IJNS

International Journal of Network Security

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

Vol. 23, No. 6 (November 2021)

INTERNATIONAL JOURNAL OF NETWORK SECURITY

Editor-in-Chief

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

Co-Editor-in-Chief:

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

Publishing Editors

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

Board of Editors

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

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

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

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

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

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

Chi-Shiang Chan
Department of Applied Informatics & Multimedia, Asia University (Taiwan)

Chen-Yang Cheng
National Taipei University of Technology (Taiwan)

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

Stefanos Gritzalis
University of the Aegean (Greece)

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

Chin-Tser Huang
Dept. of Computer Science & Engr, Univ of South Carolina (USA)

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

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

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

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

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

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

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

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

Gregorio Martinez
University of Murcia (UMU) (Spain)

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

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

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

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

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

Chuan Qin
University of Shanghai for Science and Technology (China)

Yanli Ren
School of Commun. & Infor. Engineering, Shanghai University (China)

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

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

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

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

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

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

Ming Zhao
School of Computer Science, Yangtze University (China)

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

Yan Zhang
Wireless Communications Laboratory, NICT (Singapore)

PUBLISHING OFFICE

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

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

PUBLISHER: Candy C. H. Lin

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

Volume: 23, No: 6 (November 1, 2021)

International Journal of Network Security

The Study on the Key Management and Billing for Wireless Sensor Networks

Eko Fajar Cahyadi[1,2], Cheng-Ying Yang[3], Nan-I Wu[4], and Min-Shiang Hwang[1,5]
(Corresponding author: Min-Shiang Hwang)

Department of Computer Science and Information Engineering, Asia University[1]
No. 500, Lioufeng Rd., Taichung 41354, Taiwan (R.O.C.)
Faculty of Telecommunication and Electrical Engineering, Institut Teknologi Telkom Purwokerto[2]
Jl. D. I. Panjaitan, No. 128, Purwokerto 53147, Indonesia
Department of Computer Science, University of Taipei[3]
Taipei 10048, Taiwan (R.O.C.)
Department of Digital Multimedia, Lee-Ming Institute of Technology[4]
No.2-2, Lijhuan Rd., Taishan Township, Taipei County 243, Taiwan, ROC
Department of Medical Research, China Medical University Hospital, China Medical University[5]
No. 91, Xueshi Rd., Taichung 40402, Taiwan (R.O.C.)
Email: mshwang@asia.edu.tw

(Invited Paper; First Online Oct. 31, 2021)

Abstract

The sensing items of the wireless sensor network include sound waves, electromagnetic waves, positioning, temperature, air pressure, humidity, light brightness, speed measuring devices, etc. The sensors are mainly used in difficult-to-care environments—for example, military environment, forest fire sensing, and seawater quality sensing. Nowadays, sensing devices are more used in household appliances or highways. As a result, wireless sensor networks are becoming more popular and gradually becoming an indispensable infrastructure in urban and rural areas. However, the sensor networks must overcome many security issues in wireless sensor networks. For example, an attacker can easily eavesdrop on sensor data, invade legitimate sensors, and even spread malicious sensor information in an attempt to influence data center decisions. Therefore, to resist various attacks, many scholars are devoted to the research of related information security. However, the sensor has a primary microprocessor, scarce memory, and limited power. Therefore, asymmetric fundamental encryption mechanisms or trusted third-party mechanisms are generally not suitable for use in sensor networks. In other words, the sensor networks must use a lightweight group key replacement. This article aims to propose security and pricing strategies in wireless sensor networks to solve the above problems. First, we propose a series of mechanisms for group key creation and replacement. By using group key replacement, intrusive sensors can be excluded, thereby protecting legitimate sensors. Second, we suggest using polynomial operations and hierarchical essential management techniques in wireless sensor networks. Thus, let the future wireless sensor network system be safe, effective, and practical. In addition, we also proposed directly performing user identification between sensors; There is no need to temporarily store relevant certificate information and pass it to the authentication node. Finally, through an appropriate pricing mechanism, prevent hidden leakage on the user side, and build a complete information security mechanism.

