KDD dataset," in *International Conference on Communication, Information and Computing Technology (ICCICT'15)*, pp. 1–6, Mumbai, India, Jan. 2015.

[8] R. Dong, WD. u, and Q. Zhang, "The integrated artificial immune intrusion detection model based on decision-theoretic rough set," *International Journal of Network Security*, vol. 19, no. 6, pp. 880–888, 2017.

[9] H. Duan, Q. Zhang, and M. Zhang, "Fcbf algorithm based on normalization mutual information for features selection (in Chinese)," *Journal of Huazhong University of Science & Technology (Natural Science Edition)*, vol. 45, no. 1, pp. 52–56, 2017.

[10] B. Ingre and A. Yadav, "Performance analysis of NSL-KDD dataset using ann," in *International Conference on Signal Processing and Communication Engineering Systems (SPACES'15)*, pp. 92–96, Guntur, India, Jan. 2015.

[11] E. De la Hoz, A. Ortiz, and J. Ortega, "Network anomaly classification by support vector classifiers ensemble and non-linear projection techniques," in *8th International Conference on Hybrid Artificial*

Intelligent Systems (HAIS'13), pp. 103–111, Salamanca, SPAIN, Sept. 2013.

[12] Y. Farhoui, "Design and implementation of an intrusion prevention system," *International Journal of Network Security*, vol. 19, no. 5, pp. 675–683, 2017.

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

[14] O. Osanaiye, H. Cai, K. K. R. Choo, A. Dehghantanha, Z. Xu, and M. Dlodlo, "Ensemble-based multi-filter feature selection method for ddos detection in cloud computing," *EURASIP Journal on Wireless Communications and Networking*, vol. 1, pp. 1–10, 2016.

[15] H. H. Pajouh, G. H. Dastghaibyfard, and S. Hashemi, "Two-tier network anomaly detection model: a machine learning approach," *Journal of Intelligent Information Systems*, vol. 48, no. 1, pp. 61–74, 2017.

[16] M. S. Pervez and D. M. Farid, "Feature selection and intrusion classification in NSL-KDD cup 99 dataset employing svms," in *8th International Conference on Software, Knowledge, Information Management and Applications (SKIMA'14)*, pp. 1–6, Dhaka, Bangladesh, Dec. 2014.

[17] E. Popoola, A. O. Adewumi, "Efficient feature selection technique for network intrusion detection system using discrete differential evolution and decision," *International Journal of Network Security*, vol. 19, no. 5, pp. 660–669, 2017.

[18] S. Potluri and C. Diedrich, "Accelerated deep neural networks for enhanced intrusion detection system," in *21st International Conference on Emerging Technologies and Factory Automation (ETFA'16)*, pp. 1–8, Berlin, Germany, Sept. 2016.

[19] Q. S. Qassim, A. M. Zin, and M. J. A. Aziz, "Anomalies classification approach for network-based intrusion detection system," *International Journal of Network Security*, vol. 18, no. 6, pp. 1159–1172, 2016.

[20] S. Rodda and U. S. R. Erothi, "Class imbalance problem in the network intrusion detection systems," in *International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT'16)*, pp. 2685–2688, Chennai, India, Mar. 2016.

[21] M. B. Shahbaz, WX. ang, A. Behnad, and J. Samarabandu, "On efficiency enhancement of the correlation-based feature selection for intrusion detection systems," in *7th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON'16)*, pp. 1–7, Vancouver, BC, Canada, Oct. 2016.

[22] B. Subba, S. Biswas, and S. Karmakar, "A neural network based system for intrusion detection and attack classification," in *Twenty Second National Conference on Communication (NCC)*, pp. 1–6, Guwahati, India, Mar. 2016.

[23] Z. Tan, A. Jamdagniand, X. He, P. Nanda, and R. Liu, "A system for denial-of-service attack detection based on multivariate correlation analysis," *IEEE transactions on parallel and distributed systems*, vol. 25, no. 2, pp. 447–456, 2014.

[24] M. Tavallaee, E. Bagheri, W. Lu, and A. Ghorbani, "A detailed analysis of the KDD cup 99 data set," in *Symposium on Computational Intelligence for Security and Defense Applications (CISDA'09)*, pp. 1–6, Ottawa, ON, Canada, July 2009.

[25] A. Yang, L. Sun, X. Wang, and Z. Shi, "Intrusion detection techniques for industrial control system," *Journal of Computer Research and Development*, vol. 53, no. 9, pp. 2039–2054, 2016.

[26] B. B. Zarpelao, R. S. Miani, C. T. Kawakani, and S. C. de Alvarenga, "A survey of intrusion detection in i nternet of things," *Journal of Network and Computer Applications*, vol. 84, pp. 25–37, 2017.

Biography

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

Wu Dongfang In 2015, Wu Dongfang obtained his bachelor of engineering degree from Northwest University for Nationalities. Currently, he is studying for his masters degree at Lanzhou University of Technology. His research focuses on the industrial control network security.

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

Duan Hong-xiang lecture/PhD student of School of Computer and Communication in Lanzhou University of Technology, she received M.Sc. degree in Lanzhou University of Technology in 2011. Now she is working on feature selection of high-dimensional data in multimodal human-computer interaction and image understanding and pattern recognition.

A New Approach to Quantify Network Security by Ranking of Security Metrics and Considering Their Relationships

Mostafa Behi[1], Mohammad GhasemiGol[2], and Hamed Vahdat-Nejad[2]

(Corresponding author: Mohammad GhasemiGol)

Department of Computer Engineering, Science and Research branch, Islamic Azad University[1]
Birjand, Iran
Department of Computer Engineering, University of Birjand[2]
South Khorasan Province, Birjand, A78, 97175615, Iran
(Email: ghasemigol@birjand.ac.ir)

(Received May. 20, 2016; revised and accepted Sept. 3, 2016)

Abstract

There are several characteristics in computer networks, which play important roles in determining the level of network security. These characteristics known as security metrics can be applied for security quantification in computer networks. Most of the researches on this area has focused on defining the new security metrics to improve the quantification process. In this paper, we present a new approach to analyze and quantify the network security by ranking of security metrics with considering the relationships between them. Our ranking method reveals the importance of each security metric to quantify security in the network under surveillance. The proposed approach helps the network administrators to have a better insight on the level of network security

Keywords: Correlation; Regression; Security Quantification; Security Metrics

1 Introduction

In today's digital age, every organization, regardless of its size, must have an information security program to protect its data. This program should be designed in a way to detect, prevent and significantly reduce the risks. Developing a comprehensive information security program that recognizes these risks is one of the major issues that organizations are faced with today. Identification of incidents that has an effect on the organization's assets is one of the important parts of the security program and also a difficult task. The complexity of today's computer networks has made this issue a more complicated process. Since the budgets and resources are often limited in organizations, then a mechanism should be chosen for the right direction of those matters. In [9] Verizon reports that 97% of the attacks could be neutralized by little try because they were done by amateur attackers without so much skills and tools. In spite of the big amount of money spending on security and defense in many organizations, hackers can lower their level of security and confidentiality just by using simple exploit accessible online. Then using right minimums is much better than useless maximums in security. Installing expensive firewalls and UTMs, antiviruses, intrusion detection systems and intrusion prevention systems (IDS/IPS) would be all ineffectual in network security, if one simple task such as users' loss of knowledge about security is misunderstood. In such a case an unaware user can endanger all the organization's network just by using an infected USB memory or connecting an unsecure wireless network to the organization's network or visiting unsecure websites and downloading malicious contents to the network. By network quantification, the current status of the network security would be obtained much more precisely in a way which by it can be compared to different networks' security and the security of the network itself on a timeline base. Prioritization of network attributes based on the numeric effectiveness of each attribute on the network security score causes the efforts to a higher security to be more purposive and less error-prone. Economizing time and other resources on the security is the other important role of security quantification. The quantification of network security in this paper is done by using security metrics.

According to the national institute of standards and technology metrics are tools designed to improve determination, decision and responsibility by gathering, analyzing and reporting the related functions. In other word metrics are standard of measurement which can be used to measure the security level of an organization. Security

metrics are chosen according to the organization needs and security rules. A good security metric [8] should be specific, measurable, attainable, repeatable and time-dependent. There are some different categories of security metrics which can be considered in network security quantification such as:

- Software-based;
- Network-based;
- User-based;
- Policy-based.

One of the most important problems to increase network security is the absence of solutions to measure the relative effectiveness of different security attributes and metrics on the security level of a typical computer network because what is not measured is not controllable [2, 15]. In such a situation, a network security metric is useful because it would provide quantification and measurement supplied by different network attributes. By employing security metrics in a computer network, the administrators can find out which attributes should be more concentrated to increase security while resources consumption is decreasing. A computer network has numerous attributes and metrics which many of them are less important and time consuming to be analyzed.

Then by considering security metrics relationships, the less important ones can be omitted and those which are correlated more to the network security level are kept strongly. Every vendor and company which provides security solutions such as firewall, IDS/IPS, antivirus, UTM for other companies claims that its approach to security is the best, but unfortunately there is no quantitative way to assess their products. The approach presented in this paper tries to show the effect of security metrics individually on the security of the whole computer network by evaluating the relationships between security metrics.

The remainder of this paper is organized as follows. The related works are reviewed in Section 2 and the proposed approach to security quantification is described in Section 3. An experiment of the network security quantification's solution is carried out in Section 4, and finally the paper is concluded in Section 5.

2 Related Works

Most of the works in the network security quantification is about identifying an appropriate set of security metric. Ahmed et al. [1] gathered a set of metrics based on vulnerable networks previously found. The authors quantified the present vulnerabilities and their characteristics and estimated the future vulnerabilities in the networks and its services. In another study [18] all misconfigurations and weaknesses which causes the network to be vulnerable to the attacks are studied. By introducing a new metric called VEA-bility security metric, as a comparison tool for different network configurations in order to select the best adjust in the security of the network. A network administrator tries to have the less vulnerable network configuration and of course the more secure one, therefore the writers try to deliver different network configurative comparisons to help the users to choose the best ones. Attack graph-based security metrics are used in [5] to measure the probability of network exploitation according to number of successful attacks done. The network resistance which an attacker is faced with is one of the metrics the authors used. In all the works done in security quantification, the effect of attributes in the network that an administrator is daily faced with are not taken into consideration for determining the level of security.

One important reality should not be forgotten that things which are not measurable are not controllable. In another research [19] researchers by means of attack graph and defining two security metric called probabilistic security metric and attack resistant metric to evaluate the security level of the network. Common vulnerability scoring system has an important role in risk evaluation of the network. This system as described in [13] and [12], is an important step toward network security quantification. The standardized vulnerability scores, open and clear structure for security vulnerability scoring and prioritization of risk identification are the most important features of this system. In [10] the author is describing a way to rank the security metrics based on decision theory and probability distribution. A self-assessment architecture that prepare a solution for the users to determine security metrics that are specially feasible for the user's ISMS is presented in [7]. Then a metric catalogue involving 95 metrics from different sources is provided. In [14] the basic aspects of security metrics are covered. Matters such as definition of security metrics, their value, and difficulties in generating them and a methodology for building a security metric program are expressed briefly. The author in [16] describes some important features and goals of security metrics. Attack graph-based security metrics are used in [6] to measure the probability of network exploitation according to number of successful attacks done. The network resistance which an attacker is faced with is one of the metrics the authors have used in this work.

In [17] taxonomy of Intrusion Response Systems (IRS) and Intrusion Risk Assessment (IRA), two important components of an intrusion detection solution are represented. A self-assessment framework that permits a user to determine the security metrics that are feasible for the user's ISMS is discussed in [7]. In [11] a method to improve the network security, which consists of the network management, the vulnerability scan, the risk assessment, the access control, and the incident notification is introduced. In [3] a risk estimation model based on publicly available data, the Common Vulnerability Scoring Systems (CVSS) is proposed. In [4] a big amount of information is gathered by focusing on three European countries for more than a year and a half through 5 vantage points with different access technologies to make a quantitative

Table 1: The variables that are used in correlation

r	Correlation
M_1, M_2	The value of security metrics

measurement on the behavior of users with the Internet to gain important metrics.

3 The Proposed Approach for Security Quantification

To quantify the network security, a mathematical-based approach using regression and correlation proposed in this paper. Regression and correlation makes possible analyzing the relationships between security metrics in the model of network security quantification. In Figure 1 the structure of this approach is shown.

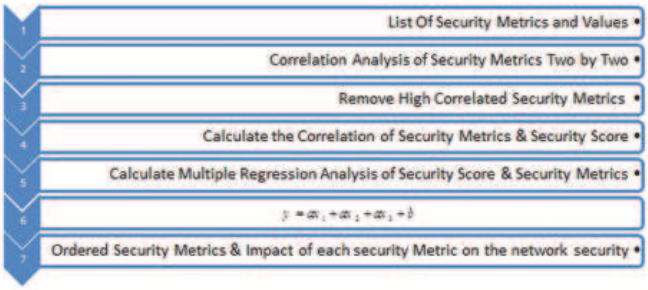


Figure 1: The structure of proposed Network Security Quantification Model

In this structure, the security of the network obtained through three important phases. First, the correlation between security metrics two by two is calculated. In this phase to prevent from multicollinearity[1], one of those metrics which are more correlated to the other one is omitted. By this strategy, just those metrics which are more important and effective to the security remains in the model of quantification. The lesser correlation value between security metrics, the more accurate would be the results of regression model of network security quantification. Figure 2 shows all the correlations between metrics that considered.

Equation (1) calculates the correlations of security metrics two by two.

$$r = \frac{\sum M_1 M_2}{\sqrt{\sum M_1^2 \sum M_2^2}} \tag{1}$$

In this equation the variables are assumed according to Table 1.

At the next phase, the correlations of security score and security metrics two by two are calculated. In this step, the more the value of correlations with the security score the more suitable would be the quantification model of network security. Since correlation is not a cause and effect concept then it just imply the presence of relationship between two security metrics therefore the effect of one specific security metric to the security score of the machine will not be concluded. Then in the next phase by means of regression, the effect of security metrics on the security score is calculated. By implementation of regression, the model of network security would be like Equation (2):

$$SecurityScore = b_0 + b_1 M_1 + b_2 M_2 + ... \tag{2}$$

In which M_i is the value of different security metrics and b_i is the coefficients that expresses the impact of security metrics on the security of the network. In the statistical models, regression evaluation used to study the relationship of variables in a cause and effect method. In the model of network security quantification the security score is the dependent variable and the security metrics are the independent variables. In a regression model, the effect of each independent variable on the dependent variable analyzed. By use of Equations (6) and (7) the coefficients in Equation (4) would be calculated.

$$b_i = \frac{\sum[(M_i - \overline{M})(SS_i - \overline{SS})]}{\sum[(M_i - \overline{M})^2]} \tag{3}$$

$$b_0 = \overline{ss} - b_1 \overline{M}. \tag{4}$$

In these equations the variables are according to Table 2.

Figure 2: The correlation of Metrics

4 Experiment

In this study, an organization's network involved about 100 machines, monitored in about two weeks to identify suitable network's security metrics. Fortunately, whole

[1] In statistics, multi-collinearity is a phenomenon in which two or more predictor variables in a multiple regression model are highly correlated.

Table 2: The list of variables used in coefficient of regression

b_i	The regression coefficients
M_i	The observation i of security metrics
SS_i	The observation i of security score
$\overline{ss}$	The average of security scores
$\overline{M}$	The average of security metrics

the network configurations and also its machines were accessible with appropriate privilege to be investigated in order to extract the security metrics else all the traffic in this period should had been saved to be interpreted later with an offline method. The point of security metrics is not to collect huge amount of data. A small set of data, understood well and usable, would be much more valuable than a pile of data left untouched on shelves or hard drives gathering dust. In this paper, the GQM method employed to develop security metrics. GQM is a simple and three-step process to gain appropriate security metrics for the network. The first step in the process involves defining specific goals that the organization hopes to achieve. These goals are those the organization by quantification is going to reach. Finally, these questions answered by identifying and developing appropriate metrics. This method guarantees that all the metrics identified are according to the goals of the organization. After an act of investigation of the network and by doing some interviews by the network's administrators and users according to the GQM method some more important security metrics chose. In [18], the writer expresses the characteristics of a good security metric such as: consistently measured, cheap to gather, expressed as a cardinal number or percentage and specific.

4.1 Web Browser Version (Browser)

Since most of the attacks that a machine is faced with is from the internet and web browsers are the first applications are to the target of attackers, then the web browser version is taken into consideration as a security metric. For each machine on the network the value for this metric is calculated by:

$$WebBrowser = LastVersion - UserVersion. \quad (5)$$

The difference between last version of a specific version and user's browser version is the value allocated to this metric for each machine on the network. In cases more than one browser is used the average of the values is allocated for this metric. Of course in the experimented network the browser which was used according to the security policy of the organization was IE and Firefox. Figure 3 shows the statistics of common browsers used in the organization's network.

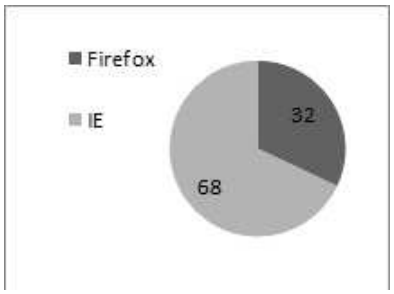


Figure 3: The web browsers statistics

Table 3: The assumed values for various OS

XP-SP1	XP-SP2	XP-SP3	7	8	8.1
5	4	3	2	1	0

4.2 Operating System Version (OS)

Since operating system is the infrastructure software for other applications and services to be executed properly, then keeping machine's OS updated is one of critical metrics, which should be considered to have a secure network. The value used for this metric in evaluations presented in Table 3.

All of the operating systems used in the experimented network were different versions of Microsoft windows and they were evaluated according to Table 3. For example, if on a machine win 8.1 is installed the value considered for it is 0 and if win XP-SP1 is installed the value would be 5. It means the higher version of OS the lower numeric specified for that version.

4.3 Vulnerabilities (VUL)

By running Nessus vulnerability scanner on the machines, the number of vulnerability on each machine is going to be in consideration as a security metric. Nessus is the world's most popular vulnerability scanner [5] and in 2005 used in 75000 organizations.

4.4 Malwares (malware)

By using licensed antivirus's reports, the number of malwares on the machine was obtained. On the experimented network an updated licensed NOD32 antivirus is installed. The server's side of this antivirus has several features, which can report the malwares penetrated into the machine. The value of this metric is the total number of observed malwares on the clients.

4.5 Defense Update (Def-Update)

Since the number of new threats continues to grow steadily, then antivirus' being-updated is so important to have a healthy network. The value allocated to this metric is the total number of days past from the last up-

Table 4: The list of variables used in the security score calculation

n	The number of machines
severity (v_i)	The severity for the vulnerability i
$Securityscore_i$	The security score for machine i

Table 5: The correlation of Security Metrics

	Browser	OS	Malware	Def-Update	Last-Scan	Update	VUL
VUL	0.09	-0.09	0.08	0.03	0.21	-0.35	-
Update	-0.30	-0.42	-0.014	-0.03	0.00	-	-0.35
Last-Scan	0.05	-0.13	-0.03	0.24	-	0.00	0.21
Def-Update	-0.11	-0.03	-0.11	-	0.24	0.03	0.03
Malware	0.11	0.19	-	-0.11	-0.03	-0.14	0.08
OS	0.43	-	0.19	-0.03	-0.13	-0.42	-0.09
Browser	-	0.43	0.11	-0.11	0.05	-0.30	0.09

date of client's side of antivirus. This value obtained by checking the update part of each antivirus.

4.6 Last on-demand scan (Last-scan)

Periodically scanning of the machines in a network is one of the main issues, which can help the network cleanness of malwares and vulnerabilities. Therefore, the total number of days past from the last scan of the network by the antivirus has been taken into account as an important security metric.

4.7 Software Updates (Software)

This metric value is the total number of OS and frequently used applications updates according to the security policies in the network. The updates can be managed and obtained by soft wares such as WSUS.

4.8 Security Score (security-score)

In order to implement the level of security in the security quantification model, a security score is going to be calculated. The more security score for each machine, the higher the level of security in the network. To calculate security score for each machine, all the vulnerabilities in each machine extracted by using Nessus vulnerability scanner.

Then to obtain the severity of each of the vulnerabilities, they mapped to the NVD[2] one by one. In the NVD, all the vulnerabilities are stored with a CVSS[3]-based severity, which is a number between 0 and 10. According to the Equation (6) the security score for each machine is calculated.

$$Securityscore_i = \sum_{i=1}^{n_k} (10 - severity(v_i)). \quad (6)$$

The variables of Equation (6) is explained in Table 4.

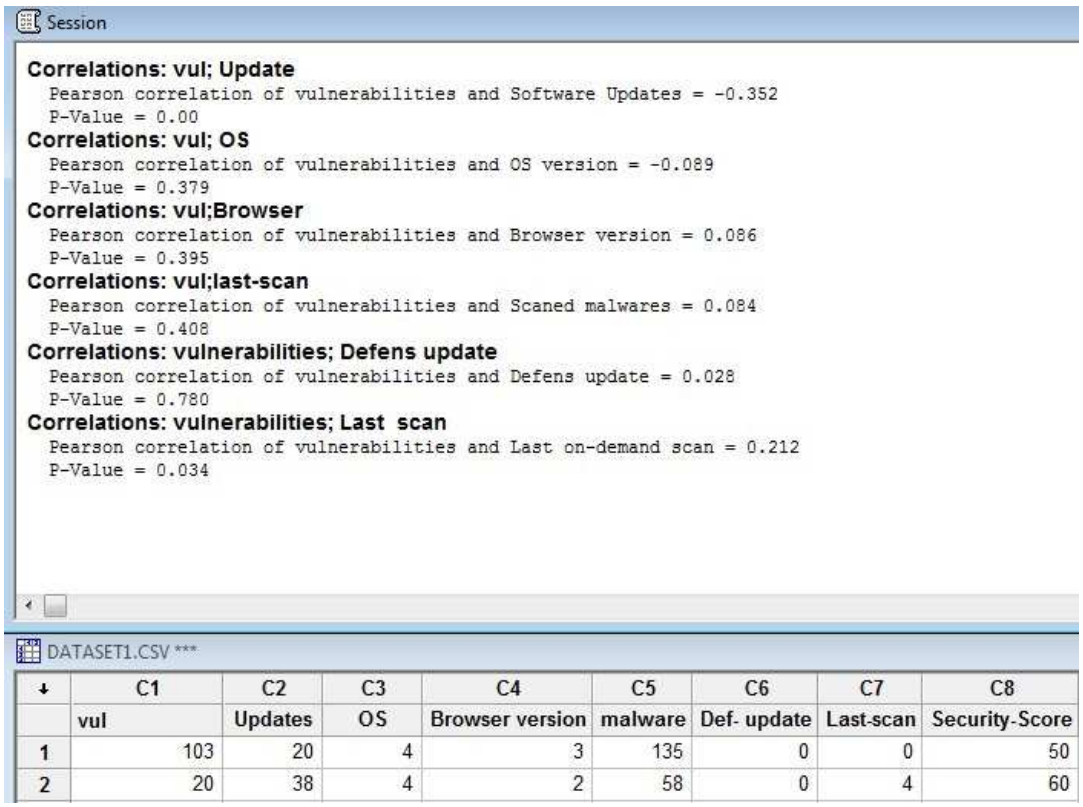
With the help of the theoretical development done in Section 4, now the numeric effect of security metrics on the network security level is going to be calculated.

In this research, the Minitab software version 16 used to interpret the relationship of security metrics. As clarified before the correlation of the security metrics is calculated to avoid multi-collinearity and the results are shown in Figure 4 and Table 5. Correlation of all the security metrics showed in Table 5 to show the non- cause and effect manner of this evaluation. As it is obvious, in Table 5, the correlation of to security metrics Update and VUL is a negative number and it shows that they are correlated in a reversed manner or better to say the more updates taken by the machines, the lesser vulnerabilities found on, or the correlation of two other security metrics VUL and Last-scan is a positive number meaning, the longer time lapsed the last scan, the more vulnerabilities found on the clients. A quick consideration of data in the last table reveals that the results exactly coincides the expectations in the real world.

After evaluating of the correlation between security metrics two by two, the correlation of security score and security metrics should be analyzed. Table 6 contains the result of correlation of security metrics and security score. In this table P-value is also considered by which the results can be better proved. The P-value would reject the null hypothesis[4], if its value was less than the alpha level which is important in the null hypothesis theory. In the other word for the P-Values less than alpha level the null hypothesis is rejected and it shows there is a meaningful relationship between two variables. Figure 5. shows the correlation of security metrics and security scores taken from Minitab. Scatter plots also can be implemented to visualize the correlation of security metrics and security scores. In Figure 6 the correlations of security metrics

[2]National Vulnerability Database:www.NVD.com
[3]Common Vulnerability Scoring System

[4]The term "null hypothesis" usually refers to a general statement or default position that there is no relationship between two measured phenomena, or no difference among groups and variables



↓	C1	C2	C3	C4	C5	C6	C7	C8
	vul	Updates	OS	Browser version	malware	Def- update	Last-scan	Security-Score
1	103	20	4	3	135	0	0	50
2	20	38	4	2	58	0	4	60

Figure 4: The correlation of security metrics generated by Minitab

Table 6: The correlation of Security Metrics

	Browser	OS	Malware	Def-Update	Last-Scan	Update	VUL
Security score	-0.26	-0.21	-0.37	0.1	-0.26	0.38	-0.23
P-value	0.01	0.03	0.00	0.3	0.3	0.00	0.02

and security scores is illustrated by scatter plot.

As it is evident in the Figure 6 according to the trend line of scatter plot the correlation of security metrics and security score is obvious. For example the positive slope of trend line in scatter plot of Def-update and security scores illustrates as Def-updates increases the security score is also increased and the negative slope of trend line for Malwares and security scores means as number of malwares increases the security score decreases.

4.9 Numeric Effect of Security Metrics on the Security Experimented Network

Since correlation is not a cause and effect evaluation, then by means of regression evaluation the exact numeric effect of each security metric on the security score is calculated. In the quantification model of network security Security-Score is dependent variable and the security metrics are the independent variables of the model in which we are going to calculate the effect of them on the security score. Finally, the multiple regression equation expresses the numeric effect of security metrics on the security level of the network. Equation (7) is the regression equation of the quantification model of the network security.

$$\begin{aligned} Security - Score = 56.4872 - (0.211009)Vulnerability \\ +(0.525058)Update + (0.343473)OS \\ -(0.345093)Browser - (0.0300952)Malwares \\ +(0.0511584)Def - update - (0.0575463)Last - scan. \end{aligned} \quad (7)$$

The Equation (5) is the quantification equation of the network security in which the coefficients are the numeric effect of each security metric on the security when the other metrics are assumed as 1. The Figure 7 illustrates this model calculated in Minitab software.

According to Equation (5) the order of security metrics according to their effect on network security is gathered in Table 7. The security metric Updates has the most effect on the security score with coefficient 0.52 and the Web browser, OS and Vulnerabilities are the next more important security metrics orderly.

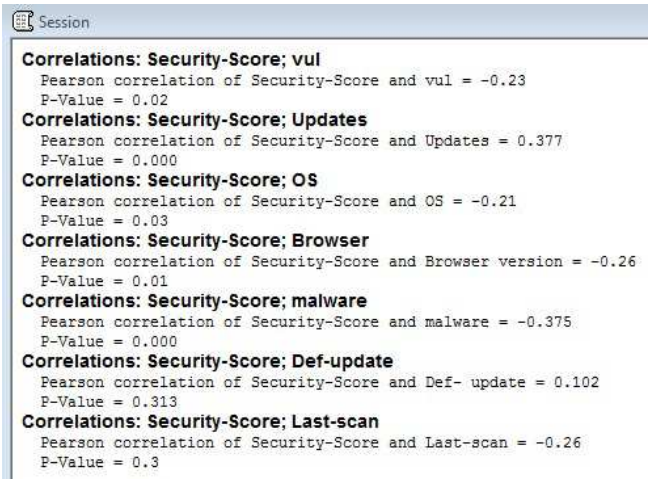


Figure 5: Correlation Of security metrics and security score

Table 7: The correlation of Security Metrics

Rank	Security Metrics	The Importance Value
1	Updates	0.52
2	Browser	0.3451
3	OS	0.3435
4	Vulnerabilities	0.21

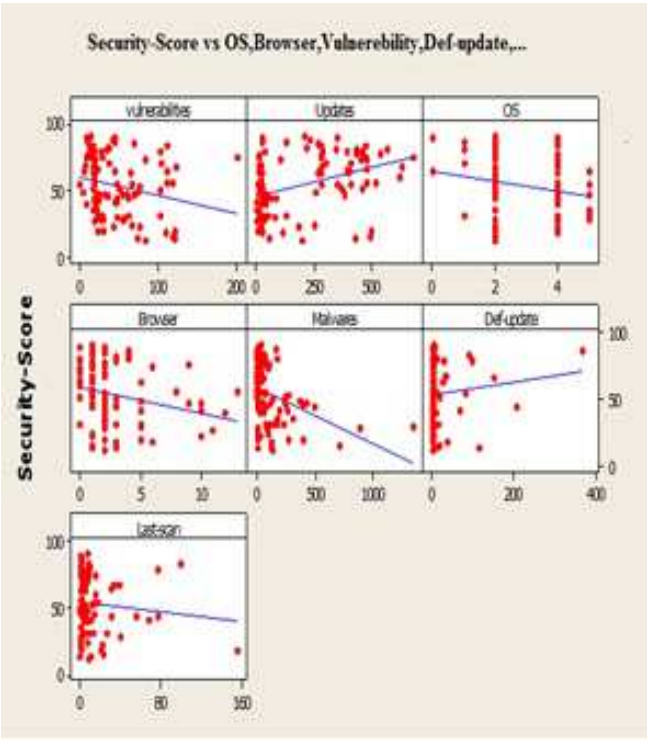


Figure 6: The scatter-plot of security score and security metrics

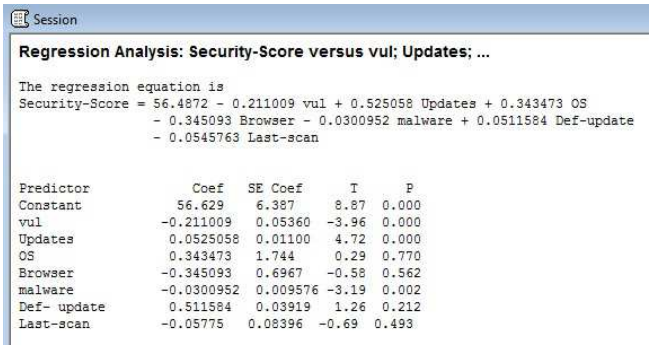


Figure 7: The regression model of network quantification

5 Conclusion

This paper is going to provide a structure for quantification of network security and prioritization of significant security metrics. A mathematical approach is developed that can help to quantify the network security and order the security metrics. By implementing regression and correlation to the network security era and security metrics the quantification of network security will be possible as shown in this paper. Once the security quantification is done, administrative efforts can be concentrated to increase security more precisely and efficiently. As it is shown in this paper there are some relationships between network attributes or security metrics which by evaluation of them the network administrators can manage the network more efficiently.

References

[1] M. S. Ahmed, E. Al-Shaer, and L. Khan, "A novel quantitative approach for measuring network security," in *The 27th IEEE Conference on Computer Communications (INFOCOM'08)*, pp. 1957–1965, 2008.

[2] A. Anurag, "Network neutrality: Developing business model and evidence based net neutrality regulation," *International Journal of Electronics and Information Engineering*, vol. 3, no. 1, pp. 1–9, 2015.

[3] G. A. Franca III, "Baseline operational security matrics for industrial control system," in *Proceedings of the International Conference on Security and Management (SAM'16)*, p. 8, The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing, 2016.

[4] J. L. García-Dorado, A. Finamore, M. Mellia, M. Meo, and M. Munafo, "Characterization of ISP traffic: Trends, user habits, and access technology impact," *IEEE Transactions on Network and Service Management*, vol. 9, no. 2, pp. 142–155, 2012.

[5] H. Ge, L. Gu, Y. Yang, and K. Liu, "An attack graph based network security evaluation model for hierarchical network," in *IEEE International Conference on Information Theory and Information Security (ICITIS'10)*, pp. 208–211, 2010.

[6] N. Ghosh and S. K. Ghosh, "An approach for security assessment of network configurations using at-

tack graph," in *IEEE First International Conference on Networks and Communications (NETCOM'09)*, pp. 283–288, 2009.

[7] B. Heinzle and S. Furnell, "Assessing the feasibility of security metrics," in *International Conference on Trust, Privacy and Security in Digital Business*, pp. 149–160, Springer, 2013.

[8] A. Jaquith, *Security Metrics: Replacing Fear, Uncertainty, and Doubt*, Addison Wesley, 2007.

[9] H. Kahtan, N. A. Bakar, and R. Nordin, "Dependability attributes for increased security in component-based software development," *Journal of Computer Science*, vol. 10, no. 8, pp. 1298–1306, 2014.

[10] M. Khan, M. Omer, and J. Copeland, "Decision centric identification and rank ordering of security metrics," in *IEEE 37th Conference on Local Computer Networks (LCN'12)*, pp. 208–211, 2012.

[11] Y. P. Lai and P. L. Hsia, "Using the vulnerability information of computer systems to improve the network security," *Computer Communications*, vol. 30, no. 9, pp. 2032–2047, 2007.

[12] P. Mell and K. Scarfone, "Improving the common vulnerability scoring system," *IET Information Security*, vol. 1, no. 3, p. 119, 2007.

[13] P. Mell, K. Scarfone, and S. Romanosky, "Common vulnerability scoring system," *IEEE Security & Privacy*, vol. 4, no. 6, pp. 85–89, 2006.

[14] F. Nielsen, "Approaches to security metrics," in *CSSPAB Workshop on" Approaches to Measuring Security*, 2000.

[15] E. U. Opara, O. A. Soluade, "Straddling the next cyber frontier: The empirical analysis on network security, exploits, and vulnerabilities," *International Journal of Electronics and Information Engineering*, vol. 3, no. 1, pp. 10–18, 2015.

[16] S. C. Payne, "A guide to security metrics," *SANS Institute Information Security Reading Room*, 2006.

[17] A. Shameli-Sendi, M. Cheriet, and A. Hamou-Lhadj, "Taxonomy of intrusion risk assessment and response system," *Computers & Security*, vol. 45, pp. 1–16, 2014.

[18] M. Tupper and A. N. Zincir-Heywood, "Vea-bility security metric: A network security analysis tool," in *Third International Conference on Availability, Reliability and Security (ARES'08)*, pp. 950–957, IEEE, 2008.

[19] L. Wang, T. Islam, T. Long, A. Singhal, and S. Jajodia, "An attack graph-based probabilistic security metric," in *IFIP Annual Conference on Data and Applications Security and Privacy*, pp. 283–296, Springer, 2008.

Biography

Mostafa Behi joined central office of Communication and Information Technology (ICT) of South Khorasan province as an IT Expert on summer 2014. He received his MS degree in computer's software engineering from Azad university of Birjand and his B.S degree from university of Birjand. He mostly researches on network security, cloud computing and data mining.

Mohammad GhasemiGol will join the Department of Computer Engineering at the University of Birjand on fall 2016. He received the B.S. degree in Computer Engineering from Payame Noor University (PNU), Birjand, Iran, in 2006. He also received the MS and PhD degree in Computer Engineering at FUM, Iran, in 2009 and 2016 respectively. November 2014 to July 2015, he was with the Department of Computer Science and Engineering, University of North Texas, Denton, TX, USA as a visiting research scholar. His research interests include network security, intrusion detection and response systems, alert management, data mining, and optimization problems.

Hamed Vahdat-Nejad is currently an assistant professor at the computer engineering department of the University of Birjand. He received his PhD from computer engineering department of University of Isfahan in 2012, his master degree from Ferdowsi University of Mashhad in 2007, and his bachelor's degree from Sharif University of Technology in 2004. He was a research scholar at the Middleware laboratory of Sapienza University of Rome in 2011. Currently, his research is focused on cloud computing, pervasive computing and security. He has (co)published about 30 papers in conferences and journals, and leads the Pervasive and Cloud computing Lab at the University of Birjand. He has served as the chairman of the 1st and 2nd International Workshop on Context-aware Middleware for Ubiquitous Computing Environments, as well as the 3rd and 4th International workshop on Pervasive and Context-aware middleware. He has served as TPC member for ICCKE, IWCMC, ISIEA, ICCIT-WCS, PerCAM, ChinaCOM, MELECON2014, COGNITIVE-2014, IBMSGS2015, EMERGING 2015, ICACCI, ADMMET'2015 ICCME-2015, CoCoNet'15, AR4MET'2016, REEGETECH'2016, ISTA16, etc. Currently, he serves as associate editor for Elsevier Computers and electrical engineering journal.

An Improved Ownership Transfer for RFID Protocol

Rui Xie[1], Bi-yuan Jian[2], and Dao-wei Liu[2]
(Corresponding author: Bi-yuan Jian)
School of Automation, Guangdong University of Technology[1]
Guangzhou 510006, China
School of Electronic and Computer Engineering, Guangzhou Vocational College of Science and Technology[2]
No. 1038, Guangcongroad Zhong Latan Baiyun District Guangzhou, Guangzhou 510006, China
(Email: jianbiyuan1983@126.com)
(Received Mar. 6, 2017; Revised and Accepted May 16 & Jun. 21, 2017)

Abstract

The ownership transfer problems occur during using the RFID tag. In view of the problems of the RFID tag ownership transfer protocol, such as security defects and high computational cost, an improved lightweight RFID tag ownership transfer protocol is proposed in this paper. The improved protocol does not depend on the trusted third party, so that the improved protocol has a wider application space. Using the challenge-response mechanism, the new owner of the tag introduces the counter count and performs the corresponding operation according to the value of count to solve the desynchronization attack problems. The analysis results show that the improved protocol not only satisfies the security requirements of the tag transfer, but also overcomes the security defects of desynchronization attack. Compared with the existing RFID tag ownership transfer protocols, this improved protocol has larger promotion in the aspect of security and efficiency.

Keywords: Internet of Things; Ownership Transfer Protocol; Rabin Algorithm; RFID, The Synchronization Attack

1 Introduction

Radio frequency identification (RFID) is a kind of non-contact information transmission with a use of radio frequency signals, in order to achieve the purpose of identification by the transmitting information. RFID technology is widely used in production, logistics, national defense, transportation and other fields owning to the advantages of small size, easy portability, low cost, long life and so on [2, 7, 9, 16]. Because the RFID tag resources are limited and are running in the open wireless environment, the RFID system communication is vulnerable to various security threats such as eavesdropping attack and replay attack. So it is essential to ensure the security of the running protocols [3, 8, 12, 15, 17].

The ownership of the entity often changes during the practical application process. For example, after the producer sells the commodities to the wholesaler, the wholesaler has the ownership of the commodities physically, but it doesn't mean that the wholesaler completely controls the ownership of the commodities [4]. If there is no change in the ownership of the tag, the producers can still scan and obtain the tag's information; thereby the wholesaler's privacy may be exposed. When the wholesaler retails the commodities to the retailer, there will be a problem whether the ownership transfers completely or not, and likewise there will be a hidden danger of exposure of the tag's private information [1, 10, 18].

Reference [11] firstly proposes the RFID tag ownership transfer protocol, which uses trusted centers that are co-trusted by the old and new owners of tags to control all of the tags' information, but it limits the range of use of tags. In [14], the security and privacy requirements of the RFID tag ownership transfer protocol are defined, and three sub-protocols are proposed to achieve the transfer of RFID tag ownership with no trusted centers. However, several scholars have pointed out that the protocol has a lot of security problem. Reference [15] gives an improvement to the protocol in [14], and it proposes an extensible RFID authentication protocol that supports the tag ownership transfer, but the improved protocol still cannot protect the backward privacy and is vulnerable to de-synchronization attacks. In [5], two transfer schemes are proposed based on the quadratic residue, but both schemes require the exchange of information between the tags and the old and new owners over and over again, which results in the low computational efficiency of the tags. The ownership transfer protocol proposed in [6] cannot resist counterfeiting attacks, and the tags are easy to be tracked because the private keys Kp and Ku used in the tags are not updated every time. The ownership transfer

protocol proposed in [19] cannot resist DoS attacks, and replaying messages will make the tag updated repeatedly so that the protocol cannot resist replay attacks of the tags as well.

The ownership transfer protocol proposed in [13] can not resist the de-synchronization attacks. The attacker can obtain the message Q by listening to a complete communication process; then, by replaying the message Q during the process of blocking the associated communication, the tag's original owner continually updates the shared key so that the shared key between the tag's owner and the tag is different, which ultimately makes the both shared information out of sync. Aiming at the security flaws in the protocol of [13], an improved RFID tag ownership transfer protocol is proposed. In this protocol, the counter count is introduced on the tag's original owner, and the value of the counter is used to solve the problem of de-synchronization attack defects in the original protocol.

The remainder of this paper is organized as follows. The second part of the paper is to conduct a security analysis of the protocol in [13]. The third part is to put forward my own ownership transfer protocol. The fourth part is to carry out security analysis of the proposed protocol. The fifth part is using the BAN logic to formally verify the proposed protocol. The sixth part is the performance comparison between the proposed protocol and other protocols. The seventh part is the summary and concluding remarks.

2 Jin-wei Shen et al.'s Protocol and Its Drawbacks

The protocol can not resist de-synchronization attacks. In [13], an improved ultra-lightweight RFID ownership transfer protocol is proposed, which claims to be resistant to de-synchronization attacks. However, the study in this paper found that the ownership transfer protocol in [13] can not resist the de-synchronization attacks. The specific attack process is as follows.

The attacker can obtain all the information? such as IDS, M, N, P, Q, X, Y –in the complete communication process of [13] by using some monitoring methods. After obtaining the above information, the attacker can immediately block the communication process of the previous five steps, so that the shared key between D_j and T can be out of sync by continually replaying the message Q.

The first replay is as follows. The attacker disguises as D_i to send the intercepted Q message to D_j. Since the authentication of Q is passed before, the replay information Q can also be authenticated. Before the message is replayed, the information stored in D_i is s_i, t_i, X, Y, $IDSold = IDS$, and $IDSnew = IDS \oplus NT \oplus NR$. After the message is replayed, D_i generates the random number S_{i+1}, calculates t_{i+1}, $X1$, $Y1$, and updates the data $IDSold = IDS$, $IDSnew = IDS \oplus NT \oplus NR$. Let $IDSnew = IDS1$, $u_i = s_i$, $v_i = t_i$, $s_i = s_{i+1}$, $t_i = t_{i+1}$. After D_i is updated, $X1$ and $Y1$ will be sent to the tag, and the attacker will block the information transmission between the both.

The second replay is as follows. After the first replay, the attacker intercepts the $X1$, $Y1$ that are transmitted to the tag by D_i. At this time the attacker prevents the information from being transmitted to the tag, and meanwhile replays the message Q again. Because D_i stores the shared keys of this and the last authentication, so Q can still be authenticated. After Q is replayed again, D_i will be set as follows.

D_i generates random number S_{i+2} and calculates t_{i+2}, $X2$, $Y2$; then updates data $IDSold = IDSnew = IDS1$, $IDSnew = IDS1 \oplus NT \oplus NR$. Let $IDSnew = IDS2$, $u_i = s_{i+1}$, $v_i = t_{i+1}$, $s_i = s_{i+2}$, $t_i = t_{i+2}$. After D_i is updated, $X2$ and $Y2$ will be sent to the tag, and the attacker will block the information transmission between the both.

The third replay is as follows. After the second replay, the attacker intercepts the $X2$, $Y2$ that are transmitted to the tag by D_i. At this time the attacker prevents the information from being transmitted to the tag, and meanwhile replays the message Q again. Because D_i stores the shared keys of this and the last authentication, so Q can still be authenticated. After Q is replayed again, D_i will be set as follows.

D_i generates random number S_{i+3} and calculates t_{i+3}, $X3$, $Y3$; then updates data $IDSold = IDSnew = IDS2$, $IDSnew = IDS2 \oplus NT \oplus NR$. Let $IDSnew = IDS3$, $u_i = s_{i+2}$, $v_i = t_{i+2}$, $s_i = s_{i+3}$, $t_i = t_{i+3}$. After D_i is updated, $X3$ and $Y3$ will be sent to the tag, and the attacker will block the information transmission between the both.

After the above three replay attacks are completed, the attacker will transmit the original intercepted message X, Y to the tag. Because the tag has never updated the shared key during the previous three replay attacks, X and Y certainly can be authenticated. After the authentication, the tag updates the shared key, $IDS = IDS \oplus NT \oplus NR$, i.e. $IDS = IDS1$; the shared key is t_{i+1}.

When analyzing the tag and the shared key ultimately stored in D_i, we can find that there is no synchronization between them. The information stored in the tag is $IDS1$, t_{i+1}, but the information stored in D_i is $IDS3$, t_{i+3}. At this time, the attacker successfully makes the shared key between D_i and the tag no longer the same by the replay attacks, so that the subsequent authentication fails. We can draw a conclusion that the original protocol can not resist the de-synchronization attacks.

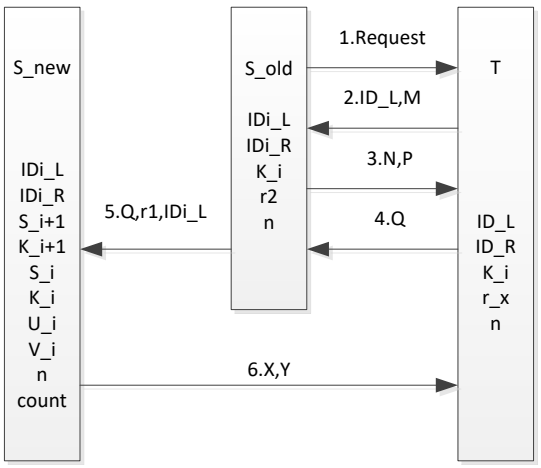


Figure 1: The improved protocol

3 Improved RFID Tag Ownership Transfer Protocol

The ownership transfer protocol proposed in [13] can not resist the replay attacks and de-synchronization attacks, so that this paper propose an improved RFID tag ownership transfer protocol. In this improved protocol, the message counter count is introduced to the tag's original owner S_old, and the replay attacks are resisted by the value of the counter count. The counter count is to record how many times the message Q is replayed. The value of count does not exist or is 0, indicating that the message Q is transmitted for the first time. If the value of count is not 0, the message Q may be a replayed message. Since the above two situations are different, the operations of the tag's original owner S_old are different as well, which not only is able to resist the replay attacks, but also avoids the asynchronous problems between the tag's original owner S_old and the tag.

As is the same to other authentication protocols, we assume that the transmission channels between the tag's original owner S_old and the tag's new owner S_new are secure. We also suppose the transmission channels between the tag's original owner S_old and the tag are insecure, and that the transmission channels between the tag's new owner S_new and the tag are insecure as well.

3.1 Symbol Description

Firstly, the meaning of each symbol in this protocol is given in Table 1.

3.2 Protocol Description

The process of the improved RFID tag ownership transfer protocol is presented in Figure 1.

The descriptions of the symbols M, N, P, Q, X, Y are

Table 1: The symbols used in the paper

Symbols	Meaning
S_{old}	The tag's original owner.
S_{new}	The tag's new owner.
T	A tag.
T_i	The i-th tag.
ID_i_L	The left half of the i-th tag identifier ID. (its length is L bits)
ID_i_R	The right half of the i-th tag identifier ID. (its length is L bits)
ID_L	The left half of the tag identifier ID. (its length is L bits)
ID_R	The right half of the tag identifier ID. (its length is L bits)
r_x	The tag data saved at the beginning. (its length is L bits)
$r1$	The random number generated by the tag. (its length is L bits)
$r2$	The random number generated by the tag's original owner. (its length is L bits)
n	Mersenne number, where the value is n=2^L-1.
L	The length of the key.
S_{-i}	The private key of the tag T_i. (its length is L bits)
K_{-i}	The public key of the tag T_i, where $K_{-i}=(S_{-i})^2 \bmod n$. (its length is L bits)
U_{-i}	The private key of the tag T_i on the last round. (its length is L bits)
V_i	The public key of the tag T_i on the last round, where $U_{-i}=(V_{-i})^2 \bmod n$. (its length is L bits)
S_{-i+1}	The random number generated by the tag's new owner, used as the private key of the authentication on the current round. (its length is L bits)
K_{-i+1}	The public key of the authentication on the current round, where $K_{-i+1} = (S_{-i})^2 \bmod n$. (its length is L bits)
$count$	The counter for the message Q of the tag's new owner.
M, N, P, Q, X, Y	The communication data in this protocol. (each length is L bits)
$MIXBITS(a, b)$	The new random number obtained by computing (a, b). (the output length is L bits)
$\oplus$	XOR operation.
&	AND operation.
$[X]_L$	Take the first L bits of the result of the operation [].

as follows.

$$\begin{aligned} M &= K\_i \oplus r1. \\ N &= r2 \oplus ID_i\_R. \\ P &= [(r1 \oplus r2 \oplus K\_i)^2 \bmod n]_L. \\ Q &= [(r1 \oplus r2 \oplus D\_R)^2 \bmod n]_L. \\ X &= S\_{i+1} \oplus r1 \oplus ID_i\_R. \\ Y &= K\_{i+1} \& r1 \& ID_i\_R. \end{aligned}$$

$[\cdot]_L$ means taking the first L bits of the result of the operation [].

The complete execution steps of this ownership transfer protocol are described below.

Step 1: The tag's original owner S_old sends a request command Request for transferring tag ownership to the tag T, and opens ownership transfer session.

Step 2: T receives the messages from S_old; then calculates r1=r_x, M=$K\_i \oplus$ r1, and sends the values of ID_L and M to S_old.

Step 3: *S_old* receives the messages from *T*; then searches the database for the result of whether ID_i_L is equal to *ID_L*. If the result does not exist, the tag is forged, and the protocol terminates immediately. If it exists, *S_old* generates an *L*-bit random number *r2*, and uses K$_i$ (corresponding to ID_i_L) to calculate $K\_i \oplus M$ and obtain a random number *r1*.Then it uses *r1*, *r2*,ID_i_R (corresponding to ID_i_L) and K$_i$ to calculate $N = r2 \oplus ID_i\_R$ and $P=[(r1 \oplus 2 \oplus K\_i)^2 \bmod n]_L$. Finally the values of *N* and *P* are sent to *T*.

Step 4: *T* receives messages *N* and *P* from *S_old*. The tag uses its own *ID_R* to calculate $N \oplus$ *ID_R* and obtain random number *r2*. Then the tag uses random number *r1* generated by itself, random number *r2* and its own public key K_i to verify the correctness of *P*, i.e.

$$P' = [(r1 \oplus (N \oplus ID\_R) \oplus K\_i)^2 \bmod n]_L.$$

If P' is unequal to *P*, then *S_old* is forged, and the protocol terminates immediately. If P' is equal to *P*, the tag correctly verifies *S_old*. Next the tag begins to update data *r_x=MIXBITS(r1, r2)*, and uses random number *r1* generated by itself, random number *r2* and its own *ID_R* to calculate the value of *Q*. Finally the value of *Q* is sent to *S_old*.

Step 5: *S_old* receives the message *Q* from the tag. *S_old* uses random number *r2* generated by itself, random number *r1* and its own ID_i_R to verify the correctness of *Q*, i.e.

$$Q = [((K\_i \oplus M) \oplus r2 \oplus ID_i\_R)2 \bmod n]_L.$$

If Q' is unequal to *Q*, then the tag is forged, and the protocol terminates immediately. If Q' is equal to *Q*, the tag correctly verifies *S_old*. Then *S_old* send all the values of *Q*, *r1*, *ID_L* to the new tag's owner *S_new* through secure channels.

Step 6: *S_new* receives the messages from *S_old*. Then *S_new* searches the database for the result of whether Q' is equal to *Q*. If the result exists and the value of the corresponding counter count is not 0, it indicates that the message *Q* has been transmitted. In order to resist replay attacks, *S_new* executes Step 7. If the result does not exist, *S_new* executes Step 8.

Step 7: *S_new* searches the database for the result of whether ID_i_L is equal to *ID_L*. If the result does not exist, the tag is forged, and the protocol terminates immediately. If the result exists, *S_new* does not make any updates, and the values of *X* and *Y* which is calculated during the last authentication are transmitted directly to the tag.

Step 8: *S_new* stores the value of *Q* into its own database, allocates a corresponding counter, and sets the counter count to 1. Then *S_new* searches the database for the result of whether ID_i_L is equal to *ID_L*. If the result does not exist, then the tag is forged, and the protocol terminates immediately. If the result exists, then *S_new* generates a *L*-bit random number S_{i+1}, uses it as the new private key in the current authentication, and calculates $K\_{i+1} = (S\_{i+1})^2 \bmod n$. After the calculation is finished, *S_new* begins to update data $U\_i = S\_i$, $V\_i = K\_i$, $S\_i = S\_{i+1}$, $K\_i = K\_{i+1}$, and uses random number S_{i+1} generated by itself, *r1* transmitted from *S_old*, K_{i+1} and ID_i_R (corresponding to ID_i_L) to calculate $X=S\_{i+1} \oplus r1 \oplus ID_i\_R$, $Y = K\_{i+1}$ & *r1* & ID_i_R. Finally the values of *X* and *Y* are sent to the tag.

Step 9: *T* receives messages *X* and *Y* from *S_new*. Next the tag uses the random number *r1* generated by itself, its own *ID_R* and *X* sent from *S_new* to calculate $X \oplus r1 \oplus ID\_R$ and obtain the private key S_{i+1}. Then it uses the private key S$_{i+1}$, random number *r1* generated by itself and its own *ID_R* to verify the correctness of *Y*, i.e.

$$Y' = ([(S\_{i+1})^2 \bmod n]_L) \& r1 \& ID\_R.$$

If Y' is unequal to *Y*, then *S_new* is forged, and the protocol terminates immediately. If Y' is equal to *Y*, the tag correctly verifies *S_new*, and then the tag begins to update data $K\_i = Y \oplus r2$. The tag ownership transfers successfully.

4 Security Analysis

4.1 Valid Target Transfer

Valid target (abbr. VT) transfer means it is the valid target that is transferred, instead of other tags in the

system. In the improved protocol, the tag's original owner *S_old* verifies the authenticity of the tag for the first time in Step 3. The tag firstly verifies the authenticity of the tag's original owner *S_old* in Step 4, and then *S_old* will verify the tag's authenticity again in Step 5. It makes the authentication security improve greatly between the tag's owner and the tag, and after mutual authentication it can make sure that the current authenticated tag is certainly the tag that belongs to *S_old*.

S_new verifies the tag's authenticity in Step 6. The tag verifies the authenticity of *S_new* in step 9. During the entire authentication process, the security of the protocol improves greatly because there is a mutual authenticity verification between *S_new* and the tag. Through the above process to achieve mutual authentication, it can ensure that the tag is certainly the target that will be transferred to *S_new*. The target tag has been authenticated several times to complete the ownership transfer from *S_old* to *S_new*. As a result, the improved protocol ensures that the transfer tag is certainly the target, not the other tags in the system.

4.2 Impersonation Attack

We assume that the attacker impersonates the tag's original owner *S_old*. Because the attacker does not know the shared key K_{-i} and ID_i_Rbetween *S_old* and the tag T_i, the attacker can not correctly calculate the values of N and P. In Step 4 the tag will promptly find that *S_old* is forged, and the protocol terminates immediately.

Then we assume that the attacker impersonates the tag's new owner *S_new*. Because the attacker does not know the shared key K_{-i} and ID_i_R between *S_new* and the tag K_{-i}, the attacker can not correctly calculate the values of X and Y. In Step 4 the tag will promptly find that *S_new* is forged, and the protocol terminates immediately.

Next we assume that the attacker impersonates the tag. The attacker knows neither the shared key K_{-i} and ID_i_R between *S_old* and the tag T_i, nor the shared key K_{-i} and ID_i_R between *S_new* and the tag T_i, so there is no way at all to correctly calculate the value of M and Q. Both *S_old* in Step 3 and *S_new* in Steps 6, 7, 8, will find that the tag is forged, and the protocol terminates immediately. Above all, the improved protocol can resist resist various impersonation attacks.

4.3 Brute Force Attack

By listening to a complete communication process, the attacker can obtain the values of M, N, P, Q, X and Y. In the improved protocol, the random numbers *r1*, *r2* are no longer transmitted in plain text, but simply encrypted with other information firstly. For instance, the attacker is unaware of the values of the shared private key K_{-i} and ID_i_R between *S_old* and the tag T_i. Moreover, the number *r2* is randomly generated by *S_old*, and the number *r1* is randomly generated by the tag as well. From the attacker's perspective, although the values of N and P are intercepted, it is impossible to make an exhaustion of any useful privacy information.Aiming at the formulas $P = [(r1 \oplus r2 \oplus K_{-i})^2 \bmod n]_L$ and $N = r2 \oplus ID_i\_R$, the attacker knows nothing about *r1*, *r2*, K_{-i}, ID_i_R. As a result, it is impossible for an attacker to analyze a specific K_{-i}. Simply knowing the values of N and P, there is no way to make an exhaustion of the specific values. For the same reason, the attacker is unable to make an exhaustion of any useful information by intercepting the values of M, N, P, Q, X, Y. Based on the above descriptions, the improved protocol can resist brute force attacks.

4.4 Replay Attack

During each execution of the improved protocol, *S_old* uses a random number *r2* generated by itself to keep the messages fresh, and similarly *S_new* uses a random number s_i generated by itself to keep the messages fresh. The tag uses a random number *r1* generated by function *MIXBIT(a, b)* to keep the messages fresh. Based on the above descriptions, the values of M, N, P, Q, X, Y are various in each step. Therefore, although the attacker replays the messages, any useful information won't be available. So the improved protocol can resist replay attacks.

4.5 De-synchronization Attack

In the improved protocol, *S_new* introduces the counter count. It's no use that the attacker replays message Q, because only when the value of count does not exist or is 0 will *S_new* generate a new random number s_i as the new shared private key, and then execute the subsequent update steps. If the value of count is not 0, it indicates that the message Q has existed before, and this Q is possibly the information which the attacker replays. In order to resist de-synchronization attack, *S_new* will not generate new random numbers, but directly use current private key to verify the correctness of Q. Then it uses current private key to update related data. During this process, according to the different values of count, the update operation also uses different mechanisms, so it avoids the differences of the shared private key between the tag and *S_new* caused by replaying the message Q. Based on the above descriptions, the improved protocol can resist de-synchronization attacks.

Table 2 shows the security comparison between this protocol and several other RFID tag ownership transfer protocols. $\surd$ indicates resistance, $\times$ indicates irresistance.

5 BAN Logic Formal Analysis

In this paper, BAN logic formalization method is used to prove the security of the improved protocol. BAN logic is proposed by Burrows et al. Using BAN logic, the proving process of the protocol is shown as follows. Because both

Table 2: Protocol security comparison

Attack Types	Ref.[11]	Ref.[13]	Ref.[14]	Ref.[15]	Our protocol
VT	√	√	√	√	√
Impersonate Attack	√	×	√	√	√
Brute Force Attack	√	√	√	√	√
Replay Attack	√	√	×	√	√
Replay Attack	×	√	√	×	√

S_new and *S_old* have part of readers, so we can see the both as a whole, as a large reader, represented with R.

5.1 Idealized Model of the Protocol

Message 1: $R \longrightarrow T$: *Q*uery;

Message 2: $T \longrightarrow R$: ID_L, M;

Message 3: $R \longrightarrow T$: N, $P5$;

Message 4: $T \longrightarrow R$: Q;

Message 5: $R \longrightarrow T$: X, Y.

5.2 Expected Target of the Protocol

The main proving target of the protocol's correctness is G1, G2, G3, G4, G5 and G6, that is, mutual authentication entity's belief in the freshness of interactive information.

G1: $R \mid\equiv Q$, R believes Q.

G2: $T \mid\equiv N$, T believes N.

G3: $T \mid\equiv P$, T believes P.

G4: $T \mid\equiv X$, T believes X.

G5: $T \mid\equiv Y$, T believes Y.

G6: $R \mid\equiv M$, R believes M.

5.3 Initial Assumptions of the Protocol

P1: $R \mid\equiv R \overset{K_{-i}}{\longleftrightarrow} T$, R believes R and T share the public key K_{-i}.

P2: $T \mid\equiv R \overset{K_{-i}}{\longleftrightarrow} T$, T believes R and T share the public key K_{-i}.

P3: $R \mid\equiv R \overset{n}{\Longleftrightarrow} T$, R believes R and T share the Mersenne number n.

P4: $T \mid\equiv R \overset{n}{\Longleftrightarrow} T$, T believes R and T share the Mersenne number n.

P5: $R \mid\equiv R \overset{ID\_R}{\Longleftrightarrow} T$, R believes R and T share the identifier *ID_R*.

P6: $T \mid\equiv R \overset{ID\_R}{\Longleftrightarrow} T$, T believes R and T share the identifier *ID_R*.

P7: $R \mid\equiv \sharp\,(r1)$, R believes the freshness of the random number *r1*.

P8: $T \mid\equiv \sharp\,(r1)$, T believes the freshness of the random number *r1*.

P9: $R \mid\equiv \sharp\,(r2)$, R believes the freshness of the random number *r2*.

P10: $T \mid\equiv \sharp\,(r2)$, T believes the freshness of the random number *r2*.

P11: $R \mid\equiv \sharp\,(S\_i{+}1)$, R believes the freshness of the random number *S_i+1*.

P12: $T \mid\equiv \sharp\,(S\_i{+}1)$, T believes the freshness of the random number *S_i+1*.

P13: $R \mid\equiv \sharp\,(K\_i{+}1)$, R believes the freshness of the random number *K_i+1*.

P14: $T \mid\equiv \sharp\,(K\_i{+}1)$, T believes the freshness of the random number *K_i+1*.

P15: $T \mid\equiv R \mid\Rightarrow N$, T believes R has jurisdiction over N.

P16: $T \mid\equiv R \mid\Rightarrow P$, T believes R has jurisdiction over P.

P17: $T \mid\equiv R \mid\Rightarrow X$, T believes R has jurisdiction over X.

P18: $T \mid\equiv R \mid\Rightarrow Y$, T believes R has jurisdiction over Y.

P19: $R \mid\equiv T \mid\Rightarrow M$, R believes T has jurisdiction over M.

P20: $R \mid\equiv T \mid\Rightarrow Q$, R believes T has jurisdiction over Q.

5.4 The Proving Process of the Protocol

From Message 4 we have that $R \lhd \{D\}$, which means R had receive the message D. According to the initial assumptions P2, P5 and the message-meaning rule $\dfrac{R \mid\equiv R \overset{K}{\longleftrightarrow} T, P \lhd \{X\}_K}{P \mid\equiv Q \mid\sim X}$, it follows $R \mid\equiv T \mid\sim D$.

Table 3: Performance comparison of protocols

Operation	Ref.[11]	Ref.[13]	Ref.[14]	Ref.[15]	Our protocol
Operation A	0T1	1T1	0T1	3T1	2T1
Operation B	2T2	1T2	19T2	8T2	9T2
Operation C	1T3	0T3	0T3	5T3	3T3
Operation D	0T4	0T4	0T4	1T4	1T4
Operation E	0T5	4T5	5T5	0T5	0T5
Operation F	0T6	0T6	3T6	3T6	3T6
Operation G	0T7	0T7	1T7	0T7	0T7
Operation H	0T8	3T8	5T8	0T8	0T8
Operation I	2T9	1T9	2T9	2T9	2T9
Storage capacity	1L	3L	4L	3L	3L

Then by the initial assumptions P7, P9 and the freshness-concatenation rule $\frac{P \mid\equiv \sharp(X)}{P \mid\equiv \sharp(X,Y)}$, it follows $R \mid\equiv \sharp\ (D)$.

Because of the conclusions $R \mid\equiv T \mid\sim D$, $R \mid\equiv \sharp\ (D)$, that we have proved above and the nonce-verification rule $\frac{P \mid\equiv \sharp(X), P \mid\equiv Q \mid\sim X}{P \mid\equiv Q \mid\equiv X}$, we have that $R \mid\equiv T \mid\equiv D$.

Finally, according to the corollary $R \mid\equiv T \mid\equiv D$, the initial assumption P20 and the jurisdiction rule $\frac{R \mid\equiv T \mid\Rightarrow Q, P \mid\equiv Q \mid\equiv X}{P \mid\equiv X}$, it can be proved that $R \mid\equiv D$. Thus, the proof of target G1 is now completed.

The target G2, G3, G4, G5 and G6 can be proved in a similar way as shown above.

6 Performance Analysis

The tag calculation complexity, the tag storage space and other several aspects are used for performance analysis.

As shown in Table 3, in [15] the tag stores t_i, with storage capacity of 1L. In [5] the tag stores *ik*, *uk* and *id*, with storage capacity of 3L. In [6] the tag stores *h(TID)*, *KTID*, *r* and *n*, with storage capacity of 4L. In this paper the tag stores *IDS*, t_i, and *Nx*, with storage capacity of 3L.

Operation A represents '+' operation, the operation time of which is represented by T1. Operation B represents '$\oplus$' operation, the operation time of which is represented by T2. Operation C represents Rabin encryption, the operation time of which is represented by T3. Operation D represents the function *MIXBITS(x, y)*, the operation time of which is represented by T4. Operation E represents Hash function, the operation time of which is represented by T5. Operation F represents *mod* operation, the operation time of which is represented by T6. Operation G represents CRC function, the operation time of which is represented by T7. Operation H represents PENG operation, the operation time of which is represented by T8. Operation I represents comparison operation, the operation time of which is represented by T9. Moreover, the time represented by T1 to T9 is different, some operations to spend a long time, some operations to take a short time. To sum up, the overhead cost of the protocol presented in this paper is acceptable.

Compared the improved protocol in this paper and the protocol in [13], both of the tag storage space are similar. In terms of the tag calculation, the improved protocol has twice less square operations than the original protocol. Although calculation complexity is not much reduced, it is found that the improved protocol solves the security flaws in the original protocol without increasing the calculation complexity of the tag. The original protocol can not resist the replay attacks and can not resist the de-synchronization attacks, however, the improved protocol can resist them. Compared with the protocols in [5, 6, 15], the tag storage space of the improved protocol is similar to theirs. What's more, it reduces the total calculation complexity of the tag, and meanwhile compensates for the security flaws in the protocols above.

7 Conclusion

An improved lightweight RFID tag ownership transfer protocol is proposed for the security problems of current ownership transfer protocol in [13]. Aiming at the problem that the tag's new owner in the original protocol can not resist the de-synchronization attacks caused by the replay messages, the improved protocol introduces the concept of the counter count for message Q. According to the value of *count*, different operations are used so as to solve the de-synchronization problems. If the value of *count* does not exist or is 0, the tag's new owner will generate new random numbers, otherwise won't, which makes it possible to avoid the problem that the shared private key between the both is not synchronized because the random number is generated after the message Q is received multiple times. Finally, a comprehensive security analysis shows that the improved protocol meets the security requirements of the tag ownership transfer.

Acknowledgments

This work was supported in part by the Natural Science Foundation of China under Grant 61472090, Grant 61472089 and Grant 61672169, in part by the Science and Technology Project of Guangdong Province under Grant 2015B010128014 and Grant2016B010107002,in part by the Science and Technology Planning Project of Guangzhou under Grant 201707010492.Grant 201604016003,Grant 201604016067 and Grant 201604016041.

References

[1] M. A. Chang-She, "Low cost RFID authentication protocol with forward privacy," *Chinese Journal of Computers*, vol. 34, no. 8, pp. 1387–1398, 2011.

[2] C. L. Chen, Y. L. Lai, C. C. Chen, Y. Y. Deng, and Yu C. Hwang, "RFID ownership transfer authorization systems conforming epcglobal class-1 generation-2 standards," *International Journal of Network Security*, vol. 13, no. 1, pp. 41–48, 2011.

[3] Y. C. Chen, W. L. Wang, M. S. Hwang, "RFID authentication protocol for anti-counterfeiting and privacy protection", in *The 9th International Conference on Advanced Communication Technology*, pp. 255–259, 2007.

[4] P. Y. Cui, "An improved ownership transfer and mutual authentication for lightweight RFID protocols," *International Journal of Network Security*, vol. 18, no. 6, pp. 1173–1179, 2016.

[5] R. Doss, W. Zhou, and S. Yu, "Secure RFID tag ownership transfer based on quadratic residues," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 2, pp. 390–401, 2013.

[6] S. Fouladgar and H. Afifi, "An efficient delegation and transfer of ownership protocol for RFID tags," *The First International Eurasip Workshop on RFID Technology*, 2007.

[7] M. S. Hwang, C. H. Wei, C. Y. Lee, "Privacy and security requirements for RFID applications", *Journal of Computers*, vol. 20, no. 3, pp. 55–60, Oct. 2009.

[8] Y. Jin, Q. Wu, Z. Shi, X. Lu, and L. Sun, "RFID lightweight authentication protocol based on PRF," *Journal of Computer Research and Development*, vol. 51, no. 7, pp. 1506–1512, 2014.

[9] G. Kapoor and S. Piramuthu, "Single RFID tag ownership transfer protocols," *IEEE Transactions on Systems Man and Cybernetics Part C*, vol. 42, no. 2, pp. 164–173, 2012.

[10] L. Lu, "Wireless key generation for RFID systems," *Chinese Journal of Computers*, vol. 38, no. 4, pp. 822–832, 2015.

[11] D. Molnar, A. Soppera, and D. Wagner, "A scalable, delegatable pseudonym protocol enabling ownership transfer of RFID tags," in *International Conference on Selected Areas in Cryptography*, pp. 276–290, 2005.

[12] Q. Qian, Y. L. Jia, R. Zhang, "A lightweight RFID security protocol based on elliptic curve Cryptography," *International Journal of Network Security*, vol. 18, no. 2, pp. 354–361, 2016.

[13] J. W. Shen and J. Ling, "Improved ultra-lightweight authentication of ownership transfer protocol for RFID tag," *Computer Science*, vol. 41, no. 12, pp. 125–128, 2014.

[14] B. Song and C. J. Mitchell, "RFID authentication protocol for low-cost tags," in *ACM Conference on Wireless Network Security (WISEC'08)*, pp. 140–147, 2008.

[15] B. Song and C. J. Mitchell, "Scalable RFID security protocols supporting tag ownership transfer," *Computer Communications*, vol. 34, no. 4, pp. 556–566, 2011.

[16] S. Wang, S. Liu, and D. Chen, "Scalable RFID mutual authentication protocol with backward privacy," *Journal of Computer Research and Development*, vol. 50, no. 6, pp. 1276–1284, 2013.

[17] 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.

[18] H. U. Wei, L. I. Yong-Zhong, and L. I. Zheng-Jie, "New defending RFID authentication protocol against dos attacks," *Application Research of Computers*, vol. 29, no. 2, pp. 676–675, 2012.

[19] M. H. Yang, "Secure multiple group ownership transfer protocol for mobile RFID," *Electronic Commerce Research and Applications*, vol. 11, no. 4, pp. 361–373, 2012.