Keywords: Authentication; Billing; Key Management; Wireless Sensor Networks

1 Introduction

A wireless sensor network (WSN) can consist of tens to hundreds or thousands of sensor nodes. Each sensor node can communicate with other nodes, data processing capabilities, and perceive the surrounding environment. Therefore, each node can transmit the sensed data back to the master node. And unified by this master node, sort out the required information. However, due to the limited hardware of the sensor, the sensor has only low computing power and limited memory and power. Generally speaking, the monitoring data and queries of most WSNs applications are collected and released by the base station or gateway nodes in the network. To improve the direct accessibility of real-time data, users must be able to directly access the required monitoring data from the sensor network without the help of gateway nodes. Therefore, a multi-user authentication mechanism that a single sensor can directly execute before users can access data is essential.

When the message is transmitted in the nodes of the wireless sensor network, the threats existing in the traditional wireless network (such as eavesdropping, modification, re-transmission) and other attacks will also threaten the wireless sensor network. However, due to the limitations of sensor hardware, it isn't easy to directly apply and implement the security architecture of general wireless networks. In the wireless sensor network, each sensor node is controlled by the cluster head. If each Cluster shares a key, when a node in the Cluster is stolen, it will suspect the private information sent by other nodes in the Cluster of leaking. In addition, even legitimate users may leak their login information, causing illegal external users to obtain the private information they want.

The purpose of this article is to propose to develop key management and multi-user authentication mechanism with perfect forward security. And the proposed scheme combined them with the payment mechanism to prevent legitimate internal users from espionage attacks and protect the privacy information of these sensor nodes from being leaked. Therefore, this research is suitable for the lightweight security architecture of the wireless sensor network environment. Different security requirements are added to deal with various possible security threats and meet the basic requirements in the related wireless sensor network applications.

The purpose of this research is to propose security and pricing strategies in wireless sensor networks. This research has the following three research motivations: 1). Research on multi-user authentication and key management; 2). Research and implementation of a wireless sensor network environment with a secure pricing mechanism; 3). Group key update and avoid malicious in wireless sensor network we will explain the research of message injection in the following subsections.

This paper is organized as follows. In Sections 2 to 6, we will propose five research issues on the key management and billing for wireless sensor networks. Finally, a conclusion is conducted in Section 7.

2 Topic 1: Research on Multi-user Authentication for WSNs

In a multi-user environment, the number of simultaneous users is dynamic. If the authentication node authenticates the user, they must temporarily store the request packet and send it to the authentication node. However, the storage space of sensor nodes is limited. Therefore, once too many data packets are requested simultaneously, it will cause data overload. So that legitimate users cannot get the sensor's response and requested information. It can also lead to security holes in DOS attacks. Therefore, user authentication is performed directly by the sensor node that receives the request packet in this research. It also reduces the number of calculations generated during the authentication process to ensure that the node can respond to user requests in real-time. The two-factor authentication analysis of passwords and smart cards is based on a one-way hash function, which is highly efficient. Therefore, it is suitable for wireless sensor networks with limited resources. However, in the environment of paid services, users are most likely to leak the information in their smart cards to other illegal users for their benefit, resulting in losses for the service provider. Therefore, this research solves the problem of users' information leakage by introducing the concept of pricing strategy. Below, we will separately introduce the research results of previous scholars.

2.1 Related Works on Multi-user Authentication for WSNs

Perrig *et al.* [24] proposed a protocol called μ TESLA. Their scheme is able to provide authentication and some level of security. Since these security protocols use source routing, they are highly vulnerable to traffic analysis during transmission [18]. The reason is that every time a sensor node receives a particular broadcast message, it will temporarily store this message, and it will not be able to authenticate the message until the following authentication key is received. Therefore, this protocol has a severe DOS attack [4]. Attackers can arbitrarily send a large number of messages to the sensor node, causing the sensor node to store these large numbers of messages within a certain period. Due to the limited storage capacity of the sensor node, this will cause other regular broadcast data packets to be unable to be authenticated by the sensor node. Later, Liu proposed some methods to improve the above shortcomings [16]. The problems solved by this method include: (1). Improve the problem that the keychain will be too long when used for a long time. (2). Improve the fault tolerance rate. (3). Be able to tolerate and resist DOS attacks.