Biography

Rui Xie received his B.S. in electrical engineering and automation from Dalian Maritime University in 2000, and the M.Sc. in Computer Science from Guangdong University of Technology in 2003. He is currently a Ph.D Candidates in Guangdong University of Technology. His research interests cover a variety of different topics including network security, machine learning, cloud computing, data mining and their applications.

Bi-yuan Jian received a master's degree in School of Computer Science and Engineering from South China University of Technology (China) in June 2011. He is a lecturer in School of Electronic and Computer Engineering in Guangzhou Vocational College of Science and Technology. His current research interest fields include information security and computer application.

Dao-wei Liu received a master's degree in School of Computers from Guangdong University of Technology (China) in June 2016. He is a lecturer in School of Electronic and Computer Engineering in Guangzhou Vocational College of Science and Technology. His current research interest fields include information security.

Dynamic Trust Model for Vehicular Cyber-Physical Systems

Hongzhuan Zhao[1,2], Dihua Sun[1], Hang Yue[3], Min Zhao[1], Senlin Cheng[1]
(Corresponding author: Dihua Sun)

Key Laboratory of Cyber Physical Social Dependable Service Computation & Chongqing University[1]
Area A, No.174, Shazheng street, Shapingba district, Chongqing, China
(Email: corresponding_d3sun@163.com)
School Architecture and Transportation Engineering & Guilin University of electronic Technology[2]
No.1, Jingji road, Qixing district, Guilin Guangxi, China
Johns Hopkins Healthcare LLC & Glen Burnie[3]
Johns Hopkins HealthCare LLC, 6704 Curtis Court Glen Burnie, MD 21060, USA
(Received Oct. 26, 2016; revised and accepted Feb. 20 & Mar. 28, 2017)

Abstract

Trust is a useful model for the interactions of Vehicular Ad Hoc Network (VANET) in Vehicular Cyber-Physical System (VCPS). Given a dynamic nature in transportation, traditional static trust models cannot effectively create the trust relationship among moving vehicles, and cannot handle quickly and dynamically the frequent vehicular interactions in a network topology. A novel trust model of VANET in VCPS is proposed to theorize the trust relationship in the dynamic traffic environment and perform a verification through an improved trust chain and some trusted computing theories. This trust model developed can improve the vehicular interaction security and the driving safety and resist malicious attacks and deceptions. Another, the vehicular trustworthiness is evaluated for the trust model development. The simulation experimental results show that the proposed trust model has a better performance for transportation applications than traditional models.

Keywords: Pervasive Trust Management Model; Trust Model; Trust-based Secure Service Discovery Model; Vehicular Ad Hoc Network; Vehicular Cyber-Physical System

1 Introduction

Transportation is a dynamic environment, consists of static roadways and moving objects (e.g. running vehicles and walking pedestrians) [39]. The dynamic transportation environment has a requirement on VANET for an efficient and effective interaction or communication among vehicles. But traditional methods in the wireless networks view moving vehicles as static or low-speed moving nodes. Using VANET the Vehicle to Vehicle (V2V) communication can obtain sufficient information in traffic cyber systems, and can realize a cooperative driving in traffic physical systems [16]. VCPS is the integration of the vehicular cyber system and the vehicular physical system [30, 43]. VANET in VCPS contains a series of wireless and independent moving nodes (i.e. vehicles) and temporarily form a network without a pre-existing infrastructure.

Intelligent Transport System (ITS) consists of different transportation systems, such as advanced traveler information system, advanced traffic management system, advanced transit system, and so on [35, 38]. For the development of the next generation ITSs, VCPSs, with the spread of mobile computing and communicating devices, are expected to solve some problems related to dynamic traffic moving objects [24]. Another, there are some security and trust problems related to VANET. The potential spoofing, eavesdropping, denial-of-service and impersonation attack lower the user trust on the cyber system services, and cause the safety and efficiency problems on the physical systems [7]. The trust approaches of static nodes, such as PKI [5] and CA [8], cannot meet with the demand of dynamic vehicular interactions.

Given a dynamic nature in transportation, it is hypothesized that a more efficient VANET in VCPS is required to guarantee dynamic trust interactions among vehicles with the adequate security and reliability [21]. The aim of this paper is to innovatively design a distributed dynamic trust model to describe the logic relationship among moving vehicles. The trust model can resist malicious attacks and deceptions based on the evaluation and certification mechanisms. Moreover, the proposed distributed trust architecture can reduce the communication load of moving vehicles.

Section 2 reviews the related literature. Section 3 depicts the detailed description and analysis of the trust

model process. The experimental simulation analysis is detailed in Section 4. Finally, Section 5 gives a brief discussion and some concluding remarks.

2 Literature Review

Some researches discusses the security issues and the potential solutions for the trust model development. For instances, the research [17] offers a novel communication method that cannot easily be subjected to network sniffing, thereby addressing the issue of security. Nevertheless, the scheme does not prevent malicious nodes from selectively forwarding packets or from other malicious behavior; although the study [6] gives an improved scheme to solve the impersonation attack and a malicious user can generate a valid signature on behalf of the other vehicles, the mechanism does not consider every nodes function and the global characteristics adequately; a secure and distributed certification system architecture for safety message authentication in VANET [26]. Moreover, this system can resist the false public-key certification. However, this study lacks the trust process in the entire system; the paper [9] uses the confidence intervals to manage uncertain information in the assessment of trust and reputation in ubiquitous network environments. Unfortunately, the characteristics of the focused environment has not been discussed in this study; in terms of the message receiver's authenticity and user privacy preservation, the research [19] eliminates the vulnerabilities in Chuang-Lee's scheme [11], and creates a trust-extended authentication scheme in VANET. Yet this research considers a little about the service levels.

In terms of the location-based verification security, the authors in the paper [28] make a summary about some previous researches related to map or Geographic Information Systems (GIS) [1, 12, 14, 22, 41, 42], and then proposes the Location Information Verification cum Security (LIVES) based on Transferable Belief Model (TBM). This map-based model includes two verification layers: the first layer depends on the concept of virtual tiles on roads and received signal strength; the second layer depends on the trust of neighboring vehicles computed using TBM. LIVES makes an improvement on the verification security in the map-based realistic network environments. However, all of the above location-based or map-based models ignore the time dimension or do not adequately employ time parameters for the trust model development. Considering the security verification of the real-time vehicular networks, these models are difficult to meet with the demand of dynamic vehicular interactions.

Another, data mining algorithms are applied into trust models. Take, for examples, the paper [29] creates a trust-based authentication scheme for the cluster-based VANETs. For the selection of cluster heads, the trust degree of each node is estimated based on the vehicles clustered. But it lacks the process of local interaction among mobile nodes; the study [4] utilizes a weighted clustering trust model algorithm for the development of Mobile Ad Hoc Networks (MANETs) trust. However, this study does not describe the local behaviors in details; the paper [32] uses the logic regression to model dynamic trust for service-oriented MANETs and dynamically estimate the service provider trust depending on the distinct behavior patterns [20]. But the proposed model does not detail the global trust relationship; a fuzzy-based dynamic trust model with the time slice scheme is developed to guarantee that a reliable node possess enough time to enjoy its services [33]. Nevertheless, the trust relationship is ignored.

Besides the above two dynamic trust models, Pervasive Trust Management (PTM) [3] describes dynamic inter-domain and trust model with the support of the Dempster-Shafer evidence theory, and uses the probability weighted average method for evaluating trustworthiness. PTM has several advantages: a) the trustworthiness of PTM rises slowly with the increase of the real trust behaviors and reduces sharply with the increase of the malicious trust behaviors; b) the trustworthiness changes dynamically according to different time and context information, and shows the dynamic characteristics of the trust process; and c) the algorithm can suitably describe the characteristics of VANET in VCPS. The disadvantages of PTM include: a) it cannot fit different requirements in various environments because of the static trust domain; and b) it cannot handle the uncertainty of certain nodes certification if missing data.

As a hybrid model, Trust-based Secure Service Discovery Model (TSSDM) [2] allows both of secure and non-secure service discoveries to handle the security issues related to the sharing communication and service. This model also permits mutual trust for the service discovery and sharing, even the service sharing with unknown entities. TSSDM have the following features: a) due to the good adaptive ability, different service levels from TSSDM have different trust levels. Also, different service levels provide different security levels. But the information load is heavy; and b) when the unknown entities join in TSSDM, this model would offer the risk certification mechanism to verify the unknown entities. However, the trustworthiness calculation only depends on the service time.

The above overall analysis reveals that previous research studies have not well developed the distributed trust model for dynamic vehicular interactions and service requirement discoveries, and have not adequately explored the evaluation and certification mechanisms for the malicious attacks and deceptions of moving vehicles. The trust model developed is expected to avoid the disadvantages of these models, but also keep their advantages. In contrast to existing models, PTM and TSSDM is better that the other models for transportation moving object-based trust verification. It is necessary to design experimental analyses for the comparison of the proposed model with both PTM and TSSDM. The next section gives the processes of the proposed trust model in details.

3 Trust Model Processes

There are four processes in the proposed trust model, and they are the trusted VANET initialization in VCPS, the service requirement discovery, the distributed evaluation and certification, and the trust transition based on the computing theories [27, 31]. The execution of these trust processes can achieve the vehicular verification and make the whole VANET trusted.

3.1 Trusted VANET Initialization

The system initialization dealer has a long-term stable trust relationship with vehicles [13]. Each vehicle obtains its own private share from the system initialization dealer. Generated by the system initialization dealer, each private share $S(i)$ is a randomly $(k-1)$ degree polynomial function, which is shown as follows:

$$S(i) = a_0 + a_1 x^1 + ... \quad + a_{k-1} x^{k-1} (\mathrm{mod}\varphi)$$

Where i is the unique identifier of Vehicle i in VANET; each private share is evaluated as $S(i)$; φ is a large prime number; and the initial key of the trusted VANET is $S_0 = a_0(\mathrm{mod}\ \varphi)$.

When an unknown vehicle was accessing the trust system, the system would send a public key with a trust evidence to the system initialization dealer, and request a certification for this public key. The system initialization dealer verifies the certification request via the calculation of the partial private share [10] in Equation (1):

$$\begin{aligned} S(i|j) &= S(j) \prod_{r=1, r\neq j}^{k} \frac{i-r}{j-r} (\mathrm{mod}\varphi) + \\ &\quad \sum_{r=1, r\neq j}^{k} \eta(j-r) S(jr) (\mathrm{mod}\varphi). \quad (1) \end{aligned}$$

$$\eta(j-r) = \begin{cases} 1, & j-r > 0 \\ 0, & j-r = 0 \\ -1, & j-r < 0 \end{cases} \quad (2)$$

Where j is the ID of the system initialization dealer; each pair of vehicles$(j,\ r)$ in the system exchange a number *S(jr)*; and *(j-r)* is the sign function.

After all vehicles obtain their individual private shares, each node of VANET would generate a partial certificate to other nodes. It would form a trust graph composed of partial certificates. Given that there are sparse social trust relationships among initial vehicles, the system would be fully functional, and no infrastructure would be expected, the system dealer would not be needed any longer [25]. Figure 1 illustrates the process of the trusted VANET initialization with the trust graph.

3.2 Service Requirement Discovery

The main role of the service requirement discovery is to handle the service applications. As a host, the trusted vehicle of VANET in VCPS evaluates if the trustworthiness of the guest vehicle satisfies the requirements of a certain service level or not via the trust certification. The host would gain the recommendations from the neighboring vehicles, and the recommendation information is the estimation information of the guest trustworthiness. It is the initialization mechanism for the trustworthiness of the unknown applied vehicles.

The different service levels have different security levels, and the security level Q_i maps the required service level S_i. The security factor with the range [0, 1] indicates the relationship between Q_i and S_i. Also, the proposed trust model uses a communication encrypted threshold CT and an authorized intervention threshold AT. When x>CT, the communication data would be encrypted; when x>AT, even though the guests satisfy the trust verification requirements, the interaction would be authorized by the host [23]. The trust mapping between Q_i and S_i is defined as follows:

$$Q_i(x) = \begin{cases} S_n, a_n < x \leq 1 \\ S_{n-1}, a_{n-1} \leq x < a_n \\ ... \\ S_k, A_T < x \\ ... \\ S_{k-1}, C_T < x \\ ... \\ 1, a_1 \leq x < a_2 \\ 0, 0 \leq x < a_1 \end{cases}$$

The security requirements are used for the selection of the above independent coefficients $a_1, a_2, ..., a_{n-1}, a_n \in [0, \varphi - 1]$. The security factors from the security requirements would improve the flexibility and self-adaptability of the proposed trust model. Besides, the communication encryption threshold and the authorized intervention threshold would determine different security and service levels.

3.3 Distributed Evaluation and Certification

On the basis of PTM and TSSDM, the collaborative trustworthiness between the host and the guest is estimated using the metrics, these metrics are related to the recommendation information from neighbored vehicles. The combination of the evaluation and certification mechanisms would be valid to identify the false recommendation. Given that each vehicle has a unique ID, the number of vehicles in VANET is n, and Vh_1,Vh_2,...,Vh_n as a vehicle set, the trust set of VANET in VCPS is given as [44]:

$$\begin{aligned} T_i = &[t(Vh_i, Vh_1) t(Vh_i, Vh_2) \\ &..., t(Vh_i, Vh_{i-1}), t(Vh_i, Vh_{i+1}), ..., t(Vh_i, Vh_n)] \end{aligned}$$

Where t(Vh_i, Vh_k) is the trustworthiness between the host Vh_i and the guest Vh_k, t(Vh_i, Vh_k) = null means

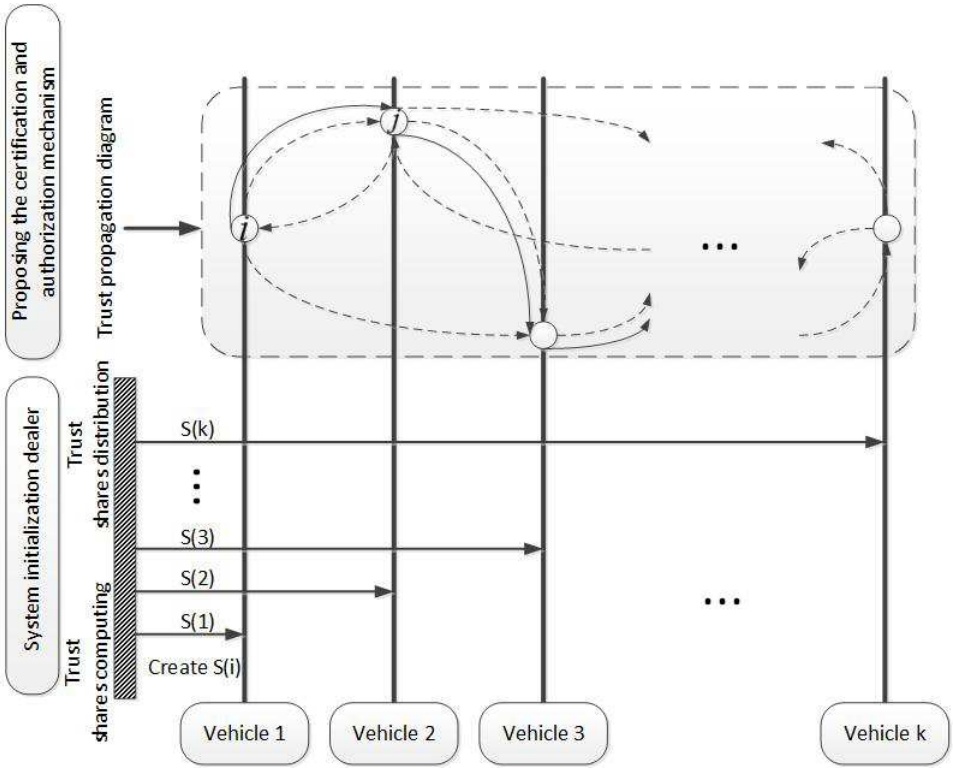


Figure 1: Trusted VANET initialization process

that the host Vh_i does not know the guest Vh_k for $1 \leq k \leq n$, $k \neq i$.

The vehicle set T_i denoted by S_{T_i} is the set of all the vehicles T with $t(Ti, T) \neq null$. The common vehicles set of T_i and T_j are given as follows:

$$\begin{cases} C(T_i,T) = [t(T_i,T_{k_1}), t(T_i,T_{k_2}), ..., t(T_i,T_{k_m})] \\ C(T,T_i) = [t(T_i,T_{k_1}), t(T_i,T_{k_2}), ..., t(T_i,T_{k_m})] \\ (T_{k_1}, T_{k_2}, ..., T_{k_m}) = S_{T_i} \bigcap S_{T_j} \end{cases}$$

Where $C(T_i, T)$ is the trustworthiness of T_i for all the common vehicles on T_i and T_j , and $C(T, T_i)$ is the trustworthiness of the common vehicles on T_i. This method can find the common vehicles on T_j.

The collaborative trustworthiness has a transitive property. If A_1 trusted A_2 and A_2 trusted B, then A_1 would trust B. Given that A_2 was an intermediary note, if A_2 was maliciously attacked and tampered, the entire trust chain from A_1 to B would be invalid. Figure 2 gives the comparison of the collaborative trust mechanism and the single trust mechanism.

The trustworthiness recommendation for the interaction between T_i and T_j is described as follows:

$$RIV(T_i,T_j) = \begin{cases} \frac{\gamma \cdot \overrightarrow{C(T_i,T)} \cdot \overrightarrow{C(T,T_j)}}{m}, S_{T_i} \bigcap S_{T_j} \neq \phi \\ 0, S_{T_i} \bigcap S_{T_j} = \phi \\ (m = |S_{T_i} \bigcap S_{T_j}|) \end{cases}$$

Where γ is the weight of the trustworthiness recommendation; $\overrightarrow{C(T_i,T)} \cdot \overrightarrow{C(T,T_j)}$ is the dot product of the collaborative trustworthiness betweenT_i and T_j; and $0 \leq RIV(T_i,T_j) \leq 1$ for any two trust sets T_i and T_j.

Given that and represent the number of the interactions between T_i and T_j with the limit of m, the confidence FIV on RIV for the vehicle sets T_i and T_j is given as follows:

$$\begin{cases} FIV(T_i,T_j) = \frac{(1-\frac{1}{m+\lambda})+(1-\frac{1}{N_{T_i}+\lambda})}{2} \\ FIV(T_j,T_i) = \frac{(1-\frac{1}{m+\lambda})+(1-\frac{1}{N_{T_j}+\lambda})}{2} \end{cases}$$

λ is an attenuation factor, the more trustworthiness comes from the higher number of common vehicles in both of sets and the higher number of the historical interactions among the vehicles. The historical interaction evaluation of T_i and T_j is described in the Equation (3):

$$HIE(T_i,T_j) = 1 - \frac{1}{\max\{\beta \cdot [\omega_{SI} SI(T_i,T_j) - \omega_{FI} FI(T_i,T_j)], 0\} + 1}, \quad \lambda = 1 \tag{3}$$

Where β is the time sensitive factor, $SI(T_i,T_j)$ is the number of successful historical interactions using the perspective of T_i, $FI(T_i,T_j)$ is the number of failed interactions. It is the same to define T_j.

As the time stamp between vehicular sets P and T, $\tau_{(P,T)}$ is the interaction weight at the current time, but it would be smaller with the elapsing time. The evaluation

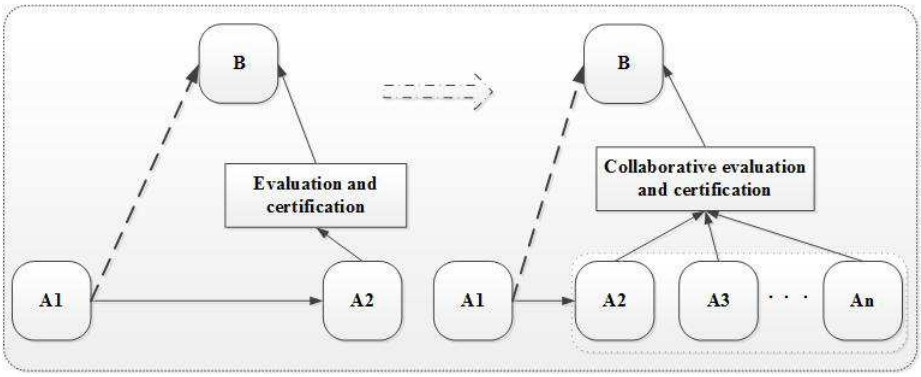


Figure 2: Collaborative and single trust mechanisms

of T_i and T_j is described as follows:

$$TTE(T_i, T_j) = \frac{m}{\sum_{z=1}^{m} \frac{\triangle\tau(T_j, T_{k_z})}{\triangle\tau}}$$

Where $\triangle\ \tau$ is the threshold time interval; the interaction between T_i and T_j at time τ; and $\triangle\ \tau(T_i, T_{k_z}) = \tau - \tau(T_i, T_{k_z}), \{T_{k_1}, T_{k_2}, ..., T_{k_m}\} = S_{T_i} \bigcap S_{T_j}$. The trustworthiness is evaluated by the weighted arithmetic mean of *RIV, FIV, HIE, and TTE.*
The trustworthiness of T_i and T_j is calculated in the following equation:

$$TR(T_i, T_j) = \frac{w_1[RIV(T_i, T_j)](\frac{FIV(T_i, T_j) + TTE(T_i, T_j)}{2}) + w_2[HIE(T_i, T_j)]}{w_1 + w_2}$$

3.4 Trust Transition

The improved Noninterference model [18] is used to design the trust transition mechanism with the support of the trusted computing theories. S represents the trust system of VANET, and the trust system contains the evaluated and certified vehicles, i.e. Vh_1, Vh_2,..., Vh_n.

$$S = \{Vh_1, Vh_2, ..., Vh_n\}$$

D is the true subset of S, D is the security domain mapped to the vehicles of the trust VANET. The trust relationship of VANET is described in Equation (4), and $Vhi \rightarrow Vhj$ means that Vhi has successfully verified Vh_j (i.e. Vh_i trusts Vh_j).

$$Vh_i \rightarrow Vh_j \in D \times D. \quad (4)$$

The trust chain of VANET is described in Equation (5), and Vhroot is the root of the trust delivering system and it is the beginning point of the trust chain.

$$Vh_{root} \rightarrow Vh_1 \rightarrow Vh_2 \rightarrow \cdots \rightarrow Vh_n, Vh_i \subset D. \quad (5)$$

The trust chain can be realized based on the preloading measurement technology of the trust computing theory (see Equation (6)). Remark(Vh_i,Vh_j) is the remark operation of the Vh_j verification using Vh_i, and expect(Vh_j) is the expected trustworthiness of Vh_j.

$$Remark(Vh_i, Vh_j) = expect(Vh_j) \Rightarrow Vh_i \rightarrow Vh_j. \quad (6)$$

Similar to the expected value expect(Vh_j), when Vh_i got the remark value of Vh_j via the remark operation, Vh_i would trust Vh_j. The trust relationship of VANET would be delivered from Vh_i to Vh_j. The role of the host would be delivered to Vh_j.

If $\forall Vh_1, Vh_2, Vh_3 \in S$, $[Vh_i \rightarrow Vh_j]$, $Vh_1, Vh_2 \in Vh_i$,$Vh_3 \in Vh_j$ and $Vh_1 \sim\succ Vh_2 \bigcup Vh_2 \sim\succ Vh_3 \Rightarrow Vh_1 \sim\succ Vh_3$, then the trust system of VANET would produce the unexpected interference via the transitivity between different domains and the trust chain would be invalid.

If $\forall Vh_i, Vh_j \subset D$, $Vh_1, Vh_2 \in Vh_i \bigcup Vh_3 \in Vh_j$ and $Vh_1 \sim\succ Vh_2 \bigcup Vh_2 \sim\succ Vh_3 \not=> Vh_1 \sim\succ Vh_3$, then $Vh_i \xrightarrow{intransitive\ noninterference} Vh_j \in D \times D$, and $Vh_i \xrightarrow{intransitive\ noninterference} Vh_j$ would be the intransitive noninterference relationship in $D \times D$. So there is not any unexpected interference in the trust VANET.

The intransitive noninterference relationship describes that there is just the direct interference relationship among the vehicles of VANET in VCPS. If the original trust chain of Trusted Computing Group (TCG) lacked enough security then it would cause an invalid trust chain, as described in Equation (7):

$$Remark(Vh_i, Vh_j) = expect(Vh_j) \not=> Vh_i \rightarrow Vh_j \quad (7)$$

With the aim of guaranteeing that vehicles move without interference and the vehicular data flows between different clusters in the trust VANET are restricted by a certain security policy, the trust chain delivery model is developed to construct effectively the trust chain. The proposed trust chain delivery model is given as follows:

$$Vh_i \xrightarrow{intransitive\ noninterference} Vh_j \bigcup Remark(Vh_i, Vh_j) = expect(Vh_j) \Rightarrow Vh_i \rightarrow Vh_j.$$

The trust chain delivery model reveals that the trust chain can be constructed and the trust relationship can be delivered, when vehicles communicate with each other and satisfy the intransitive noninterference relationships. The theorem development would fit the logic relationship of the intransitive noninterference, and guarantee the valid trust chain establishment and delivery. The four theorems are described as follows:

1) The domains of the trust VANET should keep the output-consistence. It means that the output influence of an inner interaction only relays on the view of the interaction domain in the trust VANET;

2) Caused by the interaction among vehicles, the influence of the trust VANET is just related to the previous view of the interaction domain, as described in Equation (8):

$$\begin{aligned} &Vh_i \overset{dom(Vh_{\text{int}eraction})}{\sim} Vh_j \wedge \\ &(contents(step(Vh_i, Vh_{\text{int}eraction}), Vh_{guest})) \neq \\ &contents(Vh_i, Vh_{guest}) \\ &\vee contents(step(Vh_j, Vh_{\text{int}eraction}), Vh_{guest}) \neq \\ &contents(Vh_j, Vh_{guest})) \\ &\rightarrow contents(step(Vh_i, Vh_{\text{int}eraction}), Vh_{guest}) = \\ &contents(step(Vh_j, Vh_{\text{int}eraction}), Vh_{guest}) \end{aligned} \tag{8}$$

Where contents(step(Vh_i, $Vh_{interaction}$), Vh_{guest}) is the trustworthiness of the guest Vh_{guest} in the state step(Vh_i, $Vh_{interaction}$) of the trust VANET, the single-step state transition function step(Vh_i, $Vh_{interaction}$) is the state of Vh_i after the interaction $Vh_{interaction}$ occurring among vehicles;

3) If the interaction among vehicles changed the value of the guest, then the interaction domain would modify the states of the guest, as described in Equation (9):

$$\begin{aligned} &(contents(step(Vh_i, Vh_{\text{int}eraction}), Vh_{guest})) \neq \\ &\qquad contents(Vh_i, Vh_{guest}) \\ &\rightarrow Vh_{guest} \in alter(dom(Vh_{\text{int}eraction}), Vh_i) \end{aligned} \tag{9}$$

Where alter(dom(Vh_i, $Vh_{interaction}$) is the altering set of the guest, and it can be modified under the state of Vh_i in dom($Vh_{interaction}$) of the trust VANET;

4) Any two domains in the trust VANET should satisfy the logic relationship of Equation (10):

$$\begin{aligned} &\exists Vh_{guest} \in N, Vh_{guest} \in alter(Vh_u, Vh_i) \wedge \\ &Vh_{guest} \in observe(Vh_v, Vh_i) \rightarrow \tilde{u} > v \end{aligned} \tag{10}$$

Where Vh_u and Vh_i are two interaction domains, observe(Vh_v, Vh_i) is the observing set of the guest, and this observing set can be monitored under the state Vh_i of the trust VANET in dom($Vh_{interaction}$).

Given that VI_c is the process of vehicular information collection, VI_s is the process of the vehicular information spreading,VI_p is the process of vehicular information processing and VI_d is the process of vehicular decision-making, the trust chain delivery model can be given in Equations (11) and (12). And the trust transition model of VANET in VCPS is shown in Figure 3:

$$\begin{aligned} &(VI_c^{root} \rightarrow VI_s^{root} \rightarrow VI_p^{root} \rightarrow VI_d^{root}) \bigcup \\ &VI_c^i \rightarrow VI_s^i \rightarrow VI_p^i \rightarrow VI_d^i \end{aligned} \tag{11}$$

$$Vh_{root} \bigcup Vh_1 \rightarrow Vh_2 \rightarrow \cdots \rightarrow Vh_n. \tag{12}$$

As the trust chain extends, the axiom of the trust decays is measured by the trustworthiness of a route, and the trust information of a remote vehicle is propagated by intermediate vehicles [34]. The trustworthiness $TC_{rn(cd)}(t)$ of the chain is calculated below:

$$TC_{rn(cd)}(t) = \prod (TC_t(T_i, T_j) | T_i, T_j \in Dand\ T_i \rightarrow T_j)$$

Where *Tr(c)* is the root, *Tn(d)* is the end of the trust, T_i and T_j are any two adjacent interaction vehicles, and $T_i \rightarrow T_j$ means that T_j is the next-hop node of T_i. T_r represents Vhroot, Tc represents VI_c^i, T_n represents Vh_n and T_d represent VI_d^i.

In addition, all intermediate vehicles are considered for the trustworthiness evaluation in the trust chains. The load conditions of VANET may be changed occasionally during the trustworthiness propagation, the trust would be changed accordingly. The latest arriving information would be used to calculate $TC_{rn(cd)}(t)$ of the trust chain, the scheme is adaptive to the change of VANET conditions, and the source information can be correctly delivered for a "propagation" in a timely manner. If one vehicle cheated another vehicle using false information when they interact with each other, vehicles could not accurately perceive current situation, the false information with security problem in cyber system may cause some physical safety problems, such as vehicle rear-end, crash, rollover, and so on.

4 Experimental Simulation Study

The simulation and calculation tools (Opnet Modeler 14.5 and Matlab 2014a) are used to analyze the characteristics of the trust model. There are totally 25 times simulation experiments for the calculation of the average trustworthiness. In the spatial-temporal interactions of VANET in VCPS, and are two time conversion factors, and both of their initial values are 0.5. The initial value of the spatial factor is 0. When the service level requirement factor S is 0.5, the host vehicle A and the guest vehicle B implement two times interaction during their initialization. The first interaction is the increasing trustworthiness behavior and

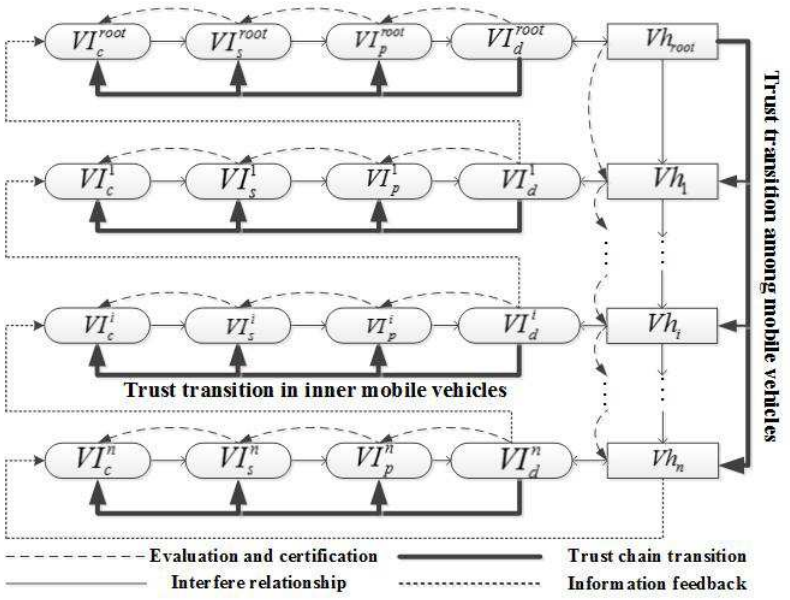


Figure 3: Trust transition model of VANET in VCPS

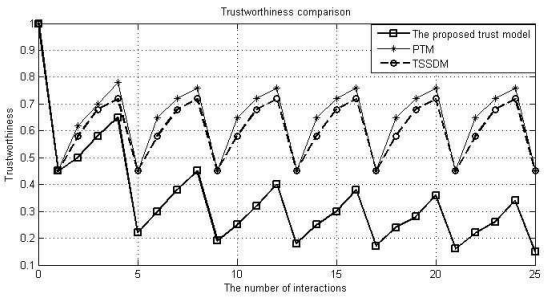


Figure 4: Capacity of resisting malicious deception in time dimension

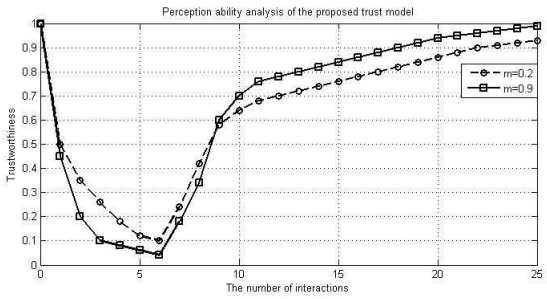


Figure 5: Perception ability in spatial dimension

the second interaction is the decreasing trustworthiness behavior. When the service level requirement factor S is 0.2, the guest B responses to the low service level requirements and continuously interacts with the host A for three times for the improvement of its trustworthiness. After obtaining the high trustworthiness, the guest B attacks the host A at a high service level domain, when the service level requirement factor S is 0.8 [15]. Figure 4 illustrates that the trustworthiness values of the proposed trust model, PTM and TSSDM between the host A and the guest B in a time dimension. The guest trustworthiness in the first scenario is some lower than two other trust models. With the trust increasing behaviors on the low service level requirements, the trustworthiness accumulation of the malicious node is much slower. With the trust decreasing behaviors on the important high service level requirements, the trustworthiness has a sharp decrease. So the proposed trust model can better resist the malicious deception in time.

Another, in a spatial dimension the proposed trust model uses the security sensitive factor to describe the perception ability of the vehicular status and traffic environment in VANET. The initialization information of the parameters is that the host vehicle A and the guest vehicle B implement their interaction for two times with λ =0.5, S=0.5 and m=0.5. One interaction is an increasing trustworthiness, and another interaction is a decreasing trustworthiness. And then they have six interactions in the decreasing trustworthiness behaviors and nineteen interactions in the increasing trustworthiness behaviors for both of m=0.2 and m =0.9 (see Figure 5). The interaction information and the tendency of the trustworthiness between A and B, the data with m=0.9 is more sensitive than m=0.2. The trustworthiness of the guest more sharply reduces with the decreasing trustworthiness interaction and more quickly increases with the increasing trustworthiness interaction in spatial dimension. The proposed trust model uses the service level requirement factor S to describe the perception ability in different service levels. It means that different service qualities may have different trustworthiness values during the interac-

tion among vehicles. The initialization information of the parameters is that the initial values of the time factors (i.e. β and λ) are 0.5, and the initial value of the spatial factor is 0. The host A and the guest B have two interactions when S=0.5. One interaction is the increasing trustworthiness behavior, and another interaction is the decreasing trustworthiness behavior. After that, the host A and the guest B have twenty-five interactions (see Table 1).

Table 1: Interaction parameters of Host A and Guest B

Interaction times	Service level requirement factor S	Trustworthiness decreasing (0) or increasing behaviors (1)
1	0.3	1
2	0.2	0
3	0.8	1
4	0.9	0
5	0.3	1
6	0.5	0
7	0.9	1
8	0.8	0
9	0.7	1
10	0.8	0
11	0.5	1
12	0.8	0
13	0.7	1
14	0.9	0
15	0.5	1
16	0.8	0
17	0.7	1
18	0.8	0
19	0.9	1
20	0.8	0
21	0.5	1
22	0.9	0
23	0.8	1
24	0.7	0
25	0.6	1

Figure 6 describes the changes of the trustworthiness between the host A and the guest B during the interaction process. The trustworthiness curve of the proposed model has a sharper fluctuation than both of PTM and TSSDM. It means that the proposed model is more sensitive to the service level requirement factor than PTM and TSSDM.

For the comparison of the proposed trust model with PTM and TSSDM in different maximum vehicular velocities. The time stamp period is denoted by (P,T), and the attenuation factor is denoted by . The moving vehicle uses the random waypoint model, and in this model each packet starts from a location to another at a random velocity [34]. The random waypoint model is used to describe the moving vehicle and Table 2 lists the fixed simulation parameters. As shown in the study [24], the simulation area consists of 66 sub-grid areas, and the range of communication endpoints (R) is 250 m.

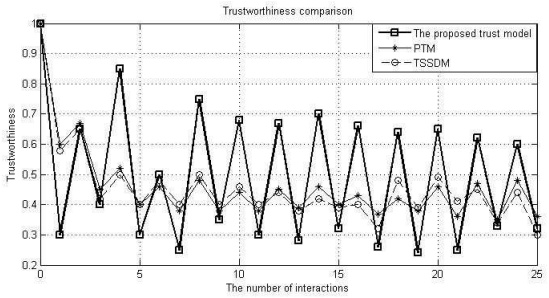


Figure 6: Perception ability in service levels

Also, three metrics, such as packet propagation ratio, average V2V latency and throughput of VANET, are used to evaluate the proposed model [33]. The calculation method of packet propagation ratio is that the number of the data packets delivered to the destination vehicles is divided by those sent by root vehicles; average V2V latency is the average time taken by the data packets from the starting points to the destinations; and throughput of VANET is the amount of the interaction information between the starting points and the destinations.

The propagation ratio of the proposed trust model is higher than PTM and TSSDM in any maximum velocity from 0 m/s to 30 m/s (see Figure 7). The packet propagation ratios in the three models have a more significant difference at low maximum velocities than at high maximum velocities. It means that the proposed trust model is more stable, this model can propagate more information to adjacent vehicles or infrastructures in a short time and can adapt to the dynamic changes related to the topology of the VANET. Figure 8 illustrates that the average V2V latency values rise with the increase of the maximum velocity. During the vehicular initialization of the trusted VANET, the trust chain roots are invalid more remarkably, and the trust chain roots initiate more trust chain rediscoveries before data interactions. The proposed trust model has a lower average V2V latency than PTM, and has a similar average V2V latency to TSSDM. It reveals that the proposed trust model can resist malicious deception more sensitively in the simulations than PTM, and it would be useful for the delay risk decline in the delivery of the failed interactive information packets. A lower packet propagation ratio means a less throughput of VANET. Figure 9 shows that the proposed trust model has a higher throughput than PTM and TSSDM over the whole range of the maximum velocity. It shows that the proposed trust model can bear more complex interaction contents, and it would reduce packet loss ratio of regional information spillover. From the above overall comparison the proposed model have a better performance than two other models in the experimental simulation analysis.

Table 2: Fixed simulation parameters

Parameters	Meaning	Value
simulation time	he period of simulation process	800 s
the number of nodes	the number of simulation vehicles	25
moving vehicular model	each packet starts from a location to another at a random speed	random waypoint
pause time	once the destination is reached, another destination is randomly chosen after a pause time	5 s
packet size	data payload size	512 bytes
$\tau_{(P,T)}$	the period of time stamp	30 s
λ	the attenuation factor	0.9
experiment times	the number of simulation times	25

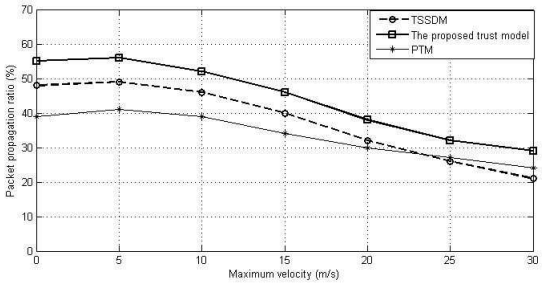


Figure 7: Packet propagation ratio

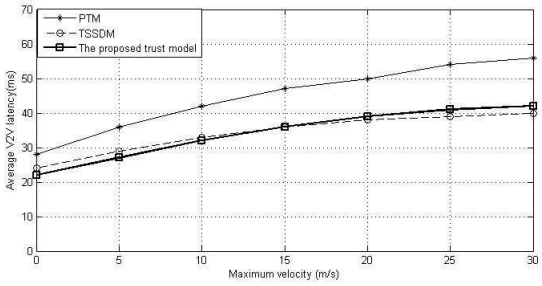


Figure 8: Average V2V latency

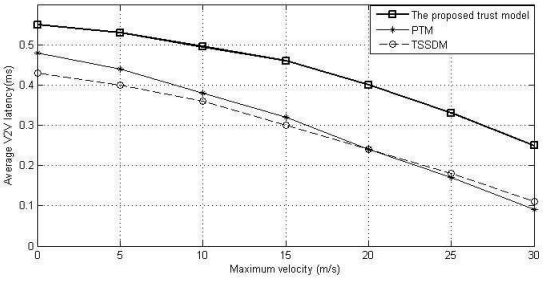


Figure 9: Packet propagation ratio

5 Concluding Remarks

This study discusses the development of the dynamic and distributed trust model for VANET in VCPS. This proposed trust model can describe the dynamic trust relationship among moving vehicles using the trust chain with a high accuracy. The trust chain can propagate the trust relationship based on the cryptography technology and the intransitive noninterference theory. Also, the verification mechanism developed can improve the reliability of the trust system. In addition, the distributed trust architecture in this new trust model can reduce the vehicular communication loads. The further research may focus on the combination of artificial intelligence and spatio-temporal databases [36, 37, 40] for the development of the proposed dynamic trust model in this study.

Acknowledgments

The authors gratefully acknowledge the research funding support from the National Natural Science Foundation of China (NSFC) (Grant No. 61573075), the core projects of Chongqing City 151 science and technology (Grant No. cstc2013jcsf-zdzxqqX0003), National Key R&D Program(Grant No.2016YFB0100904), Scientific and Technological Research Program of Chongqing Municipal Education Commission(Grant No.KJ1503301) and the Fundamental Research Funds of China Central Universities (Grant No. 106112014CDJZR178801).

References

[1] O. Abumansoor and A. Boukerche, "A secure cooperative approach for nonline-of-sight location verification in VANET," *IEEE Transaction on Vehicular Technology*, vol. 61, no. 1, pp. 275–285, 2012.

[2] S. I. Ahamed and S. Moushumi, "A trust-based secure service discovery (TSSD) model for pervasive computing," *Computer Communications*, vol. 31, no. 18, pp. 4281–4293, 2008.

[3] F. Almenrez, A. Marn, C. Campo, and C. Garcia, "PTM: A pervasive trust management model for dynamic open environments," in *The First Workshop on Pervasive Security, Privacy and Trust (PSPT'04)*, vol. 4, no. 7, pp. 1–8, 2004.

[4] M. Ashwin, S. Kamalraj, and M. Azath, "Weighted clustering trust model for mobile ad hoc networks," *Wireless Personal Communications*, vol. 3, no. 6, pp. 6–17, 2016.

[5] Q. H. Bai, Y. Zheng, L. N. Zhao, H. Chun, and C. Y. Cheng, "Research on mechanism of PKI trust model," *Applied Mechanics and Material, Trans Tech Publications*, vol. 536, pp. 694–697, 2014.

[6] 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.

[7] B. S. Bhati and V. Pallapa, "Performance analysis of location privacy preserving scheme for MANETs," *International Journal of Network Security*, vol. 18, no. 4, pp. 736–749, 2016.

[8] J. Braun, F. Volk, J. Classen, J. Buchmann, and M. Mhlhuser, "CA trust management for the web PKI," *International Journal of Network Security*, vol. 22, no. 6, pp. 913–959, 2014.

[9] G. Carullo, A. Castiglione, G. Cattaneo, A. De Santis, U. Fiore, and F. Palmieri, "Feeltrust: Providing trustworthy communications in ubiquitous mobile environment," in *IEEE 27th International Conference on Advanced Information Networking and Applications (AINA'13)*, pp. 1113–1120, 2013.

[10] J. H. Cho, A. Swami, and R. Chen, "A survey on trust management for mobile ad hoc networks," *IEEE Communications Surveys and Tutorials*, vol. 13, no. 4, pp. 562–583, 2014.

[11] M. C. Chuang and J. F. Lee., "TEAM: Trust-extended authentication mechanism for vehicular ad hoc networks," *IEEE Systems Journal*, vol. 8, no. 3, pp. 1758–1761, 2014.

[12] R. G. Engoulu, M. Bellache, and S. Pierre, "Vanet security surveys," *Computer Communications*, vol. 44, no. 1, pp. 1–13, 2014.

[13] F. Zhang, Z. P. Jia, H. Xia. Li, H. M. Sha Edwin, "Node trust evaluation in mobile ad hoc networks based on multi-dimensional fuzzy and markov SCGM(1,1) model," *Computer Communications*, vol. 35, no. 5, pp. 589–596, 2012.

[14] M. Fiore, C. Casetti, and C. F. Chiasserini, "Discovery and verification of neighbor positions in mobile ad hoc networks," *IEEE Transactions on Mobile Computing*, vol. 12, no. 2, pp. 289–303, 2013.

[15] H. Hu, R. Lu, Z. Zhang, and J. Shao, "Replace: A reliable trust-based platoon service recommendation scheme in VANET," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 2, pp. 1786–1797, 2016.

[16] D. Jia, K. Lu, and J. Wang., "On the network connectivity of platoon-based vehicular cyber-physical systems," *Transportation Research Part C: Emerging Technologies*, vol. 40, pp. 215–230, 2014.

[17] Y. Kobayashi, K. Totani, K. Utsu, and H. Ishii, "Achieving secure communication over manet using secret sharing schemes," *Journal of Supercomputing*, vol. 72, no. 3, pp. 1215–1225, 2016.

[18] X. Kong and Y. Zhuang., "Research on trust chain transfer model based on dynamic intransitive noninterference," *Journal of Convergence Information Technology*, vol. 7, no. 21, pp. 157–163, 2012.

[19] S. Kumari, M. Karuppiah, X. Li, F. Wu, A. K. Das, and V. Odelu, "An enhanced and secure trust-extended authentication mechanism for vehicular ad-hoc networks," *Security and Communication Networks*, vol. 9, no. 17, pp. 4255–4271, 2016.

[20] C. T. Li, M. S. Hwang, C. Y. Liu, "An electronic voting protocol with deniable authentication for mobile Ad hoc networks," *Computer Communications*, vol. 31, no. 10, pp. 2534–2540, June 2008.

[21] J. Lin, Y. Wu, G. Wu, and J. Xu, "An adaptive approach for multi-agent formation control in MANET based on CPS perspective," *Journal of Networks*, vol. 9, no. 5, pp. 1169–1177, 2014.

[22] C. Malandrino, F. Borgiattino and C. Casetti, "Verification and inference of positions in vehicular networks through anonymous beaconing," *IEEE Transaction on Mobile Computing*, vol. 13, no. 10, pp. 2415–2428, 2014.

[23] D. Marudhadevi, V. N. Dhatchayani, and S. Shankar, "A trust evaluation model for cloud computing using service level agreement," *The Computer Journal*, vol. 58, no. 10, pp. 2225–2232, 2014.

[24] M. N. Mejri, J. Ben-Othman, and M. Hamdi., "Survey on VANET security challenges and possible cryptographic solutions," *Vehicular Communications*, vol. 11, no. 2, pp. 53–66, 2014.

[25] M. Omar, C. Yacine, and B. Abdelmadjid, "Reliable and fully distributed trust model for mobile ad hoc networks," *Computers and Security*, vol. 28, no. 3, pp. 199–214, 2009.

[26] T. Oulhaci, M. Omar, F. Harzine, and I. Harfi, "Secure and distributed certification system architecture for safety message authentication in VANET," *Telecommunication Systems*, vol. 64, no. 4, pp. 679–694, 2016.

[27] H. Peng, D. Zhao, and Y. Yu, "Trust model based on trusted computing for distributed heterogeneous networks," *Computer Science*, vol. 10, no. 8, pp. 66–69, 2014.

[28] D. K. Sheet, O. Kaiwartya, A. H. Abdullah, Y. Cao, A. N. Hassan, and S. Kumar, "Location information verification using transferable belief model for geographic routing in vehicular ad hoc networks," *IET Intelligent Transport Systems*, vol. 11, no. 2, pp. 53–60, 2017.

[29] R. Sugumar, A. Rengarajan, and C. Jayakumar., "Trust based authentication technique for cluster based vehicular ad hoc networks (VANET)," *Wireless Networks*, vol. 29, no. 7, pp. 1–10, 2016.

[30] D. Sun, H. Zhao, H. Yue, M. Zhao, S. Cheng, and W. Han, "ST TD outlier detection," *IET Intelligent Transport Systems*, vol. 11, no. 4, 2017.
[31] Y. Uzunay and B. Kemal, "Trust-in-the-middle: towards establishing trustworthiness of authentication proxies using trusted computing," *Computer Science*, vol. arXiv preprint arXiv:1511.05682, pp. 1–32, 2014.
[32] Y. C. Wang, Y.and Lu, I. R. Chen, and J. H. Cho, "Logittrust: A logit regression-based trust model for mobile ad hoc networks," in *The 6th ASE International Conference on Privacy, Security, Risk and Trust*, pp. 1–10, Boston, MA, Mar. 2014.
[33] G. Wu, Z. Du, Y. Hu, T. Jung, U. Fiore, and K. Yim, "A dynamic trust model exploiting the time slice in WSNs," *Soft Computing*, vol. 18, no. 9, pp. 1829–1840, 2014.
[34] H. Xia, Z. Jia, X. Li, L. Ju, and H. M. Sha, "Trust prediction and trust-based source routing in mobile ad hoc networks," *Ad Hoc Networks*, vol. 11, no. 7, pp. 2096–2114, 2013.
[35] H. Yue, "Advanced traveler information inquiry, archiving, and decision making system," in *The 12th Guangzhou Convention of Overseas Chinese Scholars Science and Technology*, Guangdong, China, Dec. 2009.
[36] H. Yue, "Archiving capability of spatio-temporal data in different highway railroad grade crossing (HRGC) databases," in *The 20th Annual Intelligent Transportation System*, Houston, America, May 2012.
[37] H. Yue, E. Jones, and P. Z. Revesz, "Use of local polynomial regression models for average traffic speed estimation and forecasting in linear constraint databases," in *The 17th International Symposium on Temporal Representation and Reasoning*, pp. 154–161, Paris, France, Nov. 2010.
[38] H. Yue and R. Yang, "Development of intelligent transportation systems and plan of integrated information system," *Journal of Wuhan University of Technology*, vol. 29, no. 4, pp. 560–563, 2005.
[39] H. Yue and P. Z. Revesz, "TVICS: An efficient traffic video information converting system," in *The 19th International Symposium on Temporal Representation and Reasoning (TIME'12)*, pp. 141–148, Leicester, UK, Dec. 2012.
[40] H. Yue, L. R. Rilett, and P. Z. Revesz, "Spatio-temporal traffic video data archiving and retrieval system," *GeoInformatica*, vol. 20, no. 1, pp. 59–94, 2016.
[41] Y. Zeng, J. Cao, and J. Hong, "Secure localization and location verification in wireless sensor networks: A survey," *Journal of Supercomputing*, vol. 64, no. 3, pp. 685–701, 2013.
[42] P. Zhang, Z. Zhang, and A. Boukerche, "Cooperative location verification for vehicular ad-hoc networks," in *Proceedings of IEEE International Conference on Communications (ICC'12)*, pp. 37–41, Ottawa, ON, Canada, July 2012.
[43] H. Zhao, D. Sun, H. Yue, M. Zhao, and S. Cheng, "Using CSTPNS to model traffic control CPS," *IET Software*, vol. 11, no. 3, 2017.
[44] A. Zhou, J. Li, Q. Sun, C. Fan, T. Lei, and F. Yang, "A security authentication method based on trust evaluation in VANETs," *EURASIP Journal on Wireless Communications and Networking*, vol. 10, no. 3, pp. 59–68, 2015.

Biography

Hongzhuan Zhao(1985-), Male, is an Assistant Professor at School Architecture and Transportation Engineering, Guilin University of Electronic Technology, Guangxi, China. He gained Ph.D. from Chongqing University, China in 2016. His current work focuses on some research topics related to ITS, Cyber-Physical System (CPS), reliable perception, and trusted interaction.

Dihua Sun (1962-), Male, is a Professor at College of Automation, Chongqing University. He gained the B.S. from Huazhong University of Science and Technology in 1982, China, and obtained M.S. and PH.D. from Chongqing University, China, in 1989 and 1997, respectively. His research interests include CPS, ITS, computer-based control, data analysis and decision support.

Hang Yue (1977-), Male, obtained his M.S. degree in Transportation Engineering and Minors in Statistics and Geography from University of Nebraska-Lincoln in 2012 and M.S. degree in Computer Software Engineering from Zhejiang University in 2005. He served as a Data Analyst at AirSage Inc. in 2013, and a Data Scientist at Charter Global Inc. in 2014. He is currently a Data Analyst at Johns Hopkins Health Care LLC. His research interests are machine learning, big data business intelligence, data warehouse, data visualization, GIS, and spatio-temporal databases.

Min Zhao (1980-), Female, is an Associate Professor at College of Automation, Chongqing University. She gained Ph.D. from Chongqing University, China in 2010. Her research interests include CPS, ITS and the traffic image processing.

Senlin Cheng (1968-), Male, is an Associate Professor at College of Automation, Chongqing University. He gained Ph.D. from Chongqing University, China in 1999. His research interests include CPS, ITS and wireless location technology.

Recipient Anonymous Ciphertext-Policy Attribute-based Broadcast Encryption

Leyou Zhang, Hongjian Yin
(Corresponding author: Hongjian Yin)
School of Mathematics and Statistics, Xidian University
Xi'an, Shaanxi 710171, China
(Email: xidianyhj@163.com)
(Received Oct. 15, 2016; revised and accepted Feb. 20, 2017)

Abstract

The ciphertext-policy (CP) attribute-based broadcast encryption (CP-ABBE) is a more flexible broadcast encryption (BE), in which the broadcaster encrypts the data with an access policy and a receiver set. Only receivers in the valid set who satisfy the access policy will be able to decrypt the ciphertext. However, most existing CP-ABBE schemes only pay attention to plaintext privacy rather than access policy privacy and broadcast list privacy. It results in the fact that the adversary can determine the access policy or the broadcast set from ciphertexts and public parameters. However, in the real life, the access policy or the receiver set may be sensitive. To overcome this shortcoming, we propose a recipient anonymous CP-ABBE scheme, where it can protect the description of the access structures and broadcast sets associated with ciphertexts. The proposed scheme achieves full security based on the dual system encryption and constant size ciphertexts.

Keywords: Attribute-based Encryption; Broadcast Encryption; Fully Secure; Recipient Anonymity

1 Introduction

Broadcast encryption (BE) [7, 9, 11, 13] is a one-to-many encryption technique which is efficient to data sharing. It allows the broadcaster to send an encrypted message to a subset of privileged users only such listeners who are in this set can decrypt the ciphertext. In recent years, there have been many broadcast encryption schemes such as identity-based BE [15], attribute-based BE [18], and anonymous identity-based BE [23].

The notion of attribute-based encryption (ABE) was introduced by Sahai and Weters [21], which allows users to control their encrypted data at a fine-grained level. In ABE, the data owner can share their data with those users who have the specified attributes [4, 17]. There are two kinds of ABE involving ciphertext-policy ABE (CP-ABE) [3] and key-policy ABE (KP-ABE) [8, 19]. In a CP-ABE scheme, ciphertext is related to access structure and the private key of user is associated with an attribute set. Only the user whose private key satisfies the access structure associated with the ciphertext will be able to decrypt the ciphertext successfully. In contrast, in a KP-ABE scheme, ciphertext is related to an attribute set and the private key of user is associated with access structure [12, 24]. The user will be able to decrypt ciphertext only if the attributes associated with the ciphertext satisfy the access structure of the private key.

Attribute-based broadcast encryption (ABBE) was first proposed by David and Thomas [18], in which the broadcaster encrypted data with an access structure and a receiver list. Only receivers who satisfy the access policy and are in this list will be able to decrypt the ciphertext. As normal ABE, ABBE also allows fine-grained and flexible access control. However, compared with the traditional broadcast encryption, ABBE is a more flexible broadcast encryption and supply direct revocation by removing the revoked users from the receiver list. It is an important capability for real time applications such as Pay-TV.

After the first ABBE scheme [18], there have been proposed many efficient and provably secure ABBE schemes. Attrapadung and Imai proposed CP-ABBE and KP-ABBE [1] based on CP-ABE and KP-ABE, respectively. Both schemes are efficient revocable scheme. However, they only achieve selective security which is a weak security for ABBE. A strong secure ABBE scheme was proposed by Li and Zhang [16], where the scheme achieved full security by employing dual system encryption technique [22], but the ciphertext and decryption pairings grow linearly with the number of attributes and recipients. To improve the efficiency, Phuong et al. [20] proposed an ABBE scheme with short ciphertexts and private keys. Especially, their scheme achieves constant size ciphertexts and decryption pairings. However, its security is based on the decision n-Bilinear Diffie-Hellman exponent assumption which is a strong hardness assumption.

Nevertheless, all of the above mentioned ABBE schemes cannot achieve recipient anonymous, and it means that any intermediate user can only use public parameters to determine whether the ciphertexts are encrypted under the given access structure and receiver set or not. The recipient anonymity is an important property for encryption schemes. For instance, in Phuong's second ABBE scheme [20], an intermediate user can use some parts of the ciphertexts $C_1 = g^r$, $C_2 = (\prod_{j\in S^*} g_{n+1-j})^r$ to run the Decision Diffie-Hellman (DDH) test $e(C_1, \prod_{j\in S} g_{n+1-j}) \stackrel{?}{=} e(C_2, g)$, to determine whether the ciphertexts are encrypted under a given receiver set or not, where ν, g and g_l are the public parameters ($l = 1, 2, \cdots, 2n$). The maximum number of DDH-test is 2^{2n}, that is, the adversary run the DDH-test at most 2^{2n} times then he will be able to ascertain whether the S^* is broadcast list or not. Furthermore, the access structure of the CP-ABBE scheme [26] also can be determined by the DDH-test.

In this paper, we present a recipient anonymous CP-ABBE scheme. In the proposed scheme, both the access structure and the broadcast list are hidden. That is, any one cannot get any information about the access structure or the broadcast list by DDH-test from ciphertexts. Based on three static assumptions in composite order groups, our scheme is proven to be fully secure with the dual system encryption technique [22]. Furthermore, compared with some previously known ABBE schemes, the proposed scheme is an efficient CP-ABBE scheme in which the size of the ciphertexts and the number of pairings are at a constant size level.

The paper is organized as follows. In Section 2, some preliminaries are given. Section 3 gives the definition of recipient anonymous CP-ABBE scheme and its security model. The recipient anonymous CP-ABBE scheme is presented in Section 4. Security proof is introduced in Section 5. In Section 6, some comparisons between our scheme and previous works in security and efficiency are given. Finally, we conclude this paper in Section 7.

2 Preliminaries

Let $x \in_R X$ denote that x is randomly chosen from a set X.

2.1 Composite Order Bilinear Groups

The first composite order bilinear group was introduced by Boneh, Goh, and Nissim in 2005 [5]. Then it was used for many cryptographic constructions. This paper will use the bilinear group whose order is product of three distinct primes.

Let $\mathcal{G}(\cdot)$ be an algorithm that takes a security parameter λ as input and outputs a tuple $(N = p_1p_2p_3, \mathbb{G}, \mathbb{G}_T, e)$, where p_1, p_2, p_3 are distinct primes, $\mathbb{G}$ and $\mathbb{G}_T$ are cyclic groups of composite order $N = p_1p_2p_3$ and e: $\mathbb{G} \times \mathbb{G} \longrightarrow \mathbb{G}_T$ is a map such that

1) for all $g, h \in \mathbb{G}$ and $a, b \in \mathbb{Z}_N$, $e(g^a, h^b) = e(g, h)^{ab}$;

2) exists $g \in \mathbb{G}$ such that $e(g, h)$ has order N in $\mathbb{G}_T$.

Let $\mathbb{G}_{p_1}, \mathbb{G}_{p_2}$ and $\mathbb{G}_{p_3}$ denote the subgroups of order p_1, p_2 and p_3 in $\mathbb{G}$ respectively. And g_1, g_2 and g_3 are the generators of subgroups $\mathbb{G}_{p_1}$, $\mathbb{G}_{p_2}$ and $\mathbb{G}_{p_3}$ respectively. As Lewko and Waters [14] illuminated, when $h_i \in \mathbb{G}_i$, and $h_j \in \mathbb{G}_j$ for $i \neq j$, then $e(h_i, h_j) = 1$. This property is called orthogonal property of $\mathbb{G}_{p_1}, \mathbb{G}_{p_2}, \mathbb{G}_{p_3}$.

2.2 Complexity Assumptions

The security of our recipient anonymous CP-ABBE scheme will be reduced to three static assumptions [14]. And these assumptions are described below:

Assumption 1. *Given a group parameters generator $\mathcal{G}$, we define the following distribution: $\Theta = (N = p_1p_2p_3, \mathbb{G}, \mathbb{G}_T, e) \in_R \mathcal{G}$, $g_1 \in_R \mathbb{G}_{p_1}$, $X_3 \in_R \mathbb{G}_{p_3}$, $D = (\Theta, g_1, X_3)$, $T_1 \in_R \mathbb{G}$, $T_2 \in_R \mathbb{G}_{p_1p_3}$. Now the advantage of an algorithm $\mathcal{A}$ in breaking Assumption 1 is defined to be*

$$Adv^1_{\mathcal{A}} = |Pr[\mathcal{A}(D, T_1) = 1] - Pr[\mathcal{A}(D, T_2) = 1]|.$$

Assumption 2. *Given a group parameters generator $\mathcal{G}$, we define the following distribution: $\Theta = (N = p_1p_2p_3, \mathbb{G}, \mathbb{G}_T, e) \in_R \mathcal{G}$, $g_1 \in_R \mathbb{G}_{p_1}$, $X_i \in_R \mathbb{G}_{p_i} (i = 1, 2, 3)$, $Y_2 \in_R \mathbb{G}_{p_2}$, $D = (\Theta, g_1, X_1X_2X_3, Y_2)$, $T_1 \in_R \mathbb{G}_{p_1}$, $T_2 \in_R \mathbb{G}_{p_1p_2}$. Now the advantage of an algorithm $\mathcal{A}$ in breaking Assumption 2 is defined to be*

$$Adv^2_{\mathcal{A}} = |Pr[\mathcal{A}(D, T_1) = 1] - Pr[\mathcal{A}(D, T_2) = 1]|.$$

Assumption 3. *Given a group parameters generator $\mathcal{G}$, we define the following distribution: $\Theta = (N = p_1p_2p_3, \mathbb{G}, \mathbb{G}_T, e) \in_R \mathcal{G}$, $g_1 \in_R \mathbb{G}_{p_1}$, $X_3 \in_R \mathbb{G}_{p_3}$, $D = (\Theta, g_1, X_3)$, $T_1 \in_R \mathbb{G}_T$, $T_2 = e(g_1, g_1)^{\alpha s}$. Now the advantage of an algorithm $\mathcal{A}$ in breaking Assumption 3 is defined to be*

$$Adv^3_{\mathcal{A}} = |Pr[\mathcal{A}(D, T_1) = 1] - Pr[\mathcal{A}(D, T_2) = 1]|.$$

2.3 Access Structure

Our construction will employ AND-gate on multi-valued attributes access structure, which is similar to what used in [2, 6]. The access structure of AND-gate on multi-valued attributes is described as follows.

Let $\mathbb{U} = \{att_1, att_2, \cdots, att_n\}$ be a set of attributes. For $att_i \in \mathbb{U}$, $S_i = \{v_{i,1}, v_{i,2}, \cdots v_{i,m_i}\}$ is a set of possible values, where m_i is the number of possible values for each att_i. Let $L = [L_1, L_2, \cdots L_n]$ be an attribute list for a user where $L_i \in S_i$. Let $\mathbb{A} = [w_1, w_2, \cdots w_n]$ be an access structure where $w_i \in S_i$. The notation $L \models \mathbb{A}$ expresses that an attribute list L satisfies an access structure $\mathbb{A}$ and $\not\models$ refers to not satisfy symbol.

3 Definitions and Security Model

3.1 Definitions of CP-ABBE Scheme

A recipient anonymous ciphertext-policy attribute-based broadcast encryption (CP-ABBE) scheme consists of the following four algorithms:

Setup$(1^\lambda, \mathcal{U}, \mathcal{N})$**.** Take as input a security parameter λ, the broadcast index set $\mathcal{U}$ and the universal attribute set $\mathcal{N}$. Then this algorithm outputs a public parameters PK and a master secret key MSK.

KeyGen(MSK, k, L)**.** Take as input the master secret key MSK, the user's index $k \in \mathcal{U}$ and attribute set $L \subseteq \mathcal{N}$. Then this algorithm outputs the user's private key $SK_{(k,L)}$.

Encrypt$(PK, S, \mathbb{A})$**.** Take as input the public parameters PK, a broadcast index list $S \subseteq \mathcal{U}$ and an access structure $\mathbb{A} \in AS$, where AS is an access structure family over $\mathcal{N}$. Then this algorithm outputs a broadcast header Hdr and a massage encryption key K.

Decrypt$(SK_{(k,L)}, Hdr)$**.** Take as input the a private key $SK_{(k,L)}$ as well as a broadcast header, if $k \in S$ and $L \models \mathbb{A}$, then this algorithm outputs K.

3.2 Security Model

Following [10], we describe the indistinguishability against chosen plaintext attack (IND-CPA) definition of recipient anonymous CP-ABBE in the fully secure model. The formal secure game between adversary $\mathcal{A}$ and challenger $\mathcal{B}$ is as follows.

Setup. Assume universal attribute set $\mathcal{N}$, broadcast index set $\mathcal{U}$ and access structure family AS are predefined. The challenger $\mathcal{B}$ runs the **Setup** algorithm to obtain a public parameters PK and a master secret key MSK. Then it gives adversary $\mathcal{A}$ the public parameters PK and keeps MSK to itself.

Key Query Phase 1. The adversary queries the challenger $\mathcal{B}$ for private keys corresponding to index $k \in \mathcal{U}$ and attribute set $L \subseteq \mathcal{N}$. The challenger runs the **KeyGen** algorithm and gives the corresponding private keys $SK_{(k,L)}$ to $\mathcal{A}$.

Challenge. When the adversary decides that **Phase 1** is over, $\mathcal{A}$ outputs two same-length messages M_0 and M_1. The adversary also outputs a challenge broadcast index set S^* and access structure $\mathbb{A}^*$ such that for all index k and attribute set L queried in **Phase 1**, we have $k \notin S^*$ and $L \not\models \mathbb{A}^*$. Then $\mathcal{B}$ runs **Encrypt** algorithm to get $\langle Hdr^*, K_0 \rangle$ and randomly chooses $K_1 \in_R \mathcal{K}$, where $\mathcal{K}$ is the symmetric key space. It flips a coin $\mu \in \{0,1\}$ and gives $\langle Hdr^*, K_\mu \rangle$ to $\mathcal{A}$.

Key Query Phase 2. In this phase, $\mathcal{B}$ acts almost the same as in **Phase 1** except it is unable to ask key for attribute set L and index k such that $L \models \mathbb{A}^*$ and $k \in S^*$.

Guess. Finally, the adversary $\mathcal{A}$ outputs the guess bit $\mu' \in \{0,1\}$ for μ and wins the game if $\mu' = \mu$.

The advantage of the adversary in this game is defined as follows:

$$Game_{\mathcal{A}}(\lambda) = |Pr[\mu = \mu'] - \frac{1}{2}|,$$

where the probability is taken over the random bits used by the challenger and the adversary.

Definition 1. *A recipient anonymous CP-ABBE scheme is IND-CPA secure if for all polynomial time adversary $\mathcal{A}$, the $Game_{\mathcal{A}}(\lambda)$ is negligible.*

4 Recipient Anonymous CP-ABBE Scheme

In this section, we will present our recipient anonymous CP-ABBE scheme construction and show the recipient anonymity of our scheme by employing composite order bilinear groups. There are four algorithms in our scheme, which are defined in Section 3.1. First, we briefly summarized our idea. In order to realize recipient anonymity, some random numbers are added to each part of the ciphertexts. And these random numbers can prevent adversary from determining user information by running DDH-test. Thanks to employ composite order group, these random numbers will not affect the decryption process in our scheme. The detailed algorithms are described in the following. The abbreviations and notations used throughout the paper are shown in Table 1.

4.1 Construction

– **Setup**$(1^\lambda, \mathcal{N}, \mathcal{U})$**:** To generate the system parameters, the setup algorithm takes a security parameter λ, an universal attribute set $\mathcal{N}$ and a broadcast index set $\mathcal{U}$ where $|\mathcal{U}| = h$ as inputs. Then it runs the group generator $\mathcal{G}$ to get a description of bilinear composite order group $\Theta = (N = p_1p_2p_3, \mathbb{G}, \mathbb{G}_T, e)$. The algorithm picks random elements a, α in $\mathbb{Z}_N$, g_1, $u_{j'}$ in $\mathbb{G}_{p_1}$ and R_0 in $\mathbb{G}_{p_3}$, where $j' \in \mathcal{U}$. For each attribute $v_{i,j} \in \mathcal{N}$, the setup algorithm chooses random elements $a_{i,j}$ in $\mathbb{Z}_N$ and $R_{i,j}$ in $\mathbb{G}_{p_3}$. Then the setup algorithm computes $A_{i,j} = g_1^{a_{i,j}} \cdot R_{i,j}$, $A_0 = g_1 \cdot R_0$. The public parameters PK is defined as

$$PK = \langle e(g_1, g_1)^\alpha, A_0, A_{i,j}, \{u_{j'}\}_{j' \in \mathcal{U}} \rangle,$$

and the master secret key MSK is defined as $MSK = \langle g_1, \alpha, a, a_{i,j}, a_j \rangle$, where $1 \le i \le n, 1 \le j \le m_i$.

Table 1: Parameters declaration

Symbol	Description
p_i	The primes, where $i = 1, 2, 3$.
g_i	Generators with order p_i, over $\mathbb{G}_{p_i}$, where $i = 1, 2, 3$.
$\mathbb{Z}_N$	The set of positive integers.
MSK	The master secret key.
$SK_{(k,L)}$	The private key associated with attributes set L and user index k.
M	Message.
K	Symmetric key.
Hdr	The broadcast header.
$e(\cdot)$	Bilinear pairing.
$\|B\|$	The number of elements in set B.