But these schemes still have challenging implementation problems. Moreover, it is still inseparable from the basic structure of the Key Chain. Therefore, this research did not use this Key Chain method to solve the multicast authentication problem.

Benenson *et al.* first proposed a multi-user broadcasting technology based on a public key cryptosystem in 2005 [3], but this scheme has big problems. The reason is that the user's public key certificate still needs to be sent. This method has two disadvantages: (1). Each sensor node must verify the public key certificate. (2). After the public key certificate is authenticated, the signature of message M needs to be authenticated. In 2007, Jiang *et al.* proposed a new method called SC-PKC to improve this problem [12]. This method has the following characteristics: (1). Each sensor node must have its public key and private key pair. (2). They integrated an asymmetrical cryptographic system. Although the symmetric key cryptosystem is more effective than the asymmetric cryptosystem, the amount of calculation is concerned. But at the same time, it also has two shortcomings: difficulty to manage and lack of scalability. In wireless sensor networks, easy management and good scalability are very important. Therefore, in the multi-user Broadcast Authentication (BA) architecture, a public-key cryptosystem can solve this problem.

Among public-key cryptosystems, elliptic curve cryptosystems have attracted much attention in recent years. It is because some of its characteristics make it very suitable for applications in wireless sensor networks. We list several main reasons as follows: (1). With the same security strength as RSA, the key length is shorter. (2). Reduce computational complexity. Based on the above characteristics, elliptic curve cryptosystems are more and more widely used in sensor nodes. The first user authentication protocol based on elliptic curve cryptography (ECC) for WSNs was proposed by Yeh *et al.* [28]. However, their protocol does not have mutual authentication between the user and the sensor node [25]. Ren *et al.* proposed a hybrid broadcast authentication scheme based on ECC. We enumerate the research of scholars in this area in recent years. For example, Ren *et al.* proposed three user broadcast authentication mechanisms [25]:

1) Certificate-based authentication scheme: This method requires the user's public-key and private-key pair and public-key certificate;

2) Establish an ID-based authentication scheme;

3) Plant an identity verification scheme based on Merkle Hash Tree. The following only introduces the authentication scheme based on Merkle Hash Tree.

Ren *et al.* scheme [25] is more effective than the previous scheme in terms of calculation, but this scheme still has excellent shortcomings. For example, the scheme must update the Merkle hash tree once a new user joins or an old user is removed. It may also affect other users who have joined to obtain new AAI from CA, which is impractical. In 2013, Kheradmand proposed an enhanced energy-saving WSN by improving e Elliptic Curve Digital Signature Algorithm (ECDSA) [14]. The researchers pointed out the need to reduce the verification process by using cooperation between sensor nodes.

Zhong *et al.* proposed an improved elliptic curve digital signature scheme for use on WSNs by optimizing the signature generation module of ECDSA [33]. However, they were unable to reduce the number of point additions and point multiplication in the verification algorithm. To overcome the challenges in efficient remote monitoring, Sharavan *et al.* proposed a privacy preservation secure cross-layer protocol design for Wireless Body Area Networks (WBAN) using ECDSA [26]. However, ECDSA has been found not to be suitable for design of authentication protocol.

In 2020, Kasyoka *et al.* proposed an efficient pairing-free Broadcast Authentication (BA) scheme with message recovery based on a lightweight digital signature protocol for WSNs [13]. Their BA scheme can accelerate the authentication of WSN broadcast messages while providing user anonymity. Huang *et al.* proposes a scalable broadcast authentication scheme called DH-TESLA [11]. Their scheme achieves the infinite life cycle of the hash chain, security certification, scalability, and Strong password tolerance for message loss. Their protocol consists of a self-reinitialized hash chain scheme based on -threshold and an authentication scheme based on -left-counting-Bloom-filter.

2.2 The Proposed Multi-user Authentication for WSNs

First, in researching multi-user broadcast authentication technology, we will study its possible security issues. We use public-key cryptography to solve multi-user broadcast authentication. When a new user joins or leaves, the BS only needs to recalculate the authentication information required by each sensor node and broadcast it. This research does not require that each sensor node hold its public and private keys; only a piece of authentication information is needed. The main key parts of this research are described as follows:

1) Generate the private key of each new user: For each new user, the BS will calculate a user's private key. First, and through a secure channel to the user (see Figure 1). After that, the user can use the private key obtained by the user to sign the message to be sent and then send it to the wireless sensor network. Because the principal amount of calculation falls on the resource-rich BS, therefore, there is no impact on sensor nodes with low computing power.

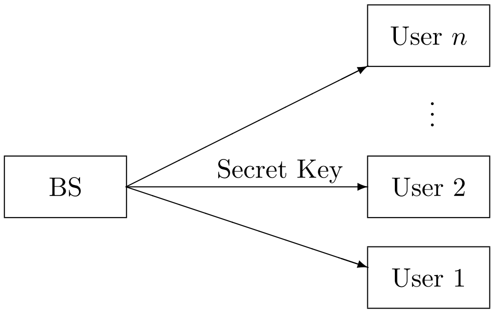


Figure 1: The private key generation

2) The authentication information generation of the sensor node: Once the user joins or leaves, the BS will generate its corresponding authentication information to the sensor node. The sensor node can authenticate the message signed by the user through the authentication information. If the authentication fails, it means that this is a malicious user. Therefore, you can refuse to provide services to this user (as shown in Figure 2).

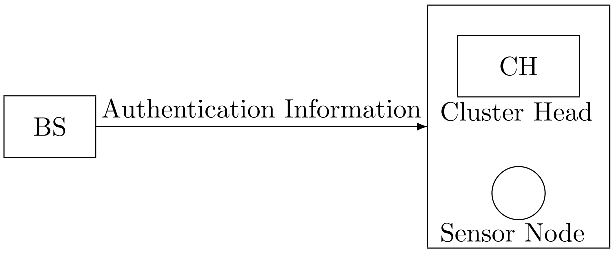


Figure 2: The authentication information generation

3) Dismissal of malicious users: Because the attacker may capture the user's mobile device, the BS must effectively deal with this problem. Compared with previous studies by scholars, sensor nodes usually must record these rejected user IDs. Once the number of rejected users is enormous, the storage capacity of sensor nodes may be severely affected.

4) Prevent DOS attacks: This research can also resist the severe DOS attacks that traditional broadcasting schemes suffer because this research scheme does not require sensor nodes to store any data packets before the following period. Each sensor node can immediately verify the user's signature.

3 Topic 2: Research on Key Management for WSNs

In the research of key management mechanisms, the key management of wireless sensor networks has always been an important and hot research topic. The most intuitive and worst method is that all sensor nodes share the same key and use these keys to encrypt and decrypt messages. However, this will be a big problem. Once the attacker captures the node or knows the key, this mechanism will completely disintegrate.

3.1 Related Works on Key Management for WSNs

Next, we briefly explain the key management methods proposed by the predecessors and explain their shortcomings. For example, Eschenauer *et al.* proposed to use the key pool method to distribute keys in 2002 [?]. First, generate a large Key Pool. There may be tens of thousands to hundreds of thousands of keys, and then each sensor node randomly takes out m keys. After deployment, any two neighboring nodes can find out whether they share a standard key. The subsequent communication will try to ensure the security of the communication channel through the public key. If a sensor node finds that it does not share a public key with neighbor nodes, it can establish a secure channel through a node that has established a secure channel. This method has some disadvantages: (1). There is a certain probability that there is no public key between the two nodes. (2). There may be multiple links using the same key. (3) Insufficient ductility.