– **KeyGen**(MSK, k, L)**:** Given a user index $k \in \mathcal{U}$, an attributes set $L = [v_{1,j_1}, v_{2,j_2}, \cdots v_{n,j_n}] \in \mathcal{N}$ and the master secret key, the key generation algorithm chooses a random element $r \in_R \mathbb{Z}_N$ and computes:

$$D_1 = g_1^{\alpha+ar} u_k^r, \qquad D_2 = g_1^r,$$
$$\{D_{3,j'} = u_{j'}^r\}_{j' \in \mathcal{U}\setminus\{k\}}, \quad D_4 = \left(g_1^{a+\sum_{v_{i,j_i} \in L} a_{i,j_i}}\right)^r.$$

Finally, this algorithm outputs the private key associated with attributes set L and user index k

$$SK_{(k,L)} = \langle D_1, D_2, D_{3,j'}, D_4 \rangle_{j' \in \mathcal{U}\setminus\{k\}}.$$

– **Encrypt**$(PK, S, \mathbb{A})$**:** Let $M \in \mathbb{G}_T$ be the message to be encrypted and let $\mathbb{A} = \wedge_{i=1}^n w_{i,j_i}$ where $w_{i,j_i} \in Att_i$, be an access policy and a broadcast set $S \subseteq \mathcal{U}$. A broadcaster randomly selects $s \in_R \mathbb{Z}_N$ and $R_1, R_2, R_3 \in_R \mathbb{G}_{p_3}$. Then this algorithm computes the symmetric key K and broadcast header Hdr as follows.

$$K = e(g_1, g_1)^{\alpha s}, \qquad C_1 = (\textstyle\prod_{v_{i,j_i} \in \mathbb{A}} A_{i,j_i})^s \cdot R_1,$$
$$C_2 = A_0^s \cdot R_2, \qquad C_3 = (\textstyle\prod_{j' \in S} u_{j'})^s \cdot R_3,$$
$$Hdr = (C_1, C_2, C_3).$$

In this system, K is used to encrypt the massage M in a symmetric encryption scheme. Note that a random element $R \in \mathbb{G}_{p_3}$ can be selected by choosing a random $\eta \in \mathbb{Z}_N$ and setting $R = g_3^\eta$ where g_3 is publicly given.

– **Decrypt**$(SK_{(k,L)}, Hdr)$**:** The decryption algorithm takes a broadcast header Hdr and a private key $SK_{(k,L)}$ as input. If the private key of the recipient $SK_{(k,L)}$ satisfies the policy of the ciphertext, then this algorithm will compute the symmetric key K as follows,

$$\frac{e(D_1 \cdot \prod_{j' \in S\setminus\{k\}} D_{3,j'}, C_2) \cdot e(D_2, C_1)}{e(D_2, C_3) \cdot e(D_4, C_2)}$$
$$= e(g_1, g_1)^{\alpha s}$$
$$= K.$$

4.2 Correctness

The correctness will subsequently be checked by applying the orthogonality property of $\mathbb{G}_{p_i}$ $(i = 1, 2, 3)$.

If the user index $k \in S$ and attributes set $L \models \mathbb{A}$, then one can obtain the below equations hold.

$$\begin{aligned} & e(D_1 \textstyle\prod_{j' \in S\setminus\{k\}} D_{3,j'}, C_2) \\ =& e(g_1^{\alpha+ar} u_k^r \textstyle\prod_{j' \in S\setminus\{k\}} u_{j'}^r, g_1^s R_0^s \cdot R_2) \\ =& e(g_1, g_1)^{\alpha s} \cdot e(g_1, g_1)^{ars} \cdot e(u_k^r \textstyle\prod_{j' \in S\setminus\{k\}} u_{j'}^r, g_1^s) \\ =& e(g_1, g_1)^{\alpha s} \cdot e(g_1, g_1)^{ars} \cdot e(\textstyle\prod_{j' \in S} u_{j'}, g_1)^{rs} \\ =& B_1. \end{aligned} \tag{1}$$

$$\begin{aligned} e(D_2, C_3) =& e(g_1^r, (\textstyle\prod_{j' \in S} u_{j'})^s \cdot R_3) \\ =& e(\textstyle\prod_{j' \in S} u_{j'}, g_1)^{rs} \\ =& B_2. \end{aligned} \tag{2}$$

$$\begin{aligned} & \frac{e(D_2, C_1)}{e(D_4, C_2)} \\ =& \frac{e(g_1^r, (\prod_{v_{i,j_i} \in \mathbb{A}} g_1^{a_{i,j_i}} \cdot R_{i,j})^s \cdot R_1)}{e\left(\left(g_1^{a+\sum_{v_{i,j_i} \in L} a_{i,j_i}}\right)^r, g_1^s R_0^s \cdot R_2\right)} \\ =& \frac{e(g_1, \prod_{v_{i,j_i} \in \mathbb{A}} g_1^{a_{i,j_i}})^{rs}}{e(g_1, g_1)^{ars} \cdot e\left(g_1^{\sum_{v_{i,j_i} \in L} a_{i,j_i}}, g_1\right)^{\alpha s}} \\ =& B_3. \end{aligned} \tag{3}$$

Then from the above three Equations (1), (2) and (3), it will be easy to obtain that

$$\frac{B_1 B_3}{B_2} = e(g_1, g_1)^{\alpha s} = K.$$

Note that in the *KeyGen* algorithm, this paper assumes $\forall L, L'(L \neq L')$, $\sum_{v_{i,j_i} \in L} a_{i,j_i} \neq \sum_{v_{i,j_i} \in L'} a_{i,j_i}$ because the parameter r has no effect on decryption. If the above condition is not met, various users associated with attribute set L, L' will have the same decryption ability [6].

4.3 Recipient Anonymous

This section will show that the proposed scheme achieve recipient anonymity in the composite order bilinear groups.

Compared with [16, 20], our scheme adds a random number to each part of the ciphertexts, these random numbers will not affect the decryption process. However, they are necessary for recipient anonymity of scheme, because if there is no such a random number, then for some access structure $\mathbb{A}^*$ and broadcast list S^* the adversary may perform the DDH-test to determine whether the ciphertext is encrypted under the $\mathbb{A}^*, S^*$ or not. In our scheme, by utilizing the DDH-test $e(C_2, \prod_{v_{i,j_i}\in\mathbb{A}^*} A_{i,j_i}) \stackrel{?}{=} e(A_0, C_1)$ to determine whether the ciphertext is encrypted under the $\mathbb{A}^*$ or not will be fail. The DDH-test $e(C_2, \prod_{v_{i,j_i}\in\mathbb{A}^*} A_{i,j_i}) \stackrel{?}{=} e(A_0, C_1)$ is same as $e(C_2, \prod_{v_{i,j_i}\in\mathbb{A}^*} A_{i,j_i})/e(A_0, C_1) \stackrel{?}{=} 1_T$, where 1_T is the identity element in $\mathbb{G}_T$, on the public parameters of attributes occur in $\mathbb{A}^*$ and the ciphertext components. The following is the detailed analysis.

$$
\begin{aligned}
& e(C_2, \textstyle\prod_{v_{i,j_i}\in\mathbb{A}^*} A_{i,j_i}) \qquad (4)\\
=& e(g_1^s R_0^s R_2, \textstyle\prod_{v_{i,j_i}\in\mathbb{A}^*} g_1^{a_{i,j_i}} R_{i,j_i})\\
=& e(g_1^s, \textstyle\prod_{v_{i,j_i}\in\mathbb{A}^*} g_1^{a_{i,j_i}}) \cdot e(R_0^s, R_{\mathbb{A}^*}) \cdot e(R_2, R_{\mathbb{A}^*}),
\end{aligned}
$$

$$
\begin{aligned}
& e(A_0, C_1) \qquad (5)\\
=& e(g_1 R_0, \textstyle\prod_{v_{i,j_i}\in\mathbb{A}} g_1^{s a_{i,j_i}} \cdot R_{\mathbb{A}}^s \cdot R_1)\\
=& e(g_1, \textstyle\prod_{v_{i,j_i}\in\mathbb{A}} g_1^{s a_{i,j_i}}) \cdot e(R_0, R_{\mathbb{A}}^s) \cdot e(R_0, R_1),
\end{aligned}
$$

where $R_{\mathbb{A}^*} = \prod_{v_{i,j_i}\in\mathbb{A}^*} R_{i,j_i}$, $R_{\mathbb{A}} = \prod_{v_{i,j_i}\in\mathbb{A}} R_{i,j_i}$.

If $\mathbb{A} = \mathbb{A}^*$, then $j_i' = j_i$ for all i, $1 \le i \le n$, and hence $\sum_{i=1}^n a_{i,j_i'} = \sum_{i=1}^n a_{i,j_i}$ and $R_{\mathbb{A}} = R_{\mathbb{A}^*}$. Therefore,

$$
\frac{e(C_2, \prod_{v_{i,j_i}\in\mathbb{A}^*} A_{i,j_i})}{e(A_0, C_1)} = \frac{e(R_2, R_{\mathbb{A}^*})}{e(R_0, R_1)}.
$$

If $\mathbb{A} \neq \mathbb{A}^*$, there exists at last one k, $1 \le k \le n$ such that $j_k' \neq j_k$. Without loss of generality, let $j_i' = j_i$, for all i, $1 \le i \le n$ except $i = k$. Then $a_{i,j_i'} = a_{i,j_i}$, $R_{i,j_i'} = R_{i,j_i}$, for all i, $1 \le i \le n$, except $i = k$. Therefore,

$$
\begin{aligned}
& \frac{e(C_2, \prod_{v_{i,j_i}\in\mathbb{A}^*} A_{i,j_i})}{e(A_0, C_1)}\\
=& \frac{e(g_1^s, g_1^{a_{k,j_k'}}) \cdot e(R_0^s, R_{k,j_k'}) \cdot e(R_2, R_{\mathbb{A}^*})}{e(g_1, g_1^{s a_{k,j_k'}}) \cdot e(R_0, R_{\mathbb{A}}^s) \cdot e(R_0, R_1)}.
\end{aligned}
$$

In both the cases, $\mathbb{A} = \mathbb{A}^*$ and $\mathbb{A} \neq \mathbb{A}^*$, the DDH-test gives a random element of $\mathbb{G}_T$ so that the adversary will be not able to determine whether the ciphertext is encrypted under the $\mathbb{A}^*$ or not. By the same way, the user index DDH-test $e(C_2, \prod_{j'\in S^*} u_{j'}) \stackrel{?}{=} e(A_0, C_3)$ will be fail, too. So both the access structure and the broadcast set are hidden, which means the proposed scheme is recipient anonymous.

5 Proof of Security

This section will show that the proposed scheme achieves the full security by employing the dual system encryption technique. In dual system encryption schemes [14, 22], ciphertexts and keys can take on two forms: normal or semi-functional. Semi-functional ciphertexts and semi-functional keys are only used in security proof, but not used in the real system. Let g_2 be a generator of the subgroup $\mathbb{G}_{p_2}$. The semi-functional ciphertexts and the semi-functional keys are created as follows.

Semi-functional ciphertexts: For an access structure $\mathbb{A} = \wedge_{i=1}^n w_{i,j_i}$, where $w_{i,j_i} \in Att_i$, and a broadcast set $S = \{1, 2, \cdots, q\} \in \mathcal{U}$, we first run the encryption algorithm *Encrypt* to obtain normal ciphertexts K', C_1', C_2', C_3'. Then choose some random elements δ, $b_{j'}$ and z_{i,j_i} in $\mathbb{Z}_N$ where $j' = \{1, 2, \cdots, q\}$, $i = \{1, 2, \cdots, n\}$. Semi-functional ciphertexts are computed as follows:

$$
\begin{aligned}
K &= K', & C_1 &= C_1' g_2^{\delta \sum_{i=1}^n z_{i,j_i}},\\
C_2 &= C_2' g_2^{\delta}, & C_3 &= C_3' g_2^{\delta \sum_{j'=1}^q b_{j'}}.
\end{aligned}
$$

Semi-functional keys: There are two types of semi-functional keys in our proof. Firstly, run the key generation algorithm *KeyGen* to get normal private key for index t and attribute set L as: $D_1, D_2, \{D_{3,j'}\}_{j'\in\mathcal{U}\setminus\{t\}}$ and D_4. Then choose random values γ, σ, σ' and $\delta_{j'}$ in $\mathbb{Z}_N$ where $j' = 1, 2, \cdots, h$ and compute two types of semi-functional private keys components as follows.

Type 1.

$$
\begin{aligned}
D_1 &= D_1' g_2^{\gamma}, & D_2 &= D_2' g_2^{\sigma},\\
\{D_{3,j'} &= D_{3,j'}' g_2^{\sigma\delta_{j'}}\}_{j'\in\mathcal{U}\setminus\{t\}}, & D_4 &= D_4' g_2^{\sigma' + \sigma \sum_{v_{i,j_i}\in L} z_{i,j_i}}.
\end{aligned}
$$

Type 2.

$$
\begin{aligned}
D_1 &= D_1' g_2^{\gamma}, & D_2 &= D_2',\\
\{D_{3,j'} &= D_{3,j'}'\}_{j'\in\mathcal{U}\setminus\{t\}}, & D_4 &= D_4'.
\end{aligned}
$$

When the semi-functional ciphertexts are used to decrypt semi-functional keys, the regular decryption will be prevented by a blind factor.

The security of the proposed scheme will be proved by using a hybrid argument over a sequence of games. Let q denote the number of secret key queries made by the adversary. The games are defined as follows.

$Game_{Real}$: It is a real CP-ABBE security game in which both private keys and challenge ciphertexts are in normal form.

$Game_0$: In this game, the challenge ciphertexts are semi-functional, but all private keys are normal.

$Game_{k,1}$: The challenge ciphertexts are semi-functional, the fist $k-1$ keys are type 2 semi-functional private keys and the k^{th} key is semi-functional of type 1. The rest of keys are replied in normal form.

$Game_{k,2}$: This game is like $Game_{k,1}$ expect for the k^{th} key is a semi-functional of type 2.

$Game_{q,2}$: In this game, the challenge ciphertexts are semi-functional, and all the private keys are in semi-functional of type 2.

$Game_{Final}$: This final game $Game_{Final}$ is the same as $Game_{q,2}$, except that the challenge ciphertext is semi-functional encryption of random message, other than neither of the two chosen massages by adversary, so the advantage of adversary in this game is 0.

We will prove that these games are indistinguishable in a set of Lemmas. Let $Game_* Adv_{\mathcal{A}}$ denote the advantage of adversary $\mathcal{A}$ in $Game_*$. Note that we have $Game_{Real} Adv_{\mathcal{A}} = Adv_{\mathcal{A}}(\lambda)$ for some fixed security parameter λ.

Lemma 1. *Suppose there exists a polynomial time algorithm $\mathcal{A}$ such that $Game_{Real} Adv_{\mathcal{A}} - Game_0 Adv_{\mathcal{A}} = \epsilon$. Then we can build a polynomial time algorithm $\mathcal{B}$ with advantage ϵ in breaking Assumption 1.*

Proof. We establish an algorithm $\mathcal{B}$ which has received $\langle \Theta, g_1, T \rangle$, where T is either an element of $\mathbb{G}$ or an element of $\mathbb{G}_{p_1 p_3}$ from the challenger. Note that a random element $d \in \mathbb{G}_{p_i}$ can be selected by choosing a random $\tau \in \mathbb{Z}_N$ and setting $d = g_i^{\tau}$, where g_i is the generator of $\mathbb{G}_{p_i}$ for $i \in \{1, 2, 3\}$.

Setup. The algorithm $\mathcal{B}$ randomly selects $R_0, R_{i,j} \in_R \mathbb{G}_{p_3}$, $\alpha, a, a_{i,j}$ and $a_{j'} \in_R \mathbb{Z}_N$. Then $\mathcal{B}$ computes $Y = e(g_1, g_1)^{\alpha}$, $A_0 = g_1 \cdot R_0$, $A_{i,j} = g_1^{a_{i,j}} \cdot R_{i,j}$, for all $1 \le i \le n$ and $1 \le j \le m_i$. The algorithm $\mathcal{B}$ produces the public parameters $PK = \langle Y, A_0, \{u_{j'} = g_1^{a_{j'}}\}_{j' \in \mathcal{U}}, \{A_{i,j} | 1 \le i \le n, 1 \le j \le m_i\} \rangle$.

Key Query Phase 1 and Phase 2. Consider the adversary $\mathcal{A}$ requires the private key for any attribute set L and index $t \in \mathcal{U}$. $\mathcal{B}$ can answer it in normal form readily because it knows the master key MSK.

Challenge. The algorithm $\mathcal{A}$ outputs an access structure $\mathbb{A}^*$ and a broadcast index set S^* to challenger $\mathcal{B}$. In order to compute the challenge ciphertexts, $\mathcal{B}$ randomly chooses $t_1, t_2, t_3 \in_R \mathbb{Z}_N$ then flips a coin $\mu \in \{0, 1\}$ and computes

$$K_0 = e(g_1, T)^{\alpha}, \qquad C_1 = T^{\sum_{v_{i,j_i} \in \mathbb{A}^*} a_{i,j_i}} g_3^{t_1},$$
$$C_2 = T g_3^{t_2}, \qquad C_3 = T^{\sum_{j \in S^*} a_{j'}} g_3^{t_3}.$$

Then $\mathcal{B}$ chooses a random symmetric key K_1 in the key space $\mathcal{K}$ and sends $\langle Hdr^*, K_{\mu} \rangle$ to $\mathcal{A}$, where $Hdr^* = (C_1, C_2, C_3)$. Here we note that if there exists $L \not\models \mathbb{A}^*$ and $t \notin S^*$ such that $\sum_{v_{i,j_i} \in \mathbb{A}} = \sum_{v_{i,j_i} \in \mathbb{A}^*}$ and $\sum_{j' \in S} a_{j'} = a_t \sum_{j' \in S^* \setminus \{t\}} a_{j'}$ hold, then the algorithm $\mathcal{B}$ aborts.

Guess. Finally, the adversary $\mathcal{A}$ outputs the guess bit $\mu' \in \{0, 1\}$ and wins the game if $\mu' = \mu$.

Note that for $i \neq j$, the values ρ modulo p_i are uncorrelated from the values ρ modulo p_j by the Chinese Remainder Theorem. If $T \in \mathbb{G}$, then it can be written as $T = g_1^s g_2^{\delta} X_3$, where g_1^s and g_2^{δ} is the $\mathbb{G}_{p_1}$ and $\mathbb{G}_{p_2}$ part of T respectively and X_3 is a random element in $\mathbb{G}_{p_3}$. This implicitly sets $z_{i,j_i} = a_{i,j_i}$ and $b_{j'} = a_{j'}$. Hence Hdr^* is a properly distributed semi-functional ciphertext, in this case, $\mathcal{B}$ simulates the game $Game_0$. If $T \in \mathbb{G}_{p_1 p_3}$, Hdr^* is a properly distributed normal ciphertext and hence $\mathcal{B}$ will simulate the game $Game_{Real}$. Therefore, if $\mathcal{A}$ can distinguish these two games then $\mathcal{B}$ will distinguish the two distributions so as to break the Assumption 1. □

Lemma 2. *Suppose there exists a polynomial time algorithm $\mathcal{A}$ such that $Game_{k-1,2} Adv_{\mathcal{A}} - Game_{k,1} Adv_{\mathcal{A}} = \epsilon$. Then we can build a polynomial time algorithm $\mathcal{B}$ with advantage ϵ in breaking Assumption 2.*

Proof. We establish an algorithm $\mathcal{B}$ which has received $\langle \Theta, g_1, X_1 X_2 X_3, Y_2, T \rangle$, where X_i and Y_i are random elements in $\mathbb{G}_{P_i}$ and T is either an element of $\mathbb{G}_{p_1 p_2}$ or an element of $\mathbb{G}_{p_1}$ from the challenger.

Setup. The algorithm $\mathcal{B}$ chooses random elements $R_0, R_{i,j}$ in $\mathbb{G}_{p_3}$, and $\alpha, a, a_{i,j}$ and a_j in $\mathbb{Z}_N$, then it computes $Y = e(g_1, g_1)^{\alpha}$, $A_0 = g_1 \cdot R_0$, $A_{i,j} = g_1^{a_{i,j}} \cdot R_{i,j}$, for all $1 \le i \le n$ and $1 \le j \le m_i$. The algorithm $\mathcal{B}$ products the public parameters $PK = \langle Y, A_0, \{u_{j'} = g_1^{a_{j'}}\}_{j' \in \mathcal{U}}, \{A_{i,j} | 1 \le i \le n, 1 \le j \le m_i\} \rangle$, and keeps the master key MSK.

Key Query Phase 1 and Phase 2. To compute the first $k-1$ private semi-functional keys, the algorithm $\mathcal{B}$ chooses random elements ϑ, r in $\mathbb{Z}_N$ and implicitly sets $Y_2 = g_2^l$ and responds to each private key request on a set of attributes L and broadcast index $t (t < k)$ from $\mathcal{A}$ by setting

$$D_1 = g_1^{\alpha + ar} u_t^r \cdot Y_2^{\vartheta}, \quad D_2 = g_1^r,$$
$$D_{3,j'} = u_{j'}^r, \qquad D_4 = \left(g_1^{a + \sum_{v_{i,j_i} \in L} a_{i,j_i}} \right)^r.$$

This implicitly sets $z_{i,j_i} = a_{i,j_i}$, $b_{j'} = a_{j'}$ and $Y_2^{\vartheta} = g_1^{\delta}$, so D_1, D_2, D_3 and D_4 are properly distributed semi-functional private key components.

To compute the k^{th} private key, the algorithm $\mathcal{B}$ will implicitly set g_1^r as the $\mathbb{G}_{p_1}$ part of T and sets

$$D_1 = g_1^{\alpha} \cdot T^{a + a_t}, \qquad D_2 = T,$$
$$D_{3,j'} = T^{a_{j'}}, \qquad D_4 = T^{a + \sum_{v_{i,j_i} \in L} a_{i,j_i}}.$$

Suppose $T \in \mathbb{G}_{p_1 p_2}$. Let $T = g_1^r g_2^{\sigma}$ for some $r, \sigma \in \mathbb{Z}_N$. Here we implicitly set $\sigma = (a + a_t), b_{j'} = a_{j'}, \sigma' = a\sigma$, and $z_{i,j_i} = a_{i,j_i}$. So the key is a semi-functional key of type 1. Similarly if $T \in \mathbb{G}_{p_1}$, the private key is normal.

Challenge. The adversary $\mathcal{A}$ submits an access structure $\mathbb{A}^*$ and a broadcast index set S^*. The algorithm $\mathcal{B}$

flips a coin $\mu \in \{0,1\}$ and sets $X_1X_2 = g_1^s g_2^\sigma$ implicitly. Then it prepares challenge ciphertexts as:

$$K_0 = e(X_1X_2X_3, g_1)^\alpha, \quad C_1 = (X_1X_2X_3)^{\sum_{v_{i,j_i} \in \mathbb{A}^*} a_{i,j_i}} g_3^{t_1},$$
$$C_2 = (X_1X_2X_3) \cdot g_3^{t_2}, \quad C_3 = (X_1X_2X_3)^{\sum_{j \in S^*} a_{j'}} g_3^{t_3}.$$

Then algorithm $\mathcal{B}$ chooses a random symmetric key K_1 in the key space $\mathcal{K}$ and sends $\langle Hdr^*, K_\mu \rangle$ to $\mathcal{A}$, where $Hdr^* = (C_1, C_2, C_3)$.

Guess. Finally, the adversary $\mathcal{A}$ outputs the guess bit $\mu' \in \{0,1\}$ and wins the game if $\mu' = \mu$.

If $T \in \mathbb{G}_{p_1}$, then $\mathcal{B}$ has properly simulated $Game_{k-1}$. If $T \in \mathbb{G}_{p_1p_2}$, then $\mathcal{B}$ has properly simulated $Game_k$. Hence, the algorithm $\mathcal{B}$ can use the output of $\mathcal{A}$ to distinguish $Game_{k-1}$ and $Game_k$. □

Lemma 3. *Suppose there exists a polynomial time algorithm $\mathcal{A}$ such that $Game_{k,1}Adv\mathcal{A} - Game_{k,2}Adv\mathcal{A} = \epsilon$. Then we can build a polynomial time algorithm $\mathcal{B}$ with advantage ϵ in breaking Assumption 2.*

Proof. This proof is very similar to the proof of the previous lemma. After receiving the challenge parameters, the algorithm $\mathcal{B}$ forms the public parameters $PK = \langle Y, A_0, \{u_{j'} = g_1^{a_{j'}}\}_{j' \in \mathcal{U}}, \{A_{i,j} | 1 \le i \le n, 1 \le j \le m_i\}\rangle$. The adversary $\mathcal{A}$ forms first $k-1$ private keys and challenge ciphertext as the previous lemma and forms last $q-k$ keys by employing master secret key respectively. For k^{th} key, the algorithm $\mathcal{B}$ chooses a random value ϕ in $\mathbb{Z}_N$ and computes:

$$D_1 = g_1^\alpha \cdot T^{a+a_t} Y_2^\phi, \quad D_2 = T,$$
$$D_{3,j'} = T^{a_{j'}}, \quad D_4 = T^{a+\sum_{v_{i,j_i} \in L} a_{i,j_i}}.$$

Let g_1^r be the $\mathbb{G}_{p_1}$ part of T. It is easy to see that if $T \in \mathbb{G}_{p_1}$, this is a well-formed type 2 semi-functional key and $\mathcal{B}$ has properly simulated $Game_{k,2}$. Otherwise, $T \in \mathbb{G}_{p_1p_2}$, this is type 1 semi-functional key and $\mathcal{B}$ has properly simulated $Game_{k,1}$. In the both cases the decryption test will be fail because the random element Y_2^ϑ cannot be cancelled out. Hence the algorithm $\mathcal{B}$ can use $\mathcal{A}$'s output to break Assumption 2 with advantage ϵ. □

Lemma 4. *Suppose there exists a polynomial time algorithm $\mathcal{A}$ such that $Game_{q,2}Adv\mathcal{A} - Game_{Final}Adv\mathcal{A} = \epsilon$. Then we can build a polynomial time algorithm $\mathcal{B}$ with advantage ϵ in breaking Assumption 3.*

Proof. We establish an algorithm $\mathcal{B}$ which has received $\langle \Theta, g_1, g_1^\alpha X_2, Y_2Y_3, Z_2, T\rangle$ and the algorithm needs to decide $T = e(g_1, g_1)^{\alpha s}$ or T is a random element of $\mathbb{G}_T$.

Setup. The algorithm $\mathcal{B}$ randomly selects $R_0, R_{i,j} \in_R \mathbb{G}_{p_3}$, $\alpha, a, a_{i,j}$ and $a_{j'} \in_R \mathbb{Z}_N$, then it computes $Y = e(g_1^\alpha X_2, g_1)$, $A_0 = g_1 \cdot R_0$, $A_{i,j} = g_1^{a_{i,j}} \cdot R_{i,j}$, for all $1 \le i \le n$ and $1 \le j \le m_i$. The algorithm $\mathcal{B}$ forms the public parameters $PK = \langle Y, A_0, \{u_{j'} = g_1^{a_{j'}}\}_{j' \in \mathcal{U}}, \{A_{i,j} | 1 \le i \le n, 1 \le j \le m_i\}\rangle$. Here $e(g_1^\alpha X_2, g_1) = e(g_1, g_1)^\alpha$.

Key Query Phase 1 and Phase 2. For attribute set L and user index t, The algorithm $\mathcal{B}$ randomly picks $r, t \in \mathbb{Z}_N$ and sets type 2 semi-functional key as

$$D_1 = g_1^{\alpha+ar} u_t^r \cdot Z_2^t, \quad D_2 = g_1^r,$$
$$D_{3,j'} = u_{j'}^r, \quad D_4 = \left(g_1^{a+\sum_{v_{i,j_i} \in L} a_{i,j_i}}\right)^r.$$

Challenge. The adversary $\mathcal{A}$ sends $\mathcal{B}$ an access structure $\mathbb{A}^*$ and a broadcast index set S^*. Then $\mathcal{B}$ flips a coin $\mu \in \{0,1\}$ and sets challenge ciphertexts as

$$K_0 = T, \quad C_1 = (g_1Y_2Y_3)^{\sum_{v_{i,j_i} \in \mathbb{A}^*} a_{i,j_i}} g_3^{t_1},$$
$$C_2 = g_1^s Y_2Y_3 g_3^{t_2}, \quad C_3 = (g_1^s Y_2Y_3)^{\sum_{j' \in S^*} a_{j'}} g_3^{t_3}.$$

This implicitly sets $Y_2 = g_2^\delta, z_{i,j_i} = z_{i,j_i}, b_{j'} = a_{j'}$. Then $\mathcal{B}$ chooses a random symmetric key K_1 in the key space $\mathcal{K}$ and sends $\langle Hdr^*, K_\mu \rangle$ to $\mathcal{A}$, where $Hdr^* = (C_1, C_2, C_3)$.

Guess. Finally, the adversary $\mathcal{A}$ outputs the guess bit $\mu' \in \{0,1\}$ and wins the game if $\mu' = \mu$.

If $T = e(g_1, g_1)^{\alpha s}$, then challenge ciphertext is a valid semi-functional ciphertext. If T is a random element in $\mathbb{G}_T$ challenge ciphertext is a valid semi-functional ciphertext for a random message. Hence the algorithm $\mathcal{B}$ can use $\mathcal{A}$'s output to break Assumption 3 with advantage ϵ. □

Theorem 1. *If assumptions 1,2,3 hold, then our scheme is fully CPA secure.*

Proof. If Assumption 1, 2 and 3 hold, by the sequence of games and Lemma from 1 to 4, the adversary's advantage in the real game must be negligible. Hence the adversary cannot attain a non-negligible advantage in breaking our scheme. □

6 Performance Analysis

In this section, we will present the comparisons between previous CP-ABBE schemes and our scheme with regard to security and efficiency.

Some previous CP-ABBE schemes are compared with ours in terms of public key size, private key size, ciphertext size, and decryption pairings cost in Table 2, access structure, full security, recipient anonymity, and hardness assumption in Table 3. Pairing denotes the decryption pairings cost. Hardness is hardness assumption. "m" and "h" are respectively used to denote the total number of attributes and users in the system. "n" and "k" represent the number of attributes in an access structure and an attribute list, respectively. "N" is maximum number of wildcard in an access structure; "m'" is maximum size of objective attribute set allowed to be associated with ciphertext.

Table 2: Efficiency comparison among different CP-ABBE schemes

Scheme	Public parameter size	Private key size	Ciphertext size	Pairing
[1]	$\mathcal{O}(m'+h)$	$\mathcal{O}(k)$	$\mathcal{O}(n)$	$\mathcal{O}(n)$
[20](Scheme 2)	$\mathcal{O}(m+h)$	$\mathcal{O}(N)$	$\mathcal{O}(1)$	$\mathcal{O}(1)$
[16]	$\mathcal{O}(m+h)$	$\mathcal{O}(h+k)$	$\mathcal{O}(n)$	$\mathcal{O}(n)$
[25]	$\mathcal{O}(log(h)+m)$	$\mathcal{O}(k)$	$\mathcal{O}(1)$	$\mathcal{O}(n)$
Ours	$\mathcal{O}(m+h)$	$\mathcal{O}(h)$	$\mathcal{O}(1)$	$\mathcal{O}(1)$

Table 3: Security comparison among different CP-ABBE schemes

Scheme	Access Structure	Full Security	Recipient Anonymity	Hardness
[1]	LSSS	No	No	n-BDHE, MEBDH
[20](Scheme 2)	AND+wildcard	No	No	n-BDHE
[16]	LSSS	Yes	No	Static
[25]	AND	No	No	n-BDHE
Ours	AND	Yes	Yes	Static

In Table 2, it is quite obvious to see that our scheme is efficient in that the ciphertext size and the costs of decryption pairing do not depend on the number of attributes. Furthermore, our scheme only needs four decryption pairing computations, $e(D_1 D_{3,j'}, C_2)$, $e(D_2, C_3)$, $e(D_2, C_1)$, and $e(D_4, C_2)$, respectively.

In Table 3, it is apparent to see that only the proposed scheme provide recipient anonymity. Recipient anonymity is an important property for encryption schemes. Recalling the example in Section 1, an intermediate user can determine receiver set by running the DDH-test. To make up for the loophole, the proposed scheme adds a random number to each part of the ciphertexts such that each side of the equation contains different random numbers, which can prevent both access structure DDH-test and user index DDH-test. In addition, our scheme adopts AND-gate access structure and achieves full security. The security of the scheme is reduced to the static assumptions.

7 Conclusions

In this paper, a recipient anonymous ciphertext-policy attribute-based broadcast encryption (CP-ABBE) scheme is introduced. In the proposed scheme, the adversary cannot learn any information about the access structure and the broadcast list just from public parameters and ciphertexts. In addition, the proposed scheme enjoys high efficiency and achieves full security in the standard model.

A drawback of the new scheme is that our access structure is restricted, where it only supports AND-gate on multi-valued attributes. So the future works are to construct a recipient anonymous CP-ABBE scheme with more flexible access structure that is holding up high efficiency under a stronger security model.

Acknowledgments

This study was supported by the Nature Science Foundation of China (61472307, 61402112, 61100165, 61100231), Natural Science Basic Research Plan in Shaanxi Province of China (2016JM6004).

References

[1] N. Attrapadung and H. Imai, "Conjunctive broadcast and attribute-based encryption," in *Third International Conference on Pairing-based Cryptography (Pairing'09)*, pp. 248–265, 2009.

[2] A. Balu and K. Kuppusamy, *Privacy Preserving Ciphertext Policy Attribute Based Encryption*, Springer Berlin Heidelberg, 2010.

[3] J. Bethencourt, A. Sahai, and B. Waters, "Ciphertext-policy attribute-based encryption," in *2007 IEEE Symposium on Security and Privacy (SP'07)*, pp. 321–334, 2007.

[4] P. S. Chung, C. W. Liu, and M. S. Hwang, "A study of attribute-based proxy re-encryption scheme in cloud environments," *International Journal of Network Security*, vol. 16, no. 1, pp. 1–13, 2014.

[5] B. Dan, E. J. Goh, and K. Nissim, "Evaluating 2-DNF formulas on ciphertexts," *Lecture Notes in Computer Science*, vol. 3378, pp. 325–341, Springer, 2005.

[6] K. Emura, A. Miyaji, A. Nomura, K. Omote, and M. Soshi, "A ciphertext-policy attribute-based encryption scheme with constant ciphertext length," *International Journal of Applied Cryptography*, vol. 5451, no. 1, pp. 13–23, 2009.

[7] A. Fiat and M. Naor, "Broadcast encryption," in *Advances in Cryptology (CRYPTO'93)*, pp. 480–491, 1993.

[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] M. S. Hwang, C. C. Lee, T. Y. Chang, "Broadcasting cryptosystem in computer networks using geometric properties of lines", *Journal of Information Science and Engineering*, vol. 18, no. 3, pp. 373–379, May 2002.

[10] J. Lai, R. H. Deng, and Y. Li, "Fully secure cipertext-policy hiding CP-ABE," in *International Conference on Information Security Practice and Experience*, pp. 24–39, 2011.

[11] C. C. Lee, T. Y. Chang, M. S. Hwang, "A simple broadcasting cryptosystem in computer networks using exclusive-OR", *International Journal of Computer Applications in Technology*, vol. 24, no. 3, pp. 180–183, 2005.

[12] C. C. Lee, P. S. Chung, M. S. Hwang, "A survey on attribute-based encryption schemes of access control in cloud environments," *International Journal of Network Security*, vol. 15, no. 4, pp. 231–240, July 2013.

[13] K. Lee and H. L. Dong, "Adaptively secure broadcast encryption under standard assumptions with better efficiency," *IET Information Security*, vol. 9, no. 3, pp. 149–157, 2014.

[14] A. Lewko and B. Waters, "New techniques for dual system encryption and fully secure hibe with short ciphertexts," in *Theory of Cryptography, Theory of Cryptography Conference (TCC'10)*, pp. 455–479, 2010.

[15] M. Li, X. Xu, R. Zhuang, and C. Guo, "Identity-based broadcast encryption schemes for open networks," in *Ninth International Conference on Frontier of Computer Science and Technology*, pp. 104–109, 2015.

[16] Q. Li and F. Zhang, "A fully secure attribute based broadcast encryption scheme," *International Journal of Network Security*, vol. 17, no. 3, pp. 263–271, 2015.