Although this simplest method has the above disadvantages, however, in terms of minimum security requirements, this is a good solution. After that, many scholars have adopted this method to solve the key distribution method. Later, Chan *et al.* proposed improved methods for the above scheme in 2003 [7]. Eschenauer *et al.*'s method only requires two neighbor nodes to share a key. Chan *et al.* proposed a method to extend the shared key to q. Although the security has improved, once the captured sensor nodes increase above the critical number, the security will be lower than that of Eschenauer *et al.*. In addition, Chan *et al.* also adopted a 2-hop method to increase the communication range of sensor nodes. Therefore, each sensor node can communicate with its neighbor nodes and expand the passable range. The disadvantage is that the message must be forwarded. Therefore, power consumption and communication volume may increase some costs. The main method of some previous studies [9, 17] is to load some private information to the sensor node before the sensor node is deployed. After deployment, sensor nodes can exchange certain information to create keys. One characteristic of these practices is that they create keys through probability. This method has a feature called λ-secure. This means that the scheme is safe before capturing more than λ sensor nodes. This study is also superior to the above methods in terms of safety. In [17], its method is similar to [9]. The main difference is that these authors used polynomial methods. It uses the permutability of polynomials to create keys. For example, $f(x, y) = f(y, x)$. In terms of security, this method is similar to [9].

The methods described above are all based on probability to create keys. In other words, there is a certain probability that a public key cannot be established between any two adjacent nodes. In addition, Zhou *et al.* also proposed using polynomials or matrices to establish keys in [34]. More specifically, in [34], the author uses a subordinate method to allow two sensor nodes to communicate securely. Therefore, research in recent years has mainly focused on how any two adjacent nodes must establish a public key.

A brief description is as follows: In [31], the author uses elliptic curve cryptography to allow any two adjacent sensor nodes to establish a key. They use bilinear mapping technology to solve this problem as long as the sensing node has the correct private key signed by the CA. We can create a key by exchanging ID and address. The safety architecture of these elliptic curves is based on the discrete logarithm problem of elliptic curves. In other words, once the attacker knows the public key, he/she cannot guess the private key information from the obtained public key. For example: $P_{pub} = kP$. Once the attackers know P_{pub}, they cannot know the private key k correctly. In [35], the author assumes that the attacker cannot steal the sensor node within T_{min}. Once this assumption is made, any two neighboring nodes can also establish a key through a polynomial. In addition, key management under hierarchical (heterogeneous) wireless sensor network architecture is also a hot research topic.

In 2013, Bechkit *et al.* proposed a new key scheme based on a hash chain for WSN, which uses a hash chain to generate new keys [2]. The scheme also requires predistribution of keys to nodes before deployment. After deployment, use the hash function stored in the node to generate a new key. In 2015, Zhang-Wang proposed a secure and efficient hierarchical key management scheme which uses the concept of auxiliary nodes [30]. Their scheme relies on the Diffie Hellman key algorithm, but it is not scalable because the key calculation requires more. The inspiration for this scheme comes from giving less computation and overhead. Messai-Seba proposed EAHKM for cluster sensor networks, a hierarchical key scheme for cluster sensor networks [21]. This scheme is also energy-efficient, but it is only suitable for hierarchical networks and does not provide pairing keys between nodes. The cluster head shares the secret key with the member nodes but does not share the secret key with other cluster heads.

In 2018, Mehmood *et al.* proposed a session key agreement scheme to regenerate keys for healthcare applications in WBAN [19]. The main contribution of this solution is that it has been tested in heterogeneous WSN applications, and homogeneous WSN can be tested to obtain better results. However, if a mathematical model is used to verify the proposed scheme, a robust model will be given. Ali em et al. proposed a WSN encryption scheme based on Diffie Hellman [1]. The improved Diffie-Hellman method is used for safe and efficient key generation to prevent man-in-the-middle attacks. This method analyzes the security operation, key generation and calculation time, and effective results of various data packets. However, the proposed scheme has a higher data response time or calculation time.

In 2021, Kumar *et al.* proposed a scalable and storage-efficient key management scheme for wireless sensor networks [15]. They established three types of keys for the network: a network key shared by all nodes, a cluster key shared by the cluster, and a paired key for each pair of nodes. Although SSEKMS is a dynamic key management system, it also supports including new nodes and refreshing keys as needed. Mehmood *et al.* proposed a secure hybrid session key management scheme for WSN [20]. Their scheme minimizes the basic steps of public-key cryptography, and most of the operations are based on symmetric-key cryptography. Their scheme is energy-efficient and provides an effective platform for protecting key protocols and management in the WSN environment.

3.2 The Proposed Key Management for WSNs

In the key management mechanism of heterogeneous wireless sensor networks, this research still uses a symmetric cryptographic system to encrypt the information that needs to be transmitted. Therefore, it has a good effect on the speed of message encryption and decryption. For each cluster, the work required by CH is to allocate and generate keys for the members under its jurisdiction. For newly joined members or rejected members, all subsequent operations that need to be performed are performed by CH. We describe the primary key parts of this research as follows:

1) The generation of the private key: Each sensor node stores some private information before joining the sensor network. Among these sensors, the most important private information is the seed. A fair BS produces this seed. Each sensor node has a different private seed. In other words, for each sensor node, it does not know the private seeds hidden in other sensor nodes. In addition, for each CH, the BS will also generate a private seed for it (as shown in Figure 3).

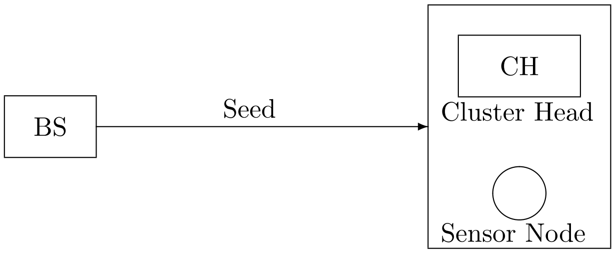


Figure 3: The private seed generation

2) Generation of communication key: Because our research goal is to have a perfect forward safety function, the key used to encrypt the message is randomly generated and discarded after being encrypted once. Other nodes that receive the message have a way to take out this random key. Therefore, with this feature, this research will have perfect front throw safety features.

3) Clustering: In hierarchical wireless sensor networks, how to group has always been a fundamental research goal. Therefore, it has been the object of research by many experts and scholars. For example, use the sensor node's transmission range, the sensor node's performance, *etc.*. This research has developed a grouping mechanism that can appropriately allocate newly added wireless sensor nodes to the jurisdiction of certain CHs (as shown in Figure 4).

4) The sensor node joins problem: For each new sensor node added, we will use the private seed owned by the user for authentication. Once the authentication is passed, these newly added sensor nodes can join the cluster they belong to. After that, data collection and data transmission are carried out.

5) Sensor node rejection problem: Sensor nodes may be captured by attackers or stolen private information held by them. In this case, the sensor node needs to be rejected. The mechanism used in this study

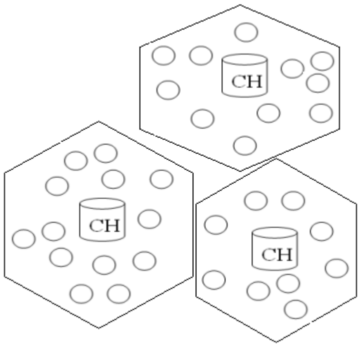


Figure 4: The wireless sensor network grouping

allows CH to reject specific sensor nodes captured by the enemy. And try to reduce the cost of rekeying required by other normal sensing nodes.

6) CH rejected the question: This study can also deal with simple CH rejection problems because the private information sent by the BS to each CH member is not in the member's hands. Therefore, the method will reject the captured CH if the attacker has captured the CH of some members. These sensor nodes will be able to be reassigned to the new cluster. Therefore, our method can use these sensor nodes rationally to avoid waste of resources.

4 Topic 3: Research on Secure Pricing Mechanisms for WSNs

In this research, we propose a pricing method based on the number of logins and apply it to the two-factor user authentication framework proposed by Das. Since this research incorporates pricing strategies into the wireless sensor network environment, it will prevent legitimate members from sharing their login information with other non-registered users for multiple logins. Furthermore, because the pricing method proposed in this research is based on service usage fees, it depends on the number of times the user logs in. Therefore, if a legitimate user deliberately leaks confidential information to other unregistered illegal users, the legitimate user and the illegal user will need to pay a service fee to log in to the WSN system. Thus, it will solve the security problems encountered by the Das protocol. Based on the existing research literature, the method proposed in this study is the first to provide a secure pricing mechanism in the wireless sensor network environment. Next, we will introduce the research results of Das [8] and Ou *et al.* [22,23].