[17] C. W. Liu, W. F. Hsien, C. C. Yang, and M. S. Hwang, "A survey of attribute-based access control with user revocation in cloud data storage", *International Journal of Network Security*, vol. 18, no. 5, pp. 900–916, 2016.

[18] D. Lubicz and T. Sirvent, "Attribute-based broadcast encryption scheme made efficient," in *Cryptology in Africa International Conference on Progress in Cryptology*, pp. 325–342, 2008.

[19] H. Ma, T. Peng, Z. Liu, "Directly revocable and verifiable key-policy attribute-based encryption for large universe," *International Journal of Network Security*, vol. 19, no. 2, pp. 272–284, 2017.

[20] T. V. X. Phuong, G. Yang, W. Susilo, and X. Chen, *Attribute Based Broadcast Encryption with Short Ciphertext and Decryption Key*, 2015.

[21] A. Sahai and B. Waters, "Fuzzy identity-based encryption," in *International Conference on Theory and Applications of Cryptographic Techniques*, pp. 457–473, 2005.

[22] B. Waters, *Dual System Encryption: Realizing Fully Secure IBE and HIBE under Simple Assumptions*, Springer Berlin Heidelberg, 2009.

[23] L. Zhang, Q. Wu, and Y. Mu, *Anonymous Identity-Based Broadcast Encryption with Adaptive Security*, Springer International Publishing, 2013.

[24] Y. Zhao, P. Fan, H. Cai, Z. Qin and H. Xiong, "Attribute-based encryption with non-monotonic access structures supporting fine-grained attribute revocation in M-healthcare," *International Journal of Network Security*, vol. 19, no. 6, pp. 1044–1052, 2017.

[25] Z. Zhou and D. Huang, "On efficient ciphertext-policy attribute based encryption and broadcast encryption: extended abstract," in *ACM Conference on Computer and Communications Security (CCS'10)*, pp. 753–755, 2010.

[26] Z. Zhou, D. Huang, and Z. Wang, "Efficient privacy-preserving ciphertext-policy attribute based-encryption and broadcast encryption," *IEEE Transactions on Computers*, vol. 64, no. 1, pp. 126–138, 2015.

Biography

Leyou Zhang is a professor in the school of mathematics and statistics at Xidian University, Xi'an China. He received his PhD from Xidian University in 2009. From Dec. 2013 to Dec. 2014, he is a research fellow in the school of computer science and software engineering at the University of Wollongong. His current research interests include network security, computer security, and cryptography.

Hongjian Yin is a master degree student in the school of mathematics and statistics, Xidian University. His research interests focus on computer and network security.

CAES Cryptosystem: Advanced Security Tests and Results

Said Bouchkaren, Saiida Lazaar
(Corresponding author: Said Bouchkaren)
Department of Mathematics and Computer Science, ENSA of TANGIER, AbdelMalek Essaadi University
P.O. Box 1818 Principal Tangier, Tangier, Morocco
(Email: saidbouchkaren1@hotmail.com)
(Received Sept. 21, 2016; revised and accepted Jan. 15 & Feb. 19, 2017)

Abstract

A robust and secure cryptosystem is an encrypting system that resists against all practical cryptanalysis methods such as statistical attacks, differential cryptanalysis and linear cryptanalysis. To prove the resistance against these attacks, the cryptosystem designer must carry out a list of robustness tests. Considering these constraints, we present in the current paper results of robustness and security tests conducted on the *CAES* (Cellular automata Encryption System) cryptosystem published in a previous article. The presented tests focus on randomness tests and on differential cryptanalysis. As results of these tests, we concluded that the cryptosystem *CAES* gives a pseudo-random output regardless the input. Also the differential attack needs huge number of chosen plaintexts which make it impractical.

Keywords: Block Cipher; Cellular Automata; Differential Cryptanalysis; Randomness Test

1 Introduction

In the new era, the use of networks to communicate becomes a necessity, which means that there is huge amounts of data transmitted between communicating entities. These data are classified as normal, secret or top secret. To transmit secret or top secret data, a secure and trusted communication channel must be created. This secret communication channel can be established using a robust and reliable cryptosystem [12].

A robust and reliable cryptosystem is an encryption algorithm that can be used to encrypt and decrypt data, if and only if the communicating entities have the encryption keys [6, 8]. In other words the cryptosystem must resists against all feasible cryptanalysis methods such as *statistical attacks*; which exploits the statistical properties of the input to guess the output; and the *differential attack*; which is a kind of statistical attack, but in lieu of exploiting statistical properties of the input, it exploits the statistical differences in inputs to guess the differences in outputs.

As method of validating the reliability of an encryption system, designers conduct series of theoretical and experimental tests.

In the current article, we present some advanced validation tests to prove the robustness of a previously published algorithm named *CAES* [3]; *CAES* is a symmetric encryption scheme based on cellular automata theories defined in [3]. It encrypts blocs of 256 bits using 256 bits keys; In this article *Randomness tests* and *differential cryptanalysis* are applied. The results obtained in this paper show that *CAES* generates pseudo-random output regardless the input which means it resists against statistical attacks and also we proved that the differential attack is practically impossible.

We remember that the previous paper [3] proved that *CAES* have a good confusion and diffusion properties and it has a hight performance rate. Also the brute force attack against *CAES* has no effects.

The rest of this article is structured as follow: the second section gives brief description of *CAES* cryptosystem, the third section describes the differential cryptanalysis, the fourth section gives an overview of statistical tests, the fifth section describes data generation for experimental tests, the sixth section gives the obtained results and discussion, the seventh section describes the results of the differential attack and the last section is a conclusion and perspectives of the work.

2 *CAES* Cryptosystem

CAES (*Cellular Automata Encryption System*) is a symmetric encryption scheme based on cellular automata defined and published previously in [3]. This algorithm uses cellular automata for encryption, decryption and sub keys generation process. As technical specification, *CAES* processes data in blocs of 256 bits and uses a key of 256 bits and the encryption or decryption is accomplished af-

ter 12 iterations. For each iteration, a sub key is generated from the encryption key using a reversible and irreversible cellular automata. The encryption and decryption processes are given respectively in Algorithm 1 and Algorithm 2.

Algorithm 1 *CAES* Encryption algorithm

procedure ENCRYPT(M, Key) ▷ M is the plaintext message block and Key is the encryption key
  $SKeys[12] \leftarrow SubKeys(K)$; ▷ Generating 12 sub keys
  for i `from` 0 `to` 11 **do**
    $M = Shift(M)$
    $M = IMix(M)$
    $M = PMix(M)$
    $M = AddKey(M, SKeys[i])$
  end for
  `return` M ▷ M contains the encrypted message
end procedure

Algorithm 2 *CAES* Decryption algorithm

procedure DECRYPT(Mc, Key) ▷ Mc is the encrypted message block and Key is the encryption key
  $SKeys[12] \leftarrow SubKeys(K)$; ▷ Generating 12 sub keys
  for i `from` 11 `downto` 0 **do**
    $Mc = AddKey(Mc, SKeys[i])$
    $Mc = invPMix(Mc)$
    $Mc = invIMix(Mc)$
    $Mc = invShift(Mc)$
  end for
  `return` Mc ▷ Mc contains the plaintext
end procedure

We remember here that a detailed description of $IMix$, $PMix$, $Shift$ are given in [3].

3 Differential Cryptanalysis Overview

Differential cryptanalysis was not known publicly until the year 1990. The first published work was the cryptanalysis of the *FEAL* algorithm by Murphy [9]. Since this time, Biham and Shamir demonstrated the feasibility of this method against a variety of encryption and hashing algorithm [1].

Today differential cryptanalysis are widely used to break some encryption algorithms and hashing functions [2]. The idea of differential cryptanalysis is to track the behaviour of pairs of plaintext blocs evolving along each iteration of the encryption process, in lieu of tracking the evolution of single plaintext block. The differential cryptanalysis is an attack of *chosen plaintext attack* family. That means the enemy needs to have the ability to encipher plaintexts using the secret key which is unknown to him.

4 Statistical Tests Overview

Statistical tests are series of mathematical operations used to prove the randomness of data samples. To prove the robustness of an encryption or hashing algorithm, **NIST** (*National Institute of Standards and Technology*) proposes 16 main tests [10]. These tests can be decomposed to sub tests, in this case we can have 189 tests in total. Brief description of the main tests is presented in [10]:

- Monobit frequency test: The purpose of this test is to determine whether the number of '1' and '0' in a binary sequence are approximately the same as would be expected for a truly random sequence.
- Frequency Test within a Block: The purpose of this test is to determine whether the frequency of '1' in an M-bit block is approximately $\frac{M}{2}$.
- Runs Test: This test determines whether the oscillation between '0' and '1' is too fast or too slow.
- Test for the Longest Run of Ones in a Block: The purpose of this test is to determine whether the length of the longest run of '1' within the tested sequence is consistent with the length of the longest run of '1' that would be expected in a random sequence.
- Binary Matrix Rank Test: The purpose of this test is to check for linear dependence among fixed length substrings of the original sequence.
- Discrete Fourier Transform Test: The purpose of this test is to detect periodic features in a binary sequence.
- Non-overlapping Template Matching Test: The purpose of this test is to detect generators that produce too many occurrences of a given aperiodic pattern.
- Overlapping Template Matching Test: Both this test and the Non-overlapping Template Matching test use an m-bit window to search for a specific m-bit pattern. The difference between this test and the Non-overlapping Template Matching test is that when the pattern is found, the window slides only one bit before resuming the search.
- Maurer's Universal Statistical Test: The purpose of the test is to detect whether or not the sequence can be significantly compressed without loss of information.
- Linear Complexity Test: The purpose of this test is to determine whether or not the sequence is complex enough to be considered random.

Table 1: Statistical tests and sub tests

Test name	Number of P-Value	Identifiers
Monobit frequency test	1	0
Frequency Test within a Block	1	1
Runs Test	1	2
Test for the Longest Run of Ones in a Block	1	3
Binary Matrix Rank Test	1	4
Discrete Fourier Transform Test	1	5
Non-overlapping Template Matching Test	148	6-153
Overlapping Template Matching Test	1	154
Maurer's Universal Statistical Test	1	155
Linear Complexity Test	1	156
Serial Test	2	157-158
Approximate Entropy Test	1	159
Cumulative Sums Test	2	160-161
Random Excursions Test	8	162-169
Random Excursions Variant Test	18	170-187
Lempel-Ziv Compression	1	188

- Serial Test: The purpose of this test is to determine whether the number of occurrences of the 2^m m-bit overlapping patterns is approximately the same as would be expected for a random sequence.
- Approximate Entropy Test: The purpose of the test is to compare the frequency of overlapping blocks of two consecutive/adjacent lengths (m and $m+1$) against the expected result for a random sequence.
- Cumulative Sums Test: The purpose of the test is to determine whether the cumulative sum of the partial sequences occurring in the tested sequence is too large or too small relative to the expected behaviour of that cumulative sum for random sequences.
- Random Excursions Test: The purpose of this test is to determine if the number of visits to a particular state within a cycle deviates from what one would expect for a random sequence.
- Random Excursions Variant Test: The purpose of this test is to detect deviations from the expected number of visits to various states in the random walk.
- Lempel-Ziv Compression: The purpose of this test is determine if the compression of a random binary sequence always give random sequence..

Table 1 gives a summary of statistical tests with the expected P-Value for each test.

5 Experimental Data Generation

To carry out statistical tests, 6 data sets are generated according to **NIST** recommendations. These data sets are generated as described in the following sub section.

5.1 Plaintext and Key Avalanche

To examine the sensibility of $CAES$ algorithm to input parameters changes (key or plaintext), 768 binary sequences of size 1048576 bits are tested. In case of key avalanche these sequences are generated as follow: Let $K_0, K_1 \ldots, K_{12287}$ be 12288 random encryption keys of 256 and a plaintext M with all bits equal to '0'. We have exactly 3145728 blocs of 256 as output of $CAES$. Each bloc is $B_i = E(M, K_i) \oplus E(M, K_i^j)$, where E is the encryption function, K_i is the i^{th} encryption key and K_i^j is the i^{th} key with the j bit is flipped for $0 \leqslant j \leqslant 255$. In case of plaintext avalanche, data are generated in the same fashion except the word 'key' is substituted with the word 'plaintext'.

5.2 CBC Encryption Mode

In category, binary sequences of 2097152 bits are generated using the CBC encryption mode. In total, we generate 200 sequences. Each sequence is created using a random key, an initialization vector (IV) with all bits equals to '0' and plaintext message with all bits equal to '0'.

5.3 Random Plaintext/Key

In this data set, we analyse 256 binary sequences. Each sequence is a concatenation of 4096 ciphertexts. These ciphertexts are generated using 4096 random plaintexts (respectively 4096 random keys) and a random key (respectively random plaintext) using CBC mode.

5.4 Plaintext/Ciphertext Correlation

To study the correlation between plaintexts and ciphertexts, 128 binary sequences of 1048576 bits are examined. Given a random key and 4096 random plaintexts, a binary

Table 2: Summary of binary sequences and size of each one

Data set	Number of sequences	Size of sequence (bits)
Plaintext avalanche	768	1048576
Key Avalanche	768	1048576
CBC encryption mode	20	2097152
Random plaintext	256	1040384
Random key	256	1040384
Plaintext/Ciphertext correlation	128	1048576
Low density plaintext	256	8421632
Low density key	256	8421632
High density plaintext	256	8421632
High density key	256	8421632

Table 3: Maximal acceptable number of sequences that maybe rejected by a test

Data set	Number of tests	Maximal (expected) number of rejected sequences
Plaintext avalanche	25	40
Key Avalanche	25	400
CBC encryption mode	25	150
Random plaintext	25	175
Random key	25	175
Plaintext/Ciphertext correlation	25	125
Low density plaintext	25	175
Low density key	25	175
High density plaintext	25	175
High density key	25	175

sequence is formed by concatenating the sum of plaintexts blocs and the corresponding ciphertexts blocs using XOR operator. The ciphertexts are calculated using ECB mode. By keeping plaintexts unchanged and changing the random key, we obtain the rest of the data sets.

5.5 Low Density Key/Plaintext

In this category, two data sets are created which can be used either as plaintexts or as keys. Each set is formed of 256 sequences. Each sequence consists of 32897 ciphertexts blocs calculated using ECB mode. Ciphertexts are formed by a plaintext (or key) of 256 bits with all bits are '0', 256 plaintexts (or keys) one bit equal to '1' and other bits equal '0' (each plaintexts corresponds to a given position of bit '1'), and 32640 plaintexts with two bits equal to '1' and other bits equal to '0' (all possible combination).

5.6 High Density Key/Plaintext

Data of this category are generated in the same manner as of the previous category, except that data of this category is the binary negation of data of the previous category.

6 Statistical Tests: Results and Discussion

Statistical tests are the most advanced tests that must be achieved to prove the robustness of a given cryptosystem. these tests are also used to test the reliability of encryption algorithms such as AES [11], hashing functions such as SHA-3 [5] and pseudo random number generator such as Blum-Blum-Shub [7] and the algorithm described in [4]. In this section we present the results of these tests when applied to $CAES$ algorithm.

6.1 Empirical Analysis

In our experimental analysis, the significance level is fixed at $\alpha = 0.01$, that is, to say a test is successful if the rate of rejected sequences is less or equal to 1%, which is the ideal case. In practice, the interval of confidence is used. In this case, the maximal number of rejected sequence is $n(\alpha + 3\sqrt{\frac{\alpha(1-\alpha)}{n}})$ where n is the number of binary sequences and α is the significance level. Table 2 gives a summary of data sets sizes and Table 3 gives the number of carried out tests and maximal number of rejected sequences.

6.2 Results and Discussion

After running various statistical tests using data categories and sequences defined previously we got the results shown in Figures 1, 2, 3, 4, 5, 6, 7, 8, 9 and 10.

It is observed, in these figures, that the number of rejected sequences is less than the maximal (expected) number of rejected sequences, which means that the test was successful.

According to these results, it is clear that the $CAES$ algorithm generates pseudo-random outputs regardless the inputs. This result demonstrates a highly sought after property in robust cryptosystems to resists against cryptanalytic attacks. As consequence, the $CAES$ resists perfectly against statistical attacks and can be used to send safely secret data over a public network.

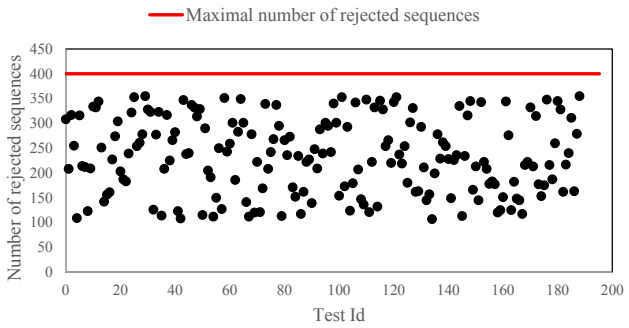


Figure 1: Statistics results using "Plaintext avalanche" data set

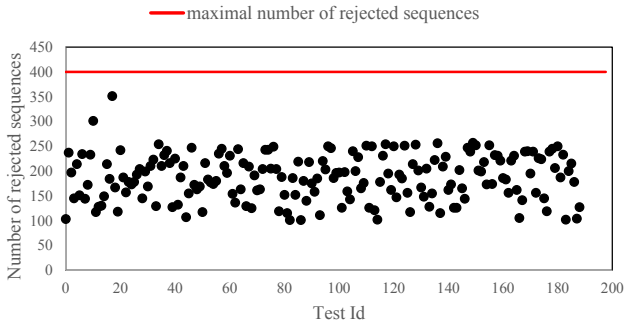


Figure 2: Statistics results using "Key avalanche" data set

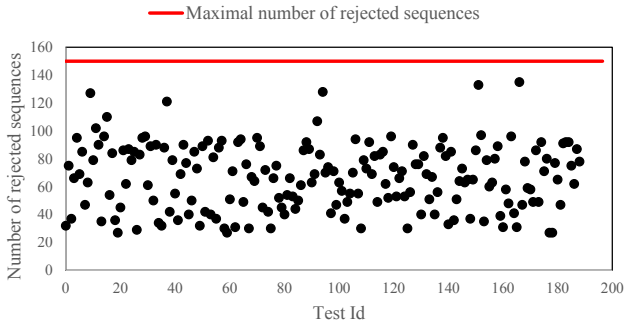


Figure 3: Statistics results using "CBC encryption mode" data set

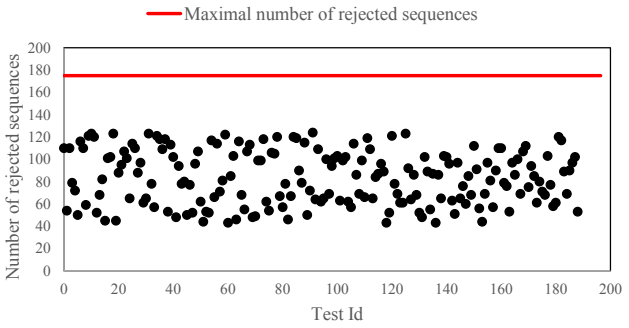


Figure 4: Statistics results using "Random plaintext" data set

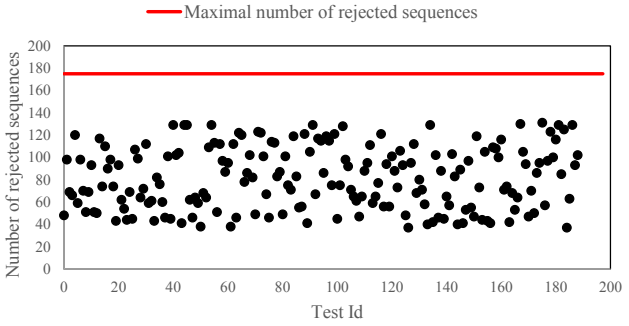


Figure 5: Statistics results using "Random key" data set

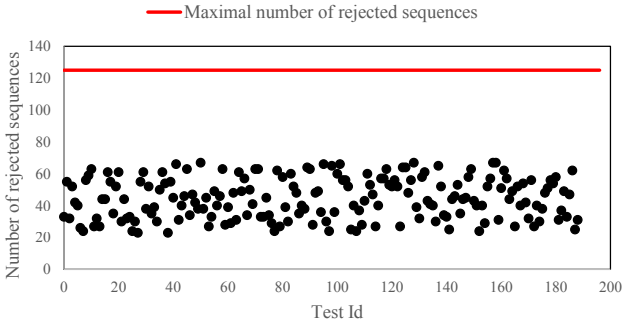


Figure 6: Statistics results using "Plaintext/Ciphertext correlation" data set

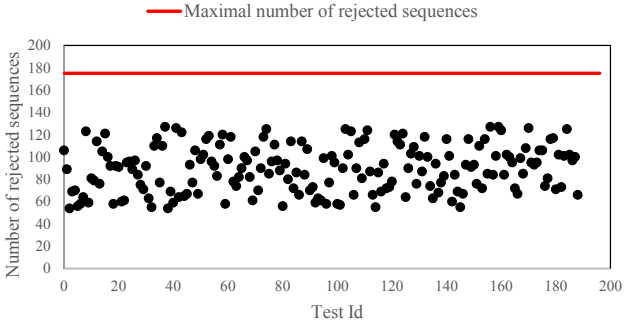


Figure 7: Statistics results using "Low density plaintext" data set

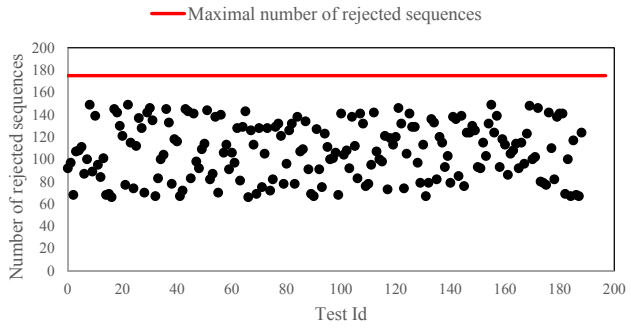


Figure 8: Statistics results using "Low density key" data set

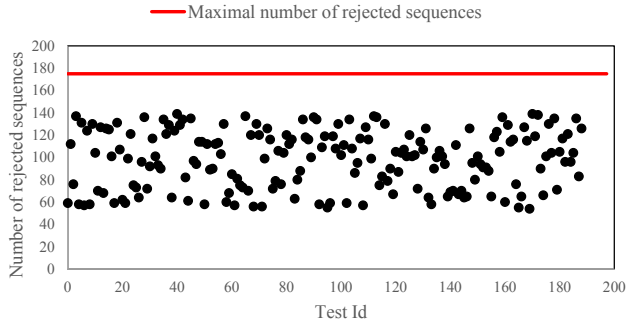


Figure 9: Statistics results using "High density plaintext" data set

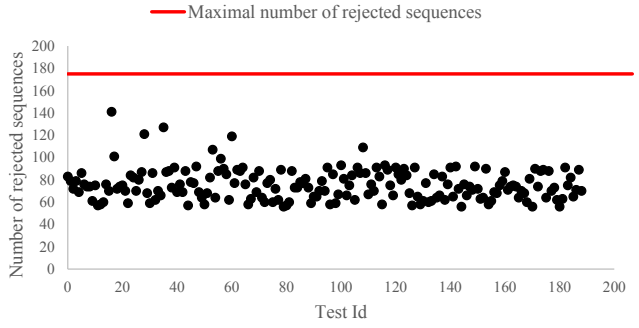


Figure 10: Statistics results using "High density key" data set

7 Differential Cryptanalysis Results

To prove the resistance of $CAES$ against differential cryptanalysis, several tests and calculations are carried out. These tests and calculations focused on non linear transformations, i.e $IMix$ and $PMix$. Our goals are to find plaintext messages m_i with a difference $X_i = m_i \oplus m_{i+1}$ producing ciphertext messages c_i with a difference $Y_i = c_i \oplus c_{i+1}$ with high probability. Table 4 gives the probability distribution of all possible differences X_i and the corresponding differences Y_i of $IMix$ and $PMix$ (they have the same distribution difference table).

Table 4: Differences distribution of $IMix$ and $PMix$ transformations

X_i \ Y_i	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	1	0	0	0	0	2	1	0	1	0	0	1	1	1	0
2	0	0	0	0	1	0	0	1	0	1	1	0	2	0	0	2
3	0	1	1	1	0	0	1	0	1	0	2	0	0	0	1	0
4	0	2	0	0	0	0	0	0	2	1	0	1	1	0	1	0
5	0	0	0	0	1	1	1	1	1	0	1	0	0	1	0	1
6	0	0	1	0	2	1	1	1	1	0	0	0	0	0	1	0
7	0	0	2	1	0	0	1	0	1	1	0	1	0	0	0	1
8	0	0	1	0	1	0	0	0	0	4	1	0	1	0	0	0
9	0	0	0	0	1	2	0	1	0	0	0	0	0	1	1	2
10	0	0	0	1	0	2	1	0	0	0	0	1	0	2	1	0
11	0	0	0	2	1	0	0	1	1	0	2	1	0	0	0	0
12	0	2	1	1	0	1	0	1	1	0	0	1	0	0	0	0
13	0	1	0	1	1	0	1	0	0	0	0	0	1	1	1	1
14	0	0	1	0	0	0	0	1	0	0	0	1	2	2	1	0
15	0	1	1	1	0	1	0	0	0	0	1	2	0	0	0	1

According to Table 4, the output difference $Y_i = 0$ is caused by the input difference $X_i = 0$ with probability $\frac{8}{16} = \frac{1}{2}$. If $X_i = m_i \oplus m_{i+1} = 0$ then $m_i = m_{i+1}$ therefore $c_i = c_{i+1}$ and as consequence no useful information about the key can be extracted using this highest value (8).

The highest exploitable value on the distribution table is 4, so an output difference Y_i of $PMix$ and $IMix$ is likely caused by an input difference X_i with probability $\frac{4}{16} \times \frac{2}{16} = \frac{1}{32}$.

To prove the robustness of $CAES$ against differential attack, we have chosen plaintexts messages m_i of difference X_i producing ciphertexts messages c_i of difference Y_i using the highest probability according to table 4 ($\frac{1}{32}$). The plaintexts messages are generated using these steps:

1) Choose the difference X_i from 4 which have the highest probability.
2) Generate a random message.
3) XOR the data from Step (1) and Step (2).

Table 5 gives an example of plaintext messages generation process.

Table 5: Example of plaintext message generated from a given difference

Difference	00800000000000000000000000000000 00000000000000000000000000000000
Random message	F622919DE18B1FDAB0CA9902B9729D49 2C807EC599D5E980B2EAC9CC53BF67D6
Resulting message	F6A2919DE18B1FDAB0CA9902B9729D49 2C807EC599D5E980B2EAC9CC53BF67D6

Suppose that the probability to have an output difference Y_i caused by an input difference X_i is exactly the probability given by the distribution table 4. In this ideal case, we need to generate and encrypt at least 2^{71} plaintext messages or $6.9 \times 10^{10} TB$ (Tera Byte) of data, which is higher than all stocked data on the internet. Therefore, we can assume that differential attack against complete version of $CAES$ cryptosystem is very difficult if not impossible.

In practice, we have been able to cryptanalyze the reduced version of CAES (one iteration version and without $Shift$ transformation) using 67 chosen plaintext messages. For a version of $CAES$ with higher number of iteration (> 2) tracking the encryption evolution at each iteration of the algorithm become very difficult. Indeed, we found that the probability to have the expected value of the output at the second iteration is $\frac{1}{128}$. As conclusion, differential attack against the full version of $CAES$ is impossible at this time.

8 Conclusion and Perspective

In the current paper we presented several tests to prove the robustness of $CAES$ encryption algorithm. The obtained results prove that the output of $CAES$ is random regardless the input, which prove that the algorithm hide all useful information about the original data. And also, we presented results of differential attack against $CAES$, the results proved that this attack have no effects against this algorithm. As perspectives, other tests and attacks; such as linear attacks and timing attacks; will be carried out in the near future.

References

[1] E. Biham and A. Shamir, *Differential Cryptanalysis of the Data Encryption Standard*, London, UK, Springer-Verlag, 1993.

[2] C. Blondeau, G. Leander, and K. Nyberg, "Differential-linear cryptanalysis revisited," in *International Workshop on Fast Software Encryption*, pp. 411–430, 2014.

[3] S. Bouchkaren and S. Lazaar, "A new iterative secret key cryptosystem based on reversible and irreversible cellular automata," *International Journal of Network Security*, vol. 18, no. 2, pp. 345–353, 2016.

[4] K. Charif, A. Drissi, and Z. Guennoun, "A pseudo random number generator based on chaotic billiards," *International Journal of Network Security*, vol. 19, no. 3, pp. 479–486, 2017.

[5] A. Doganaksoy, B. Ege, O. Koçak, and F. Sulak, "Statistical analysis of reduced round compression functions of sha-3 second round candidates.," *IACR Cryptology ePrint Archive*, vol. 2010, p. 611, 2010.

[6] T. Gulom, "The encryption algorithm GOST28147-89-PES16-2 and GOST28147-89-RFWKPES16-2," *International Journal of Electronics and Information Engineering*, vol. 6, no. 1, pp. 1–11, 2017.

[7] P. Junod, *Cryptographic Secure Pseudo-Random Bits Generation: The Blum-Blum-Shub Generator*, 1999. (`http://crypto.junod.info/bbs.pdf`)

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

[9] S. Murphy, "The cryptanalysis of FEAL-4 with 20 chosen plaintexts," *Journal of Cryptology*, vol. 2, no. 3, pp. 145–154, 1990.

[10] A. Rukhin, J. Soto, J. Nechvatal, M. Smid, and E. Barker, *A Statistical Test Suite for Random and Pseudorandom Number Generators For Cryptographic Applications*, Technical Report SP800-22, National Institute of Standards & Technology, 2001.

[11] J. Soto and L. Bassham, *Randomness Testing of the Advanced Encryption Standard Finalist Candidates*, Technical Report NIST IR 6483, National Institute of Standards and Technology, 2000.

[12] C. N. Zhang, Q. Yu, and X. W. Liu, "A hybrid fault tolerant approach for AES", *International Journal of Network Security*, vol. 15, no. 4, pp. 291–297, 2013.

Biography

Said Bouchkaren received his state engineer degree in software engineering and PhD degree in information security and cryptography from AbdelMalek Essaadi University, Morocco, in 2010 and 2016 respectively. In 2011, He joined the department of Computer sciences and mathematics as a professor. His research focuses on cryptography and information security.

Saiida Lazaar started her scientific career with a research contract funded by CNRS in France where she prepared her Ph.D. in applied mathematics. She has held positions as a researcher with IFP in France and with ONDRAF in Belgium. Currently, she is a full Professor at the University of AbdelMalek Essaadi in Morocco and Head of Master Cybersecurity and cybercriminality. She published various works, special issues in international journals, and a book on Security of networks and Cryptography

Dynamic Ant Colony System with Three Level Update Feature Selection for Intrusion Detection

Helmi Md Rais and Tahir Mehmood
(Corresponding author: Tahir Mehmood)

Department of Computer and Information Science, Universiti Teknologi Petronas
32610 Seri Iskandar, Perak, Darul Ridzuan, Malaysia.
(Email: tahirmehmood.seecs@gmail.com)
(Received Oct. 4, 2016; revised and accepted Feb. 2 & Mar. 11, 2017)

Abstract

The current era is known as the age of digital information and general medium of access to this information is computer networks. The uses of network technology also make information insecure. Intrusion Detection System (IDS) has been proven effective against such attacks. The anomaly-based detection method is good to detect new attacks. One of the foremost shortcomings in the anomaly-based detection is the irrelevant and redundant features to the classification algorithm that results in low detection rate. Therefore, the primary objective of the feature selection process is to enhance the classification accuracy by removing redundant and irrelevant features. In this research a new feature selection algorithm called, Dynamic Ant Colony System with Three Level Update Feature Selection, has been proposed. The proposed method uses a different level of pheromones that help ants to find the robust features. The method also utilizes the information of each individual ant during feature selection process and incorporates the accuracy of the classification algorithms. Results showed that proposed feature selection algorithm outperformed compared to the previous feature selection algorithms.

Keywords: Ant Colony Optimization; Feature Selection; Intrusion Detection System

1 Introduction

Network security is becoming a crucial and elementary task for the organizations. Due to which many tools are being developed to overcome the threats to the security of the network and system. Intrusion detection system is one of the measure taken for the detection of the intrusion [50]. It mainly detects for the compromising of either data confidentiality, integrity, or availability. Based on the location of the intrusion detection system implementation, it is categorized into two types [20]; network based intrusion detection system and host based intrusion detection system. Network based intrusion detection system detects intrusion in the network segment, whereas host based intrusion detection detects intrusion in the host system. Despite the division of the intrusion detection system according to their implementation, intrusion detection system is further categorized according to the implementation of the detection method [54]. There are two types of intrusion detection method namely; signature based and anomaly based. Signature based detection method uses the stored signatures of the attacks for the detection of the intrusion. Due to the utilization of the stored signature of the attacks, this method has high true positive rate. This method, however, cannot detect zero day attack as no signature exist for zero day attack. On the other hand, anomaly based detection method can detect novel attacks as it works by taking the behavior of the network into consideration. Network anomaly detection method makes a baseline for the normal activity, any activity that deviates from that baseline is considered as a possible intrusion [38]. Anomaly based detection method, however, has a high false positive rate as it is difficult to map the normal behavior of the network. Introduction of new attacks in network changes the behavior of the network while the normal data behavior remains same. Anomaly detection is therefore, depends on the behavior of both normal data and anomalous data. Learning the boundary between normal behavior and anomalous behavior of the network is therefore required. In regard to separate the normal and anomalous behavior, many techniques including supervised classification algorithms are adapted for this purpose [14, 24]. These classification algorithms, however, highly depends on the input features of the data. Redundant, irrelevant, and noisy features make it difficult for the classification algorithm to build a detection model with high accuracy rate. Feature selection approach is therefore used, which selects the features that contribute more information about the class while not compromising the accuracy of the classification algorithm.

In network intrusion detection, features are extracted

from protocols header at different layers of network architecture and contents of data packets. Due to this reason noise in channels propagate to extracted features, this leads to false intrusion alarm. These noisy features should be removed using feature selection. In this work we used Ant colony optimization (ACO) for feature selection of the network data. ACO has optimal solution, because ACO can search in the feature space up to meeting an optimal solution. ACO has the historical linkage to previous iterations. Results of next iteration is based on the amount of pheromone, which is left in previous iterations. Every ant aimed to get the best local optimal solution and among those solution global optimal solution is found. The optimal feature set was then validated using Support Vector Machine (SVM) for classification of normal and intrusive activities in network.

The paper is organized as follows, Section 2 gives the brief literature review of the feature selection method for intrusion detection. Section 3 discusses the proposed feature selection algorithm and Section 4 discusses the results. The work is concluded in Section 5. .

2 Previous Work

Lin et al. [27] Used Simulated Annealing (SA) with Support Vector Machine (SVM) and Decision Tree (DT) together for feature selection and anomaly detection. Optimal feature set was selected using simulated annealing whereas Decision Tree was used to get rules from the dataset. SVM was used for classification of the data. Simulated annealing was also used to adjust automatically the parameter setting for SVM and DT. George [17] used Principle Component Analysis (PCA) along with SVM for intrusion detection. PCA was used for dimension reduction (feature selection) while SVM used for classification. Evaluation, based on the SVM with PCA approach, gave less misclassification compared to SVM method. Ganapathy et al. [16] along with survey of the different feature selection and classification algorithm, proposed new feature selection algorithm. This feature selection algorithm combined information gain ratio of the feature and rule based approach. Liu et al. [29] has used principle component analysis for feature selection along with neural networks for classification purpose. Features with highest eigenvalues were selected in the proposed approach. The mentioned approach, however, may leave the important features for inclusion. Solely relying on the high value of eigenvalues might not be enough for feature selection [4]. Baig et al. [5] proposed two-phase technique for the classification of KDD99 dataset into normal and anomalous class. This approach used three feature ranking techniques, gain ratio, information gain, and global method for data handling (GMDH) for feature selection.