4.1 Related Works on Secure Pricing Mechanisms for WSNs

Das proposed a two-factor authentication protocol in a wireless sensor network environment with a secure pricing mechanism [8]. Das's authentication protocol is divided into the following three phases:

1) Registration phase: When a user wants to access WSN data, the user must first register with GW-node. Therefore, the user first selects his/her password PW_i and user ID_i. Then send (ID_i, PW_i) to GW-node through a secure communication channel. When GW-node receives the registration request, GW-node calculates $N_i = H(ID_i||PW_i) \oplus H(K)$. Then, GW-node stores the following information in the smart card, including $H(\cdot)$, ID_i, N_i, $H(PW_i)$, and xa. Finally, the smart card is sent to the user through a secure method.

2) Login phase: When the user wants to log in to the network, he/she needs to insert his smart card into the card reader and enter ID_i and PW_i. Then, the smart card first checks whether the parameters entered by the user match the information stored in the card. If so, the smart card will perform the following steps:

   S1: Smart card calculation $DID_i = H(ID_i||PW_i) \oplus H(xa||T)$. Here T is the current timestamp of the user system.

   S2: The smart card calculates $C_i = H(N_i||xa||T)$, and sends a login message $\{DID_i, C_i, T\}$ to the GW node.

3) Authentication phase: When the current timestamp of the GW-node is T^*, the login message $\{DID_i, C_i, T\}$ sent by the user side is received. GW-node will first check the validity of the timestamp. If $(T^* - T)$ is greater than the legal transmission tolerance interval, to avoid retransmission attacks, GW-node will reject the user's login request. Conversely, if the login message passes the timestamp verification, GW-node performs the following steps:

   S1: Calculate $H(ID_i||PW_i)^* = DID_i \oplus H(xa||T)$ and calculate $C_i^* = ((H(ID_i||PW_i)^* \oplus H(K))||xa||T)$. If $C_i^* = C_i$, GW-node will accept the user's login request, otherwise it will reject it.

   S2: GW-node calculates $A_i = H(DID_i||S_n||xa||T')$ and sends it to the sensor node area that the user wants to query. Here T' is the timestamp of the current system of GW-node, and S_n is some sensor nodes near the user.

   S3: When S_n receives the message sent by GW-node, S_n will first verify the validity of the time stamp T' and calculate $A_i' = H(DID_i||S_n||xa||T')$. If $A_i' = A_i$ and the timestamp are correct, S_n will respond to the user's query request, otherwise, it will refuse to respond and terminate the connection.

However, it is known from the dual authentication protocol proposed by Das [8]. When a legitimate user leaks the secret information N_i and xa stored in his/her smart card to other unregistered illegal users, the illegal user can use the secret information to log in to the WSN system unlimited times. In this way, a log in ID will be logged in by multiple users simultaneously, and GW-node cannot detect and know the threat. The shortcomings are described in the following steps:

S1: Suppose that after the illegal user knows the secret information N_i and xa of the legal user, he/she can generate a legal timestamp T^* and calculate the legal login message $\{DID'_i = H(ID_i||PW_i) \oplus H(xa||T'), C'_i = H(N_i||xa||T'), T'\}$, then illegal users can send login information to the GW node.

S2: When GW-node receives a login message $\{DID'_i, C'_i, T'\}$ from an illegal user, GW-node will calculate $C^*_i = H((H(ID_i||PW_i)^* \oplus H(K))||xa||T')$ and verify that C^*_i is equal to C'_i. If it matches, GW-node will accept login requests from illegal users.

According to the above attack steps, the disadvantage of the method proposed by Das [8] is that GW-node does not store the passwords of legitimate users. Therefore, $H(ID_i||PW_i)$ information does not have any user authentication function for GW-node. The authentication feature of the Das protocol is to check whether the user has the two secret information $H(xa||T')$ and $H(K)$. Although the illegal user does not know the password of the original legal user $H(ID_i||PW_i)$, it is from the end N_i can know the secret information of $H(K)$. Therefore, illegal users can successfully pass GW-unlimited times under our proposed attack without knowing the user's password.

Because of this, this research proposes a secure pricing mechanism based on a two-factor user authentication protocol. Therefore, in addition to preventing legitimate users from deliberately leaking confidential information to other illegal users due to multiple logins, users can also charge users for the number of times they log in and access the WSN system. At present, in terms of pricing strategy, Ou *et al.* [22, 23] proposed three pricing strategies in UMTS (Universal Mobile Telecommunication Systems). They are prepaid services, real-time payment services, and login sessions. In the case of the prepaid method, the user first stores value or prepays some amount to use a specific service. In addition, users can make online or real-time payments in real-time payment methods while enjoying the service. Finally, in the payment method for the number of logins, the user pays for the services enjoyed during each login.

4.2 The Proposed Secure Pricing Mechanism in WSNs

Aiming at the research and realization of the security pricing mechanism, we have developed a set of safe and efficient encryption technology to improve the security of wireless network transmission. The calculation/storage of the intelligent card information required for the registration process needs to be carefully analyzed and redesigned. Thus, prevent secret information from being calculated and derived by illegal users. In addition, these login request messages need to go through a specific encryption step before being transmitted over the public network to prevent the plaintext from leaking. In addition, wireless sensor network nodes will also dynamically establish session keys to improve communication security. Finally, to avoid fraudulent service attacks by illegal users, add an appropriate mutual authentication and login frequency mechanism in the authentication process. The overall architecture diagram is shown in Figure 5.

It can be seen from Figure 5 that when the user wants to use the wireless sensor network service, there will be two-way mutual authentication between gateway nodes. In addition, the gateway node records the user's login information, and then the gateway node can charge the user based on the login record. To achieve mutual authentication between users and gateway nodes, we use a challenge-response mechanism in the communication process to avoid those mentioned above common passive/active attacks, such as retransmission attacks and impersonation attacks. In addition, the user and sensor can establish a dynamic session key between the user and the sensor node in each session to ensure the security of the communication between the user and the sensor node. Finally, we can add a random number to the dual authentication protocol to prevent illegal users from obtaining plaintext messages through guessing and attacks. It is also because the one-way hash function disrupts the messages sent by both parties. Therefore, except for the certification of both parties, no illegal third party may be derived.

5 Topic 4: Research on Group Rekeying for WSNs

We have designed three parts: sensor registration system, sensor layout system, and key agreement, as well as the improvement of update efficiency. The sensor registration system mainly establishes a suitable mechanism for the distribution and management of gold alloys and the loading of related parameters. When a new sensor needs to be added to the network, relevant parameters can be effectively configured to facilitate the establishment of a key in the future. In addition, the sensor layout system is mainly hoped to establish an evaluation system to analyze which distribution method needs to be adopted in a particular environment.

5.1 Related Works on Group Rekeying for WSNs

Due to the low cost and simple structure of the sensor, it is widely used in various unattended environments. However, since the sensor is often used in a malicious

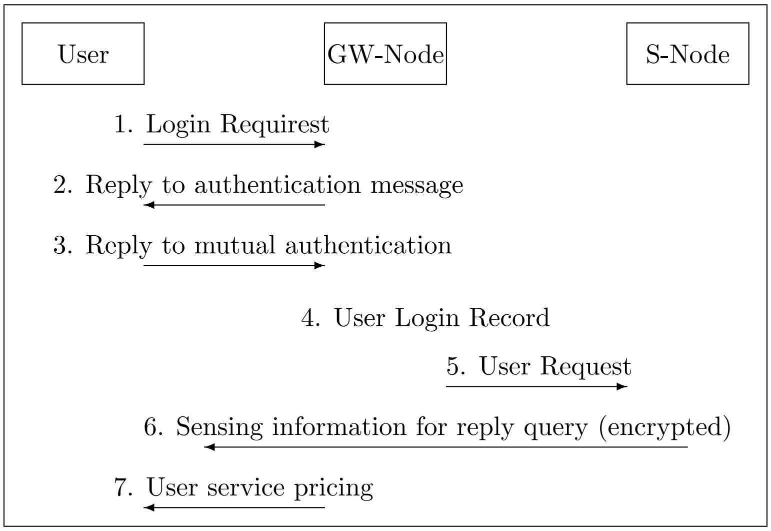

User
GW-Node
S-Node
1. Login Requirest
2. Reply to authentication message
3. Reply to mutual authentication
4. User Login