Ghali [18] used the rough set theory along with the artificial neural network. The aim was to reduce the dataset for intrusion detection which resulted in less consumption of computer resources. RSNNA (Rough Set Neural Network Algorithm) was used for feature reduction, which found dependencies among the features while feed forward neural network was used for the classification of the data. Sheikhan et al. [45] proposed a method that used fuzzy association rule for the generation of feature subsets. While fuzzy logic with ARTMAP (adaptive resonance theory neural networks) was used for the validation of the feature subset using classification. Kannan et al. [23] proposed feature selection method based on genetic algorithm. The purpose of the study was to remove unimportant features thus reducing training time of the classification or clustering. In addition to that Fuzzy based SVM was used for the validation of the feature subset. The proposed genetic algorithm was based on weighted sum, increasing global search capability resulted in better attribute collaboration. Rufai et al. [42] combined membrane computing (MC) and bee algorithm (BA) for their work. Motivated by membrane structure and operations of living cells MC, gives the solution for BA to find the best feature subset. Thus, it improved BA for feature selection. BA was run on different membranes in the main membrane to get the initial solution. Zainal et al. [53] used a 2-tier approach, which included rough set and particle swarm optimization (PSO), Rough-PSO. SVM was used for classification while fitness function was used to find out the fitness of the proposed feature subset.

3 Dynamic Ant Colony System with Three Level Update Feature Selection (DACS3-FS)

Nature inspired science for solving many hard problems that are existed for humans. Researchers, therefore, mimics their properties to solve real world problems. One of them is following the foraging behavior of ants. Real ants have the property to solve very complex problems by utilizing information of each ant. Ants use a chemical substance called, pheromone, for indirect communication with each other. Pheromone is laid by ant on the way back from food to nest and vice versa. It works as a guidance to other ants. High pheromone intensity attracts more ants. Intensity of pheromone depicts the importance of the path. Using this property ants are able to select shortest path from nest to food source.

ACO was first used to solve Traveling Salesman Problem, which is a NP-complete problem [35]. Ant system is one of the variation of ant colony optimization technique. ACO is used for many optimization problems due to its high optimal solution for optimization problems [47]. Ant colony optimization (ACO) has less complexity in terms of time and memory requirement. ACO uses heuristic information and pheromone value to compute next move. During traversing an edge ant updates the pheromone value on the edges. Pheromone update also called global pheromone update for ant system. Ant system is one of the variation of ant colony optimization technique which

includes the pheromone update level after completion of the tour by all ants. This method was improved by introducing another level of pheromone update called; local pheromone update in ant colony system (ACS) [10].

Yi and Gong [25] introduced improved version of ACS called, Dynamic Ant Colony system (DACS). Improved version of ACS avoid the growth of pheromone level too high by introducing the dynamic decay parameter (1 - $\rho[\tau(r,s)]$). The dynamic decay parameter is applied at both level of pheromone updates such as local pheromone updating rule and global pheromone updating rule. Helmi et al. [39], improved the DACS algorithm by introducing updating of pheromone at three level; local level, intermediate level, and at global level. Local pheromone is updated when all ants start their tour. Intermediate pheromone updating is done by retrieving the best knowledge of the best individual ants of the group after completing a tour and then it is divided into best of the group and worst of all groups. This is followed by the global pheromone update in which the ant that having best tour is being considered and is divided into worst of the global best and best of the global best. This method provides better searching guidance in the effort to search for better solution.

In this research, the proposed work of Helmi et al. [39], is modified to adapt for the feature selection. Block diagram of the DACS3-FS algorithm is given in Figure 1. The purpose of declaring different pheromone update levels are to take the advantage of the pheromone intensity since the small increment in the pheromone values will not guarantee that ACO will give optimal solution [33]. Also by increasing amount of pheromone too high causes to converge the solution too early.

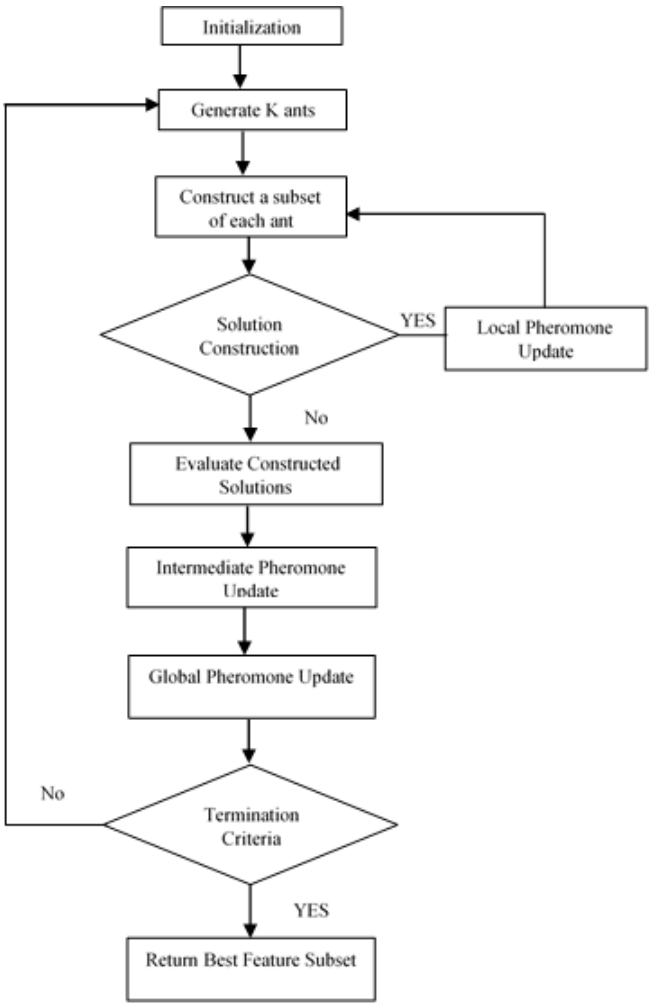


Figure 1: Flow chart of DACS3-FS

3.1 Features Representation

Suitable representation of the problem domain for ant colony optimization implementation is important. Some of the previous work used graphical representation for feature selection [40, 44], while some researchers used other methods like represented features in a binary form i.e. 1 and 0 [2, 46]. In this study, completely connected graph representation is used, Figure 2. Dark doted lines in the figure depicts, how features are selected by the best ant in feature selection process. Thus each ant have chance to select any node based on pheromone and heuristic value. Features are represented by a node and are connected with each other by edges. Pheromone and heuristic values are related with the features thus not laid on the arcs. Number of ants used were equal to total number of the features. Using too many ants would lead to quick convergence which may result to bad solution while few ants would not be able to utilize the cooperation of synergistic effect due to pheromone decay process [7].

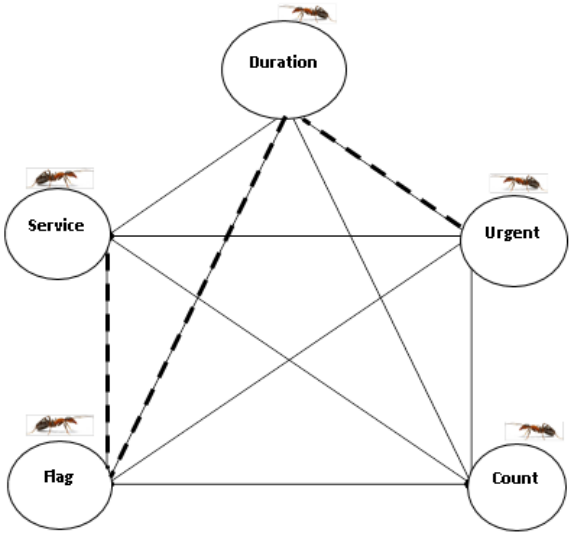


Figure 2: Graphical representation of features

3.2 Heuristic Function

Heuristic information plays a vital role to generate high quality results [37]. It describes priori desirability of the move. It has great influence on the performance of the ant colony optimization algorithms [28]. Artificial ants can lead to bad result and can reinforce it in each tour as initial random pheromones do not lead them. To avoid such deadlock heuristic information helps to recover from that deadlock [37]. Heuristic information related to the feature must be used [1]. It helps for positive constructive step thus helps to improve the performance of the ants [9]. There are two type of heuristic information [30] i.e. static heuristic and dynamic heuristic. Static heuristic is calculate and initialized at the start of the algorithm run and remains same throughout whole algorithms run like distance between cities in TSP. The advantage of the static heuristic is that it is calculated once and easy to compute [11]. Dynamic heuristic is calculated at every step as it depends on partial computed solution. Dynamic heuristic is, therefore, computationally expensive. In the proposed work we have used correlation as heuristic information. So the static heuristic is used in this work. Correlation of each feature respective to class is used and the values are constant thoughout whole algorithms run. So the ants can get some extra information for constructing solution. The importance of the heuristic is controlled by β.

3.3 Transition Probability

Like real ants artificial ants must evaluate the intensity of pheromone to do decision for next move [41]. In real ants the greater the amount of pheromone more the probability that ants will select that path. In computation problem completely relying on pheromone value for the path decision can lead to false result [36]. But in graph problem artificial ants decide the next move based on the probabilistic choice from a set of allowed nodes. This probabilistic choice depends on two parameters i.e. heuristic value and pheromone value [34]. Heuristic value is regarded as the visibility of the path in future perspective while pheromone value is regarded as common memory in past perspective of the path. Both values together controls the movement of the ant for solution constructions. In the proposed methodology we used correlation values of the features as a heuristic value and number of times features visited as initial pheromone value. Correlation values of the features to the classes will help move probability function to consider the importance of the feature to the prediction of the classes. In feature selection problem, the heuristic value should involve some kind of evaluation function for movement of feature to feature [21]. This will avoid ant to select bad feature each time by considering heuristic value since it is the priori desirability of the move. In the proposed method, initial pheromone value (τ_o)is set to 1 so no feature get biased pheromone at the start of the algorithm run. Thus allowing ant to select next feature unbiased regarding its pheromone value at start. Only feature will be selected once by a single ant while a single feature can be visited by many ants. Move probability for the proposed DACS3-FS is given as;

$$P_{ij} = max[(\tau_j)^{\alpha} * (\eta_j)^{\beta}] \quad (1)$$

Where τ_j represents pheromone value of the next feature while heuristic information of the next feature is represented by, η_j. Parameters α and β controls the trade-off between pheromone and heuristic information. The values of these parameters, therefore, influences the result of probability function. The trade-off between intensification and diversification is influenced by modifying the values of parameters [48]. Moreover, as static heuristic approach has been used in the proposed method, therefore, it will effect random probability function at the start of the construction. Selecting a single feature in each iteration, for feature subset construction, requires more computing cycles. Moreover, single feature by itself does not much help to find the class of the data, therefore, group of features are selected in this step.

3.4 Pheromone Value

Artificial pheromone is the cumulated numerical information associated with the edges that is laid by different ants during solution construction [12]. This pheromone information imitate the search experience of ants by changing the pheromone value by visiting edges each time. Pheromone information helps ant to probabilistically decide the next edge move. High intensity pheromone attracts more ants [8]. Pheromone evaporated with time that helps to diversify the search which improves the chances of other nodes to be explored by ants. Pheromone ants change this information each time they visit the edge to imitate and store their memory in the form of pheromone. This pheromone information is changed by different ants while traversing. Pheromone value, however, should be evaporated with some degree to diversify the traversing of the ants. At the end, each feature consists of the pheromone that has been laid by ants while traversing that feature. Feature which is visited by many ants will have high amount of pheromone.

3.4.1 Local Pheromone Update

Local search is the basic start of the ant colony optimization algorithm. Each arc in ant colony optimization get initialized by equally non-negative small value [51, 13]. An arc in ACO get local pheromone update with each iteration. Pheromones values get dwindle with each iteration. In ant colony system local pheromone is used to make visited edges less desirable and thus increase chance to explore the edges that are not visited yet [32, 19]. While in ant system we deposit the pheromone after decaying pheromone with some constant factor [49]. In this work local pheromone update, using Equation (2), is used while

ant traverse the feature.

$$\tau(r,s) \leftarrow (1-\rho).\tau(r,s) + \rho.\Delta\tau_o \quad (2)$$

Where

$$\Delta\tau_o = \begin{cases} \frac{1}{F_n} & if\ F_n\ \epsilon features\ visited\ by\ n\ ants \\ 0 & Otherwise \end{cases}$$

ρ is a pheromone decay parameter and $0<\rho<1$. Too much pheromone produced by the local pheromone update causes the local optimization solution in which it ignores the optimization solution [22]. To avoid local optimization solution problem the $\Delta\tau_o$ is controlled by the number of times the specific feature is used.

3.4.2 Intermediate Pheromone Update

In the proposed method we have used intermediate pheromone update which helps to reinforce the pheromone value. Available knowledge of each member of the ant′s group is used for intermediate pheromone update. Each ant′s subset was evaluated using naive bayes classifier. Recalling that feature subset selected by ant was evaluated using naive bayes classifier. Feature subset that produced high accuracy for naive bayes classifier was allowed to deposit pheromone using Equation (3). This step is important because it encourages ants to produce best feature subset for global pheromone update level.

$$\tau(r,s) \leftarrow ((1-\rho.\tau(r,s)).\tau(r,s)) + \rho.\Delta\tau_o \quad (3)$$

Where

$$\Delta\tau_o = \begin{cases} \frac{1}{accuracy} & if\tau(r,s)\epsilon Local\ Group\ Best\ Tour \\ -\frac{1}{accuracy} & if\tau(r,s)\epsilon Local\ Group\ Worst\ Tour \\ 0 & Otherwise \end{cases}$$

3.4.3 Global Pheromone Update

The accuracy from one classification algorithm is not enough. This is because one classification algorithm does not necessarily be able to find the correct relationship between the feature and class label. As pheromone update rule does not necessarily produce best solution [33]. Similarly, subset that produced good accuracy in intermediate pheromone update level is not necessary be able to produce good accuracy result for other classification algorithms. Another level of pheromone deposit, therefore, helps towards finding of best solution. Global pheromone also help to tackle the local optima problem. In the proposed ACO algorithm, accuracy of Support Vector Machine (SVM) was used for the global pheromone update. Refer to the intermediate pheromone update level, only a single ant is able to deposit the pheromone which is selected feature set produced high accuracy for SVM. Best of the best i.e. the subset which was able to produce high accuracy for both classification algorithm can deposit the global pheromone using Equation (4).

$$\tau(r,s) \leftarrow ((1-\rho.\tau(r,s)).\tau(r,s)) + \rho.\Delta\tau_o \quad (4)$$

Where

$$\Delta\tau_o = \begin{cases} \frac{1}{accuracy} & if\tau(r,s)\epsilon Global\ Group\ Best\ Tour \\ -\frac{1}{accuracy} & if\tau(r,s)\epsilon Global\ Group\ Worst\ Tour \\ 0 & Otherwise \end{cases}$$

3.4.4 Stopping Criteria

ACO runs several times until some stopping criteria met. Stopping criteria can be number of time algorithms run, number of subsets evaluation, maximum number of iterations, or for specific number of time the best solution is not changed and so on [31]. In this work maximum number of iterations were used as stopping criteria. Apart from that if any ant is unable to improve accuracy in three runs than this ant is destroy. This is necessary to avoid trap into local optima.

4 Results and Discussions

KDD99 dataset is the benchmark dataset used for the evaluation of anomaly detection methods in network intrusions [43]. Many research groups validated their detection model using KDD99 dataset [3, 6, 26, 52]. The dataset came from DARPA98 IDS evaluation program [15]. Training data is collected from seven weeks of data in which few weeks data are attack free while other weeks of data consist of attacks. Despite of it, two weeks of data resulted testing dataset which consist of attack data and normal data. Kdd99 has huge records thaś why its subset is widely used and is called kddcup.data_10_percent (kdd99_10%). 22 attacks are in the training set, 16 additional attacks are in the testing set. The training set contains 494020 instances while 311029 instances are for testing dataset. KDD99 contains four attack classes, User-to-Root (U2R), Probe, Denial of Service (DoS), Root-to-Local (R2L), and one legitimate data class called, Normal.

The experiments were carried out on system with core i7 and running windows 10 with 16GB of RAM. Matlab and weka tools were used for the experiment. Moreover, KDD99 dataset contains redundant data which were removed in this experiment. Support Vector Machine (SVM) was used for the validation of the feature subset selected using DACS3-FS. SVM was used for both binary and multiclass classification. Results of the feature subsets are empirically compared with the benchmark results.

The features that have been used in previous studies are shown in Table 1. As discussed earlier KDD99 dataset has 41 features thus we also used the result of the KDD99 full feature set for comparison purpose. This is due to the purpose of the study to increase true positive rate (TPR), when the data is correctly classified in its own class, and precision, the portion of the true positive over all the positive instances the detection method has detected as anomalous, and accuracy meanwhile minimizing false positive rate (FPR), when data of some other

class is incorrectly accepted, as low as possible. The F-measure is the harmonic mean of precision and recall. A good classifier is expected to obtain F-measure as high as possible. These features were validated for both binary and multiclass classification.

Table 1: Features selected using different feature selection methods

FS Method	Given Features	Authors
Information Gain	2,5,8,10,14,15,19, 26,27,30,31,32,33, 34,35,36,37,38,40	Ganapathy et al[16]
Rough Set	5,6,23,24,32, 33,36	Ghali [18]
Genetic Algorithm	2,3,4,5,6,10,12,23, 25,29,30,35, 36, 37,38,40	Kannan et al.[23]
Membrane Computing	2,3,8,13,20,24,32, 37,37,39,40	Rufai et al.[42]
KDD99	41	—
DACS3-FS	2,3,5,6,23, 33	—

4.1 Binary Classification

These features were validated for both binary and multiclass classification. As discussed in earlier section, KDD99 dataset has one legitimate class called, Normal, and four attack classes, DoS, R2L, PROBE, and U2R. For binary classification, these four attack classes were combined into a single class i.e. Attack class.

Table 2 gives the comparison of the different feature selection method. As shown in table, feature set selected using DACS3-FS algorithm had out performed resulted accuracy 98.7087% while full feature set resulted accuracy of 98.5172%. Moreover, Table 3 gives the detail comparison for normal class result, it can be seen that DACS3-FS based feature set had performed well for binary classification. It had though TPR 99.1% slightly less than the TPR 99.2% of rough set based feature set but FPR for DACS3-FS was less compared to other feature sets results. Result of attack class for different feature set approach is given in Table 4. Rough set based feature set had FPR and precision slightly better than DACS3-FS based feature set but the classification algorithm was unable to classify attack class data.

Table 2: Binary accuracy comparison

FeatureAlgorithms	Features	Accuracy%
IG	19	97.6348
Rough Set	7	98.0191
MC	10	95.9747
GA	17	98.3645
KDD99	41	98.5172
DACS3-FS	6	**98.7087**

Table 3: Normal class result comparison

Algorithm	TPR %	FPR %	Precision %	F-measure
Information Gain	98.9	4.3	97.2	0.981
Rough Set	99.2	3.8	97.6	0.984
Membrane Computing	99.1	8.9	94.5	0.968
Genetic Algorithm	99.1	2.7	98.3	0.987
DACS3-FS	99.1	1.9	98.8	0.989

Table 4: Attack class result comparison

Algorithm	TPR %	FPR %	Precision %	F-measure
Information Gain	95.7	1.1	98.3	0.969
Rough Set	96.2	0.8	98.7	0.974
Membrane Computing	91.1	0.9	98.5	0.947
Genetic Algorithm	97.3	0.9	98.5	0.979
DACS3-FS	98.1	0.9	98.6	0.984

4.2 Multiclass Classification

In the previous section, we used binary SVM for the binary classification of KDD99 dataset, in which we combined all attack classes into a single attack class. In this section, result for multiclass classification is given. SVM was used for multiclass, although it is a binary classifier but it can be used for multiclass using cascading different binary SVM.

For normal class result is given in the Table 5. DACS3-FS based feature set resulted low FPR and high precision compared to other feature sets. For DoS attack class different feature sets result is given in Table 6. It can be seen the genetic algorithm based feature set had better result compared to DACS3-FS based feature set and information gain based feature set. TPR for these three feature set were same.

DACS3-FS based feature set performed well for R2L class as given in Table 7. It had high TPR 80% and FPR 0%. Rough set based feature set had high precision and low FPR but had low TPR 60.6% for Probe class as given in Table 8. U2R class result using different feature sets is given in Table 9. It can be rough set and DACS3-FS and rough set based feature sets had same TPR 52.2% and FPR 0.0% but DACS3-FS based feature set resulted high precision. Accuracy for different feature selection algorithm is given in Table 10, for different classes. It can be seen that DACS3-FS resulted high accuracy of 98.7359% while full feature set resulted accuracy of 98.6013%.

Table 5: Normal class result comparison

Algorithm	TPR %	FPR %	Precision %	F-measure
Information Gain	99.5	5.1	96.8	0.981
Rough Set	99.4	3.3	97.9	0.987
Membrane Computing	99.4	8.5	94.7	0.970
Genetic Algorithm	99.7	3.2	98.0	0.988
DACS3-FS	99.5	1.5	99.1	0.993

Table 6: DoS class result comparison

Algorithm	TPR%	FPR%	Precision%	F-measure
Information Gain	99.4	0.6	99.0	0.992
Rough Set	99.6	1.0	98.2	0.989
Membrane Computing	96.6	2.6	95.3	0.959
Genetic Algorithm	99.6	0.4	99.2	0.994
DACS3-FS	99.6	0.7	98.7	0.992

Table 7: R2L class result comparison

Algorithm	TPR %	FPR %	Precision %	F-measure
Information Gain	17.9	0.0	98.7	0.303
Rough Set	55.7	0.1	91.7	0.693
Membrane Computing	1.5	0.1	15.0	0.028
Genetic Algorithm	68.4	0.0	95.8	0.798
DACS3-FS	80.9	0.0	97.0	0.882

Table 8: PROBE class result comparison

Algorithm	TPR %	FPR %	Precision %	F-measure
Information Gain	81.1	0.2	91.9	0.861
Rough Set	60.6	0.1	92.2	0.731
Membrane Computing	15.4	0.1	84.1	0.260
Genetic Algorithm	71.6	0.2	90.5	0.800
DACS3-FS	80.8	0.2	90.4	0.853

Table 9: U2R class result comparison

Algorithm	TPR %	FPR %	Precision %	F-measure
Information Gain	0.0	0.0	0.0	0.000
Rough Set	52.2	0.0	80.7	0.634
Membrane Computing	0.0	0.0	0.0	0.000
Genetic Algorithm	3.8	0.0	57.1	0.072
DACS3-FS	52.2	0.0	90.1	0.661

Table 10: Multiclass accuracy comparison

Feature Algorithms	Features	Accuracy%
IG	19	97.4769
Rough Set	7	97.834
MC	10	94.7481
GA	17	98.2571
DACS3-FS	6	**98.7359**

5 Conclusion

The performance of the classification algorithms is highly depends on input features of the data. Poor selection of features can affect classification accuracy badly which leads toward high rates of false negatives and false positives. This problem can be handled effectively by using optimized selected features. Many feature selection methods are unable to identify the complex relationship between the features which in result unable to produce the useful features. Some ranking feature selection methods tried to find all the features relevant to class attribute but failed to identify the redundant features. In this study novel feature selection algorithm called, Dynamic Ant Colony System with Three Level Update Feature Selection, a variant of ant colony optimization was proposed. The proposed algorithm is a wrapper based feature selection approach using two machine learning algorithm for the evaluation of the feature set during feature selection process. The proposed feature selection algorithm resulted in an optimal feature set that produced efficient detection model in terms of accuracy compared to the previous feature selection algorithms.

References

[1] N. Abd-Alsabour, H. Hefny, and A. Moneim, "Heuristic information for ant colony optimization for the feature selection problem," in *Conference Anthology*, pp. 1–5, IEEE, 2013.

[2] N. Abd-Alsabour and M. Randall, "Feature selection for classification using an ant colony system," in *Sixth IEEE International Conference on e-Science Workshops*, pp. 86–91, 2010.

[3] A. Abraham, C. Grosan, and C. Martin-Vide, "Evolutionary design of intrusion detection programs.," *International Journal of Network Security*, vol. 4, no. 3, pp. 328–339, 2007.

[4] I. Ahmad, M. Hussain, A. Alghamdi, and A. Alelaiwi, "Enhancing SVM performance in intrusion detection using optimal feature subset selection based on genetic principal components," *Neural Computing and Applications*, vol. 24, no. 7-8, pp. 1671–1682, 2014.

[5] Z. A. Baig, S. M. Sait, and A. Shaheen, "Gmdh-based networks for intelligent intrusion detection," *Engineering Applications of Artificial Intelligence*, vol. 26, no. 7, pp. 1731–1740, 2013.

[6] M. H. Bhuyan, D. K. Bhattacharyya, and J. K. Kalita, "Towards generating real-life datasets for network intrusion detection.," *International Journal of Network Security*, vol. 17, no. 6, pp. 683–701, 2015.

[7] E. Bonabeau, M. Dorigo, and G. Theraulaz, *Swarm intelligence: from natural to artificial systems*, No. 1, Oxford university press, 1999.

[8] M. Ciba and I. Sekaj, "Ant colony optimization with re-initialization," *Automation, Control and Intelligent Systems*, vol. 1, no. 3, pp. 59–63, 2013.

[9] C. R. Conti, M. Roisenberg, and G. S. Neto, "ACO - v-an algorithm that incorporates the visibility heuristic to the aco in continuous domain," in *IEEE Congress on Evolutionary Computation*, pp. 1–8, 2012.

[10] M. Dorigo, M. Birattari, and T. Stutzle, "Ant colony optimization," *IEEE Computational Intelligence Magazine*, vol. 1, no. 4, pp. 28–39, 2006.

[11] M. Dorigo and T. Stützle, "The ant colony optimization metaheuristic: Algorithms, applications, and advances," in *Handbook of Metaheuristics*, pp. 250–285, Springer, 2003.

[12] M. Dorigo and T. Stützle, "Ant colony optimization: overview and recent advances," in *Handbook of Metaheuristics*, pp. 227–263, Springer, 2010.

[13] D. D. Duc, H. Q. Dinh, and H. H. Xuan, "On the pheromone update rules of ant colony optimization approaches for the job shop scheduling problem," in *Pacific Rim International Conference on Multi-Agents*, pp. 153–160, Springer, 2008.

[14] M. A. Eid, H. Artail, A. I. Kayssi, and A. Chehab, "Lamaids: A lightweight adaptive mobile agent-based intrusion detection system.," *International Journal of Network Security*, vol. 6, no. 2, pp. 145–157, 2008.

[15] M. A. Faisal, Z. Aung, J. R. Williams, and A. Sanchez, "Data-stream-based intrusion detection system for advanced metering infrastructure in smart grid: A feasibility study," *IEEE Systems Journal*, vol. 9, no. 1, pp. 31–44, 2015.

[16] S. Ganapathy, K. Kulothungan, S. Muthurajkumar, M. Vijayalakshmi, P. Yogesh, and A. Kannan, "Intelligent feature selection and classification techniques for intrusion detection in networks: A survey," *EURASIP Journal on Wireless Communications and Networking*, vol. 2013, no. 1, p. 1, 2013.

[17] A. George, "Anomaly detection based on machine learning: dimensionality reduction using pca and classification using SVM," *International Journal of Computer Applications*, vol. 47, no. 21, 2012.

[18] N. I. Ghali, "Feature selection for effective anomaly-based intrusion detection," *International Journal of Computer Science and Network Security*, vol. 9, no. 3, pp. 285–289, 2009.

[19] S. Gilmour and M. Dras, "Understanding the pheromone system within ant colony optimization," in *Australasian Joint Conference on Artificial Intelligence*, pp. 786–789, Springer, 2005.

[20] G. Javadzadeh and R. Azmi, "Idufg: Introducing an intrusion detection using hybrid fuzzy genetic approach.," *International Journal of Network Security*, vol. 17, no. 6, pp. 754–770, 2015.

[21] R. Jensen, *Combining Rough and Fuzzy Sets for Feature Selection*, PhD thesis, Citeseer, 2005.

[22] X. JunYong, H. Xiang, L. CaiYun, and C. Zhong, "A novel parallel ant colony optimization algorithm with dynamic transition probability," in *International Forum on Computer Science-Technology and Applications (IFCSTA'09)*, vol. 2, pp. 191–194, 2009.

[23] A. Kannan, G. Q. Maguire Jr, A. Sharma, and P. Schoo, "Genetic algorithm based feature selection algorithm for effective intrusion detection in cloud networks," in *2012 IEEE 12th International Conference on Data Mining Workshops*, pp. 416–423, 2012.

[24] F. Kuang, W. Xu, and S. Zhang, "A novel hybrid kpca and SVM with ga model for intrusion detection," *Applied Soft Computing*, vol. 18, pp. 178–184, 2014.

[25] Y. Li and S. Gong, "Dynamic ant colony optimisation for tsp," *The International Journal of Advanced Manufacturing Technology*, vol. 22, no. 7-8, pp. 528–533, 2003.

[26] Z. Li and A. Das, "The utility of partial knowledge in behavior models: An evaluation for intrusion detection," *International Journal of Network Security*, vol. 1, no. 3, pp. 138–146, 2005.

[27] S. Lin, K. Ying, C. Lee, and Z. Lee, "An intelligent algorithm with feature selection and decision rules applied to anomaly intrusion detection," *Applied Soft Computing*, vol. 12, no. 10, pp. 3285–3290, 2012.

[28] Y. Lin and J. Zhang, "Ant colony optimization with adaptive heuristics design," in *Proceedings of the 15th Annual Conference Companion on Genetic and Evolutionary Computation*, pp. 3–4, ACM, 2013.

[29] G. Liu, Z. Yi, and S. Yang, "A hierarchical intrusion detection model based on the pca neural networks," *Neurocomputing*, vol. 70, no. 7, pp. 1561–1568, 2007.

[30] M. J. Meena, K. R. Chandran, and J. M. Brinda, "Integrating swarm intelligence and statistical data for feature selection in text categorization," *Pheromones*, vol. 1, p. 15, 2010.

[31] D. Merkle and M. Middendorf, "Swarm intelligence," in *Search Methodologies*, pp. 213–242, Springer, 2014.

[32] R. Montemanni, L. M. Gambardella, A. E. Rizzoli, and A. V. Donati, "Ant colony system for a dynamic vehicle routing problem," *Journal of Combinatorial Optimization*, vol. 10, no. 4, pp. 327–343, 2005.

[33] A. Moraglio, F. E. Otero, and C. G. Johnson, "The ACO encoding," in *International Conference on Swarm Intelligence*, pp. 528–535, Springer, 2010.

[34] L. Nunes de Castro, "Fundamentals of natural computing: an overview," *Physics of Life Reviews*, vol. 4, no. 1, pp. 1–36, 2007.

[35] Z. A. Othman, H. M. Rais, and A. R. Hamdan, "Embedding malaysian house red ant behavior into an ant colony system," *Journal of Computer Science*, vol. 4, no. 11, p. 934, 2008.

[36] W. Pan and L. Wang, "An ant colony optimization algorithm based on the experience model," in *2009 Fifth International Conference on Natural Computation*, vol. 3, pp. 13–18, 2009.

[37] S. Parsons, *Ant Colony Optimization by Marco Dorigo and Thomas StÜTzle*, MIT Press, ISBN 0-262-04219-3, 2005.

[38] Q. S. Qassim, A. M. Zin, and M. J. Ab Aziz, "Anomalies classification approach for network-based intrusion detection system," *International Journal of Network Security*, vol. 18, no. 6, pp. 1159–1172, 2016.

[39] H. M. Rais, Z. A. Othman, and A. R. Hamdan, "Improved dynamic ant colony system (dacs) on symmetric traveling salesman problem (tsp)," in *International Conference on Intelligent and Advanced Systems (ICIAS'07)*, pp. 43–48, 2007.

[40] M. H. Rasmy, M. El-Beltagy, M. Saleh, and B. Mostafa, "A hybridized approach for feature selection using ant colony optimization and ant-miner for classification," in *8th International Conference on Informatics and Systems (INFOS'12)*, pp. BIO–211, 2012.

[41] I. J. Riadi, *Cognitive Ant Colony Optimization: A New Framework In Swarm Intelligence*, PhD thesis, University of Salford, 2014.

[42] K. I. Rufai, R. C. Muniyandi, and Z. A. Othman, "Improving bee algorithm based feature selection in intrusion detection system using membrane computing," *Journal of Networks*, vol. 9, no. 3, pp. 523–529, 2014.

[43] C. Scheper and W. J. Roberts, "Anomaly detection using an mmpp-based glrt.," *International Journal of Network Security*, vol. 17, no. 6, pp. 672–677, 2015.

[44] W. Shahzad, *Classification and Associative Classification Rule Discovery Using Ant Colony Optimization*, PhD thesis, National University of Computer & Emerging Sciences, 2010.

[45] M. Sheikhan, M. S. Rad, and H. M. Shirazi, "Application of fuzzy association rules-based feature selection and fuzzy artmap to intrusion detection," *Majlesi Journal of Electrical Engineering*, vol. 5, no. 4, 2011.

[46] Q. Shen, J. Jiang, J. Tao, G. Shen, and R. Yu, "Modified ant colony optimization algorithm for variable selection in qsar modeling: Qsar studies of cyclooxygenase inhibitors," *Journal of Chemical Information And Modeling*, vol. 45, no. 4, pp. 1024–1029, 2005.

[47] R. K. Sivagaminathan and S. Ramakrishnan, "A hybrid approach for feature subset selection using neural networks and ant colony optimization," *Expert Systems with Applications*, vol. 33, no. 1, pp. 49–60, 2007.

[48] C. Solnon and D. Bridge, "An ant colony optimization meta-heuristic for subset selection problems," *System Engineering using Particle Swarm Optimization, Nova Science*, vol. 729, 2006.

[49] T. Stützle and M. Dorigo, "ACO algorithms for the traveling salesman problem," *Evolutionary Algorithms in Engineering and Computer Science*, pp. 163–183, 1999.

[50] M. Uddin, A. A. Rahman, N. Uddin, J. Memon, R. A. Alsaqour, and S. Kazi, "Signature-based multilayer distributed intrusion detection system using mobile agents.," *International Journal of Network Security*, vol. 15, no. 2, pp. 97–105, 2013.

[51] J. Van Ast, R. Babuška, and B. De Schutter, "Generalized pheromone update for ant colony learning in continuous state spaces," in *IEEE Congress on Evolutionary Computation*, pp. 1–8, 2010.

[52] W. Wang, X. Zhang, S. Gombault, and S. J Knapskog, "Attribute normalization in network intrusion detection," in *10th International Symposium on Pervasive Systems, Algorithms, and Networks*, pp. 448–453, 2009.

[53] A. Zainal, M. A. Maarof, and S. M. Shamsuddin, "Feature selection using rough-dpso in anomaly intrusion detection," in *International Conference on Computational Science and Its Applications*, pp. 512–524, Springer, 2007.

[54] J. Zeng and D. Guo, "Agent-based intrusion detection for network-based application.," *International Journal of Network Security*, vol. 8, no. 3, pp. 201–210, 2009.

Helmi Md Rais is a senior lecturer at Universiti Teknologi PETRONAS (UTP) under the Faculty of Science and Information Technology (FSIT). He received his PhD degree in Science and System Management in 2013 from Universiti Kebangsaan Malaysia (UKM), Malaysia. He received his BSc degree from Drexel University, USA in 1999 and his Master degree from Griffith University, Australia in 2001. His research interests include swarm intelligent, database and management information systems.

Tahir Mehmood received his BS in Computer Science from University of Peshawar, Pakistan. He is currently pursuing MS in Information Technology from Universiti Teknologi PETRONAS, Malaysia. His research interests include machine learning, big data, image processing, data science, and network security.

Array Erasure Codes with Preset Fault Tolerance Capability

Dan Tang, Ya-Qiang Wang, Hao-Peng Yang
(Corresponding author: Dan Tang)

Software Engineering College & Chengdu University of Information Technology
No.24 Block 1, Xuefu Road, Chengdu 610225, China
(Email: tangdan@foxmail.com)
(Received Jan. 3, 2017; revised and accepted Mar. 11, 2017)

Abstract

The array erasure code, an ideal method for fault tolerance in storage systems, however, is obstructed by its impossibility to set the fault tolerant ability according to dynamic application environment for practical purpose. In view of this, this paper presents a new class of array erasure codes, with the greatest contribution to the array codes which can be obtained according to the preset number of fault tolerance and storage efficiency in dynamic structure, and fault tolerance capability can be presented is not limited in theory. In addition, the new array code has the advantages of simple structure, easy to realize, with no strong constraints to satisfy in structure. Only binary XOR operations are required for coding and decoding for a high operational efficiency, and fixed update penalty and repair cost which will not increase with the expansion of system size or increase of fault tolerance capability.

Keywords: Array Erasure Codes; Preset Fault Tolerance; Strip Size; Weak Constraints

1 Introduction

At present, we have entered into the era of big data, with all industry data explosively growing. More and more data have become an important part of the normal operation of society. In order to deal with the data storage reliability problem caused by rapid growth of data quantity, increasing the stability of a single storage node is certainly a kind of theoretically feasible method. But more effective approach is to use multiple storage nodes to build a storage system, which can make full use of existing equipments, increase storage capacity and improve data access efficiency in parallel, and with a certain fault tolerance strategy, can also effectively enhance the reliability of the whole storage system. In this case, the number of storage nodes is usually used to represent the size of the storage system. At present, large storage systems with more than 100 nodes have become more and more popular, Google and other companies even have some PB level storage systems with more than 3,000 nodes[8] . Compared with the past, although the reliability of a single storage node has been enhanced, in a large storage system with a great amount number of nodes, the probability of failure of multiple nodes in a period is still very high. According to statistics from the Carnegie Mellon University, annual failure and replacement rate of a large storage system with more than 300 nodes is about 5.1% [9]. Of course, this is the probability of failure of nodes in the storage system under normal operation conditions, and if floods, landslides, earthquakes and other natural disasters and administrators misoperation, hacker attacks and other human factors are taken into consideration, fault tolerance and storage system reliability enhancement is more a concern.

Replication technology is the most mature fault tolerance technology in the field of reliability enhancement of the storage system. A fault tolerance scheme depending on replication technology has multiple copies of the same data stored in different nodes of the system. As copies of each node are exactly the same, usually it is not required to strictly distinguish between the parity data (redundant data) and the original data. When a node fails, a copy in any node which is not failed can recover the missing data. The fault tolerance method depending on replication technology is simple and intuitive, easy to realize and dynamically expandable, but the low storage efficiency is a huge defect. Assuming replication technology is used to construct a storage system of t fault tolerance, it needs to copy the original data into a $t+1$ copies to be stored in different nodes, that is, the storage efficiency is only $1/(t+1)$. Especially for large storage systems, this extremely low effective utilization of storage space is very difficult to accept.

Storage system as a whole shall consider not only a certain performance parameter, but the most important performance indicators according to the application environment. Erasure codes can balance all main perfor-

mances of storage system to a certain extent, and it is a kind of fault tolerance method which is more and more important. At present, the RS erasure code is the most widely used in the storage system. The relevant mathematical theory of RS codes is mature, with regular codeword structure and unlimited fault tolerance capability in theory, and the codes have the MDS property, satisfying the theoretically optimal code rate (corresponding to the storage efficiency in a storage system). However, RS codes coding and decoding are performed on a multivariate finite field by complex operations, especially for multiplication and inversion in a finite field. Therefore, the main problem to be solved by fault tolerance method based on RS codes for storage system is not to optimize the coding process, but how to improve the operational efficiency over the finite field. Of course, there are also some scholars who have made a very fruitful work on this issue, increasing operational efficiency over a finite field greatly, and one of the most representative is the reduction computing program for a finite filed[7] and the finite field calculation scheme GF-Complete[6] by Plank et al. In spite of this, the current storage system is still difficult to bear the cost of RS code operations, especially the large storage system.

Array erasure code (usually referred to as array code) is a kind of erasure codes using binary XOR coding and decoding operation, with high operational efficiency. The special two-dimensional coding structure is seemingly complex but can be used completely corresponding to two-dimensional data layout structure normally used in the current storage system of multiple nodes, so it is suitable for use in the storage system. However, in addition to the small size centralized RAID6 system, the EVENODD code[1] is chosen as one of the alternative methods for tolerating 2 failures, few array codes are used in the current commercial storage system. Such a situation can be attributed to the fixed fault tolerance capability of most array codes which are not easy to expand. Such as EVENODD codes[1], X codes[12] are 2 fault tolerance array codes, as long as the constraints of these coding structure are satisfied and in accordance with the coding method, it can make the node fault tolerance of storage system at 2, but only 2. In other words, as long as the use of EVENODD codes or X codes as a fault tolerance method, regardless of the size of a storage system, the maximum fault tolerance is only 2. Similarly, with the use of Star codes[3] and extended X codes[5] and other 3 fault tolerance array codes as a fault tolerance method, the maximum fault tolerance of the storage system is also fixed to 3. Grid codes[4] use two array codes with typical horizontal or vertical data layout as the matched codes, which can be used to obtain high fault tolerance with coding in the horizontal and vertical directions simultaneously. But once two matched codes are determined, the fault tolerance capability of the Grid codes is also determined. Weaver codes[2] can determine the codeword structure according to the requirement of fault tolerance capability, with maximum fault tolerance up to 12, but the fault tolerance capability of the codes is lack of theoretical support, that is, the fault tolerance capability is obtained by computer check. In addition, the storage efficiency of the codes is always lower than 50%, and will quickly decline with the increase of fault tolerance. Reference[10] proposes a new way to reversely determine the structure of coding by fault tolerance, but specific implementation method is not provided.

In view of above problems, this paper presents a new class of array erasure codes, which use the horizontal data layout. All calculations in the new array codes completely are binary field XOR, with high operational efficiency. The repair cost and update penalty are fixed constants, which will not increase with the expansion of storage system size. The codes can be constructed according to the fault tolerance requirements for running storage system, with fault tolerance not restricted in theory, but also the storage efficiency can be set in a particular fault tolerance capability. After a small transformation, the new codes can also be used in areas other than storage[11]. All of these properties make the new array codes already available on the basis of practical application in large storage systems.

2 Array Codes Capable of Presetting Fault Tolerance

Some basic concepts in the field of storage coding, such as data, check, parity, redundancy, element, strip, stripe, horizontal array codes, vertical array codes, coding, decoding, data reconstruction, can refer to literatures[1, 12, 3, 5, 2] for definitions. This paper will continue to use the definitions of these concepts which will not be repeated. The new array code proposed in the paper uses the horizontal data layout structure, so the entire storage array can be divided into data element array and check element array in logic. A column in the storage array corresponds to a storage node, and in the horizontal data structure each column or all shall be data elements, or check elements. A node failure shall mean that all elements the node corresponds to become unknown. For convenience in description, this paper defines the following symbols. In case of no special note afterward, the meaning of these symbols is used here. It is agreed that f is the fault tolerance of the proposed array code to be constructed, n for the number of columns of data element array, r for the number of columns of check element array, m for the number of rows of storage array. Obviously f, m, n, r are positive integers.

2.1 The $m = 2$ Case

For the f fault tolerance array codes, at least f groups is required with n linear independent check equations in each group, that is, the deployment of f group coding chains with different slopes. This paper will use the positive and negative expansion of numbers to determine

the different slopes of deployed coding chains, that is, all slopes of coding chains are taken from the set {1, -1, 2, -2,... }. Actually, the nature of the coding chain is the linear relationship between data elements and check elements. The specific concepts and definitions of the coding chain and its slope shall refer to the literature[10], which will not be repeated hereof. When $f = 1$, it only needs to deploy 1 group of coding chains. In a storage array of $n = 2$, a group of coding chains with slope of 1 is deployed, as shown in Figure 1, in which the XOR sum of elements with the same background color is 0. Of course, the relationship between data elements and check elements can also be expressed by linear equations, and the check relationship among various elements in the storage array in Figure 1 can be expressed by the equation group (1).

$$\begin{cases} d(1,1) + d(2,2) = c(1) \\ d(1,2) + d(2,1) = c(2) \end{cases} \tag{1}$$

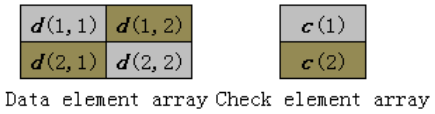


Figure 1: Relationship among elements with $f = 1$

In Figure 1, if any one of columns in the storage array fails, in each equation in the equation group (1) one element will become unknown. Obviously, the unknown element can be obtained by the other two known elements in XOR equation. Therefore, any one invalid column in the array can be effectively recovered.

When fault tolerance is required to reach 2, then 2 sets of coding chains with different slopes shall be deployed. In a storage array with $n = 4$, two sets of coding chains with slopes at 1 and -1, respectively, as shown in Figure 2, are deployed. As to why the n must be equal to or greater than 4, answers will be given below. Figure 2 (a) and (b) show the deployment of 2 coding chains with slopes of 1 and -1 respectively, in which the XOR sum of elements with the same background color is 0. Of course, the check relationship can also be expressed by equation group (2).

$$\begin{cases} d(1,1) + d(2,2) = c(1); d(1,2) + d(2,3) = c(2) \\ d(1,3) + d(2,4) = c(3); d(1,4) + d(2,1) = c(4) \\ d(1,1) + d(2,4) = c(5); d(1,2) + d(2,1) = c(6) \\ d(1,3) + d(2,2) = c(7); d(1,4) + d(2,3) = c(8) \end{cases} \tag{2}$$

After any 2 columns in storage array shown in Figure 2 fail, it can always find at least 2 unknown elements in an equation with only 1 unknown element, and the element values could be known by XOR. After this, if there are other unknown elements, then all exist in the equation of only 1 unknown element, and can be easily recovered.

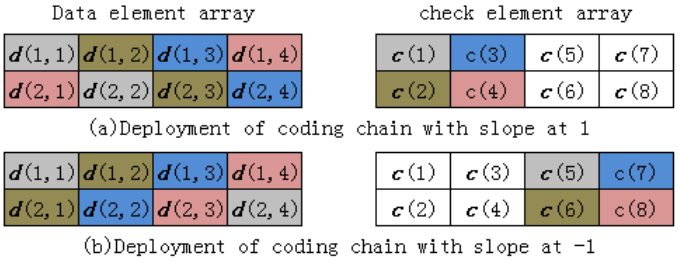


Figure 2: Relationship among elements with $f = 2$

2.2 General Coding and Data Reconstruction Method

In a storage array, $d(i, j)$ is used to represent the element at row i and column j in the data element array, and $c(t)$ as the t check element in the check element array. Among them check elements are arranged at priority, where i, j and t are positive integers, and $1 \le i \le m$, $1 \le j \le n$, $1 \le t \le m \cdot r$. The element identification on the storage array is shown in Figure 3.

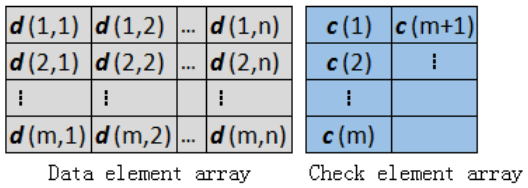


Figure 3: Element ID in storage array

In the identification system of above data elements and redundant elements, each check element may be calculated by expression (3) .

$$c(t) = \sum_{i=1}^{m} d(i, l_c + i \cdot (2 \cdot (l_r \% 2) - 1) \cdot \lceil l_r/2 \rceil - 1) \% n + 1). \tag{3}$$

Where, l_r represents the l_r coding chain used in the present check element, l_c represents that the slope of coding chain which is used for the l_c time, which can be obtained by expression (4), in which, the operator "[]" represents rounding while " %" modulus.

$$\begin{cases} l_r = [(t-1)/n] + 1 \\ l_c = (t-1) \% n + 1 \end{cases} \tag{4}$$

Example 1. *Maximum fault tolerance capability of storage system as 3, constructing the corresponding array code.*

It is assumed that the storage array strip size $t = 3$, and 7 columns of data element array in the storage array. As previously stated, the maximum fault tolerance capability f is 3, the size of data element array is 3×7, the number of check elements is 21, so the number of columns in the check element array is also 7. The maximum fault tolerance capability is 3, so coding chains with slopes of 1, -1 and 2 are deployed. Specific deployment process as

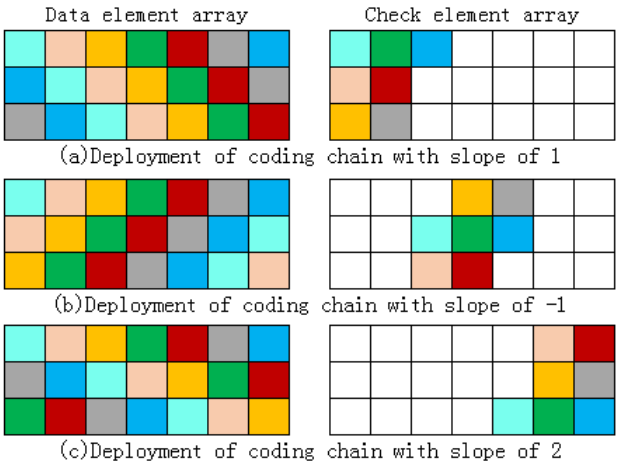


Figure 4: Relationship among elements with $f = 3$

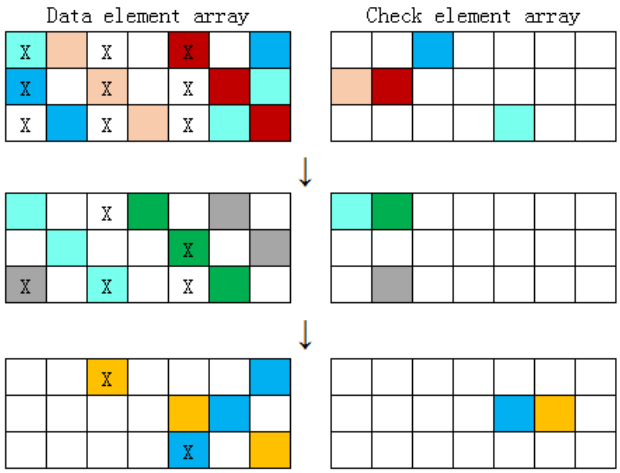


Figure 5: Successful recovery of failure node data

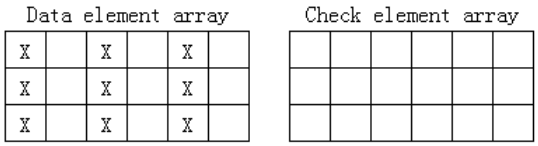


Figure 6: Failed failure node data recovery

shown in Figure 4, XOR sum of elements with the same background color is 0.

We can use the method of this paper to construct a specific fault tolerance array code according to the requirement of application environment. And the next will be a brief introduction of how to restore and reconstruct after the failure of nodes. Again, this paper only considers the case of node error, that is, the entire storage node failure causes the loss of all data elements on the node, which also corresponds to the column failure in the storage array. The basic idea of data recovery on the failure node can be summed up as, that is, to find the coding chains with only one failure element, obviously the failure element on the coding chain can be calculated by other valid elements. Continue to repeat this process until all the invalid elements are fully recovered.

Example 2. *It is assumed that the nodes 1, 3 and 5 in the storage system fail, with all failure nodes replaced and renewed, the elements at 1, 3 and 5 in the data element array in storage array become unknown. The whole data recovery process is shown in Figure 5. Elements marked with "X" represent invalid element with unknown value, and elements with the same background color shall be a coding chain with only 1 failure element, which could be recovered by XOR of other elements on the coding chain.*

Continue to extend the example. It is assumed that the size of data element array is 3×6, and elements on 1, 3, and 5 columns fail. However, as shown in Figure 6, it is impossible to find any coding chain that contains only a failure element. In this case, all data in the storage system are lost.

2.3 Constraint Conditions and Fault Tolerance Capability Guarantee

From the example 2 in the previous section it can be concluded that, if you need to grant the storage array of strip size 3 with fault tolerance of 3, in addition to the deployment of 3 coding chains of different slopes, the number of columns of data element array shall be at least not less than 7, otherwise fault tolerance is not guaranteed. By further expansion of this conclusion, the use of f coding chains with different slopes to construct array code cannot guarantee the fault tolerance f, the number of fault tolerance and array size also need to meet a certain conditions. According to the data reconstruction method of the previous section, to ensure that at least 1 failure element in one or more coding chains with only 1 failure element, then the fault tolerance f, strip size m and the number of columns of data element array n shall meet the conditions:$n \geq m \cdot f - f + 1$. The fault tolerance capability can be guaranteed in the case of satisfying the constraints as proved below.

For array codes of horizontal layout, the occurrence of node level failure can be divided into the following three types: 1. All failure nodes in the check element array; 2. Failure nodes in check element array and data element array; 3. All failure nodes in the data element array. When all of the failure nodes are in the check element array, it just needs to re-do a coding operation after the replacement of the failure nodes. This is the simplest case of data reconstruction. When the data element array and the check element array have failure nodes, it is relatively easy to deal with. In the array code with f fault tolerance, each data element is passed by f coding chains, so a data element is related to f check elements. According to the constraint conditions it can be known that, so two check elements associated with the same data element must not appear in the same column. As a result, all of the failure elements in data element array can be recovered completely, and then all the failure elements in the

check element array can be recovered by coding. And for all of the f failure nodes in the data element array, the following mathematical induction is used to prove.

First of all, we mark the f failure columns in data element array as $E_1, E_2 \cdots, E_f$, where d_i shall be the distance between the failure column E_i and the failure column on its right. It is assumed that $\max(d_1, d_2 \cdots, d_f) = d_f$, where $i = 1, 2 \cdots, f$. Therefore, by the principle of the pigeon cage, it is known that $d_f \geq m$ is established. At that time, when there was only one failure column, the column will be recorded as E_1. By the coding chain deployment method it can be known, each data element is passed by one coding chain, apparently all the failure elements on the column can be successfully restored. Assuming that all the failure elements can be recovered when $f = k$, where k is a positive integer. In the following, a discussion on $f = k + 1$ is conducted. Be known by the precondition, $\max(d_1, d_2 \cdots, d_k) = d_k \geq m$. Therefore, if the inequality $d_1 \geq \lceil m/2 \rceil$ is established, it is clear that all elements on the first failure column E_1 can be recovered by the coding chains with slopes of 1 or -1. In the same way, if $d_k \geq \lceil m/2 \rceil$ is established, all the elements on the final failure column E_{k+1} can be recovered by coding chains with slopes of 1 or -1. At this point, the above two situations can be changed into $f = k$, and according to the aforementioned assumptions, all the failure data can be recovered. When $d_1 < \lceil m/2 \rceil$ and $d_k < \lceil m/2 \rceil$, the first failure element in all of the failure columns can be recovered by coding chains with slopes of $\pm 1, \pm 2, \cdots, \pm k, k+1$. And when all the failure elements in the first row on all failure columns are recovered successfully, only to repeat the same steps, the remaining failure elements can be effectively restored. Quod erat demonstrandum.

The inequality $n \geq m \cdot f - f + 1$ is the constraint condition that is needed to satisfy when we are constructing the specific fault tolerance capability of array codes in this paper.

2.4 Satisfaction of Expected Storage Efficiency

Storage efficiency is an important performance index to measure the effective utilization of storage space. As mentioned earlier, in the storage system reliability enhancement field, replication technology is the method of fault tolerance without any operations, and the principle is simple and easy to implement. It is usually not recommended the use of replication technology in large storage systems in consideration of storage efficiency. Compared with that, the fault tolerance system based on erasure codes can greatly improve the storage efficiency under the condition of the same fault tolerance capability. Of course, in this process, the calculations are required. For array codes, as mentioned before, with the storage efficiency as the standard, it can be divided into MDS codes and non MDS codes. MDS codes have a theoretical optimal value in storage efficiency, and its typical representation includes EVENODD codes[1], X codes [12], Star codes[3] and extended X codes[5] and so on. But the fault tolerance capability of array codes with MDS property is only 2 or 3, which obviously cannot meet the demand of modern large storage system reliability enhancement. In order to improve the fault tolerance of array codes, researchers have designed some array codes without the MDS property, with fault tolerance capability greatly improved compared with array code with MDS property, but mostly with great sacrifice in storage efficiency, such as the Weaver codes with fault tolerance at 10, the storage efficiency is less than 20%.

From the last section, we can know that the total number of data elements in the array erasure codes generated by this method is $m \cdot n$, and check elements $f \cdot n$. Obviously, the storage efficiency of array erasure codes constructed by the new method is $m \cdot n/(m \cdot n + f \cdot n) = m/(m + f)$. Therefore, the factor affecting the storage efficiency in specific fault tolerance is the strip size m. Figure 7 shows the storage efficiency varies with the size of strip size with the fault tolerance at 20, 30, 50 and 100. Obviously, with fault tolerance capability unchanged, increasing strip size can effectively improve the storage efficiency. Thus, under a certain fault tolerance, of course, it can also set the storage efficiency.

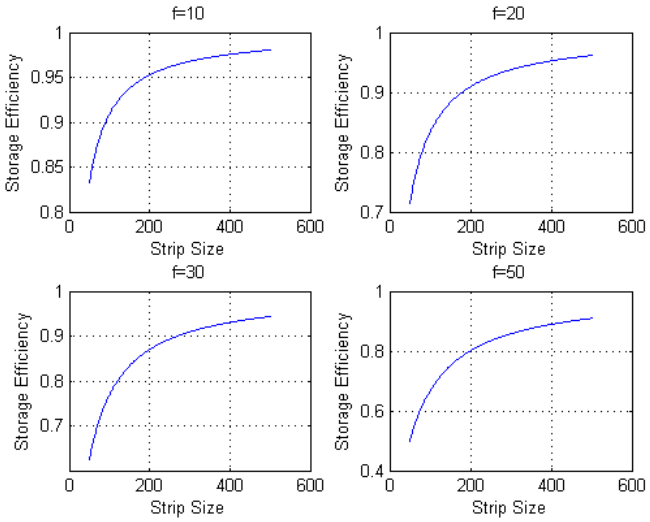


Figure 7: Influence of strip size on storage efficiency

Example 3. *Fault tolerance capability is 4, and storage efficiency is not less than 80%, constructing the corresponding array code.*

According to the previous description of array codes construction method, the storage efficiency can be expressed as $m \cdot n/(m \cdot n + f \cdot n) = m/(m + f)$, then the inequality $m/(m + f) \geq 0.8$ needs to be established. Then the inequality $m \geq 4 \cdot f$ holds, and from the preconditions it is known that when $f = 4$, $m \geq 16$. In this example $m = 16$, then it is inferred that $n \geq 61$ when $n \geq m \cdot f - f + 1$. In this example, $n = 61$, we can know the total number of check elements $f \times n = 4 \times 61 = 244$, and it is inferred that the check node number is $\lceil 244/16 \rceil = 16$.

And then, the basic information is known throughout the storage array. The array has 16 rows, namely the strip size is 16. The data element array has 61 columns, check element array has 16 columns, a total of 976 data elements, and 244 check elements, with storage efficiency of exactly 80%. According to the proof of the previous section, respectively with the 4 coding chains with slope of $\pm 1, \pm 2$ deployed, the fault tolerance capacity of the array code is up to 4.

3 Experiments and Analysis

3.1 Operational Burden

Because most of the code word structures of array codes are irregular and it is difficult to use a precise formula to express the overall workload of coding or data reconstruction. Usually the times of XOR calculations of generating a check bit is used to measure the complexity of coding, and the times of XOR calculations to recover a lost data element for the complexity of data recovery. It can be known from the preceding text that the length of each coding chain is $m+1$, so that the operational effort required to generate a single check element or to recover a single failure element is $m-1$. However, stripe size grows with the enlargement of strip size m or fault tolerance f, which means that the number of check elements will also increase, so does the overall operational burden. Therefore, changes in the strip size and fault tolerance would affect coding and data recovery calculations, as one of the factors we need to consider. Assuming that the strip size is 200, fault tolerance 50, with the storage of 1G data as an example, XOR calculations needed to generate all check elements or recover failure data on 50 nodes is about 4.2×10^9, and it takes a general personal computer with dominant frequency of 3.2G 16s to complete. Figure 8 shows the influence of strip size increasing on array code coding and data recovery operation and Figure 9 the influence of fault tolerance on coding and data recovery operational burden.

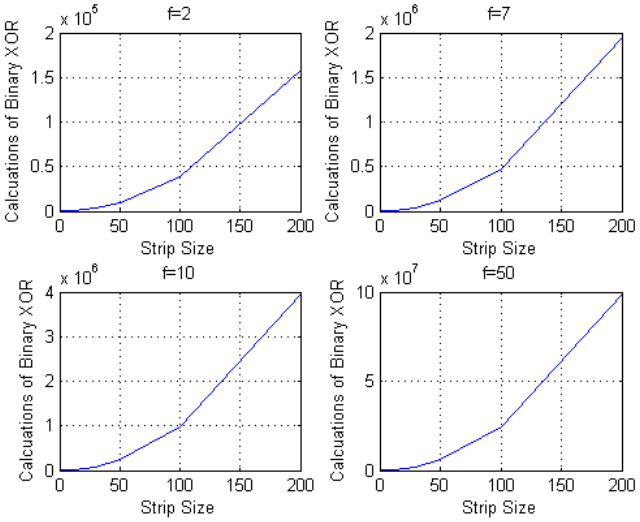


Figure 8: Influence of strip size on operational burden

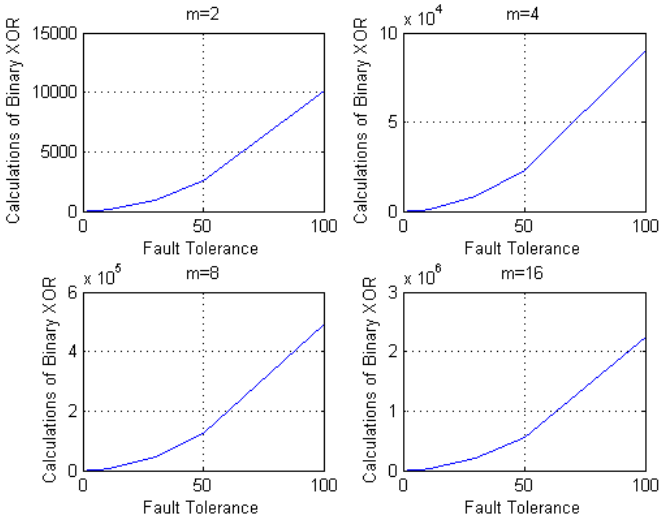


Figure 9: Influence of fault tolerance on operational burden

3.2 Constraint Conditions of Array Code Construction

Most of the array erasure codes in the construction process will restrict the storage array size namely stripe size or strip size. Satisfying the constraints is the basic condition of coding to reach a certain fault tolerance capability. Table 1 shows the comparison of several kinds of array codes during construction with constraints on stripe size and strip size, where p represents a prime number, s_r stripe size, s_c strip size, obviouslys_r, s_c are positive integers greater than 1, which will no longer be listed in the table. From this table it is not difficult to find that most array erasure codes have very strict constraints on stripe size or strip size, and usually requires stripe size as prime or satisfies a linear relationship with a prime. Constraints on the stripe or strip size will greatly limit the application and extension of array codes. In the practical process of array codes, in order to deal with the situation mentioned above, researchers have proposed a number of compromise solutions, the most typical approach is when the number of nodes in storage array is not a prime, add 1 or more virtual nodes with storage data values as 0 to complement the prime. In the process of coding or data recovery these virtual nodes participate in operations. However, with the expansion of storage scale, increase of storage node data, the interval between adjacent primes is expanding, virtual node number will increase dramatically to complement the prime, along with the invalid calculations of encoding and data reconstruction increase. As a result, such schemes are greatly restricted in large storage. Weaver codes have a high fault tolerance, without any special requirements on stripe size and strip size, but the code is not based on any systematical coding method and lack of theoretical support. The fault tolerance ca-

pability of Weaver codes depends on computer engineering test, so it is difficult to apply in large storage system. Grid codes have better fault tolerance capability, to achieve simple storage with efficiency up to 80%, but the code in the construction process need to find two other array codes (must be the typical horizontal code or vertical code) to form matched codes. Fault tolerance capability of Grid codes depend on the fault tolerance capability of the matched codes selected. In addition, the Grid code is not a typical horizontal or vertical array code, and its storage layout combines the characteristics of horizontal and vertical arrays. Such a pattern brings in high fault tolerance and at the same time also makes its expansion subject to certain constraints in practical situation.

Table 1: Constraints on construction of array erasure codes

Array Codes	Stripe Size	Strip Size
EVENODD Codes	$s_r = p$	$s_c = s_r - 1$
X Codes	$s_r = p$	$s_c = s_r$
Liberation Codes	N/A	$s_c = p, p \geq s_r$
B Codes	TBD by the mathematical problem solving results	$s_c = 2 \cdot s_r \| s_c = 2 \cdot s_r + 1$
P Codes	$s_r = p \| s_r = p - 1$	$s_c = s_r/2$
RDP Codes	$s_r = p - 1$	$s_c = s_r$
Star Codes	$s_r = p$	$s_c = s_r - 1$
Weaver Codes	N/A	N/A
Grid Codes	TBD by matched codes	TBD by matched codes
Array codes presented in this paper	$s_r - \lceil f \cdot n/m \rceil \geq f \cdot (s_c - 1) + 1$	N/A

Unlike construction of most other array codes, the new method is based on the preset fault tolerance to construct the corresponding array codes, so there is no special constraint similar to prime for the array size. But as mentioned earlier, fault tolerance capability f, strip size m and column number n in data element array shall satisfy a linear constraint $n \geq m \cdot f - f + 1$. The number of fault tolerance f and block size m are usually determined by application environments, and when the two parameters are determined, n is determined. And under specific environment, when it is needed to limit the n, by the fault tolerant quantity and above inequality the value of m is calculated. Obviously, the constraint condition is easily satisfied, that is, the strength of the constraint condition is very weak.

3.3 Repair Cost and Update Penalty

Repair cost usually refers to the total number of storage nodes that are required to reconstruct a failure data element. The repair cost is an important performance index in the storage system, which is closely related to data reconstruction, data update and degraded reading and writing. By the coding chains deployment method, each chain has $m+1$ elements, including m data elements and 1 check element. Therefore, the number of nodes for reconstruction of a failure data element is m, which will not increase with the increase of storage system scale and fault tolerance capability.

Update penalty is an unique index in the horizontal layout array code, which refers to the number of check nodes that need to change 1 minimum data bit. When the data update is more frequent, the update penalty is too high, which will lead to the check node's access overheat and reduce the overall I/O performance of the storage system. In this paper, we can know that each data element has f different code chains, that is, each data element is related with f different check elements. In other words, when any one data element is changed, it always involves the f check nodes, that is, the update penalty is f, and it has reached the theoretical optimal value of the f fault tolerance array code.

4 Conclusion

Same with most current array codes, the proposed array codes still use only binary XOR in coding, with high operational efficiency. Differing from most previous array code construction methods, the method proposed in this paper is a kind of array code construction method based on specific fault tolerance, which can construct any array code of any fault tolerance capability according to environmental requirements. And satisfying the specific fault tolerance, by changing the strip size it adjusts the storage efficiency. In addition, the proposed array codes are subject to weak constraints when constructing, which are easy to meet. Finally, the array codes constructed by the proposed method also have the theoretical optimal update penalty and the repair cost which does not change with the storage size and fault tolerance capability. It is hoped that the method proposed in this paper can be used for the practical application of array codes in the field of storage, and it plays a positive role in other fields.

Acknowledgments

This paper is supported by the National Natural Science Foundation of China (Grant No. 61501064) and Sichuan Provincial Science and Technology Project (Grant No. 2017JQ0057, 2016GZ0122). And I want to take my deepest gratitude to Professor WANG Xiao-Jing for his guidance.

References

[1] M. Blaum, J. Brady, J. Bruck, and J. Menon, "Evenodd: An optimal scheme for tolerating double disk failures in raid architectures," *ACM Sigarch Com-*

puter Architecture News, vol. 22, no. 2, pp. 245–254, 1995.

[2] J. L. Hafner, "Weaver codes: Highly fault tolerant erasure codes for storage systems," in *Conference on File and Storage Technologies*, pp. 16–16, 2005.

[3] C. Huang and L. Xu, "Star:an efficient coding scheme for correcting triple storage node failures," in *Conference on Usenix Conference on File and Storage Technologies*, pp. 15–15, 2005.

[4] M. Li, J. Shu, and W. Zheng, "Grid codes: strip-based erasure codes with high fault tolerance for storage systems," *ACM Transactions on Storage*, vol. 4, no. 4, pp. 1–22, 2009.

[5] Q. C. Meng, *Research of Erasure Codes Applied in Distributed Storage System*, PhD thesis, Beijing: Graduate University of the Chinese Academy of Sciences, 2007.

[6] J. S. Plank, K. M. Greenan, and E. L. Miller, "Screaming fast galois field arithmetic using intel simd instructions," in *Fast: Usenix Conference on File and Storage Technologies*, 2013.

[7] J. S. Plank and L. Xu, "Optimizing cauchy reed-solomon codes for fault-tolerant network storage applications," in *IEEE International Symposium on Network Computing and Applications*, pp. 173–180, 2006.

[8] M. Sathiamoorthy, M. Asteris, D. Papailiopoulos, A. G. Dimakis, R. Vadali, S. Chen, and D. Borthakur, "Xoring elephants: Novel erasure codes for big data," *Proceedings of the Vldb Endowment*, vol. 6, no. 5, pp. 325–336, 2013.

[9] B. Schroeder and G. A. Gibson, "Disk failures in the real world: what does an MTTF of 1,000,000 hours mean to you?," in *Usenix Conference on File and Storage Technologies*, p. 1, 2007.

[10] D. Tang and H. P. Shu, "A class of array erasure codes with high fault-tolerance," *Scientia Sinica Informationis*, vol. 46, no. 4, pp. 523–538, 2016.

[11] J. Wang, X. Yu, and M. Zhao, "Fault-tolerant verifiable keyword symmetric searchable encryption in hybrid cloud," *International Journal of Network Security*, vol. 17, no. 4, pp. 471–483, 2015.

[12] L. Xu and J. Bruck, "X-code: Mds array codes with optimal encoding," *IEEE Transactions on Information Theory*, vol. 45, no. 1, pp. 272–276, 1999.

Biography

Dan Tang received PhD degree from Graduate University of Chinese Academy of Sciences, associate professor. His research interests include coding theory, secret sharing.

Ya-Qiang Wang received PhD degree from the College of Computer Science, Sichuan University. He is a lectur at Chengdu University of Information Technology. His research interests include big data analysis, fault tolerance.

Hao-peng Yang undergraduate student of Chengdu University of Information Technology. His research interests include big data analysis, fault tolerance.

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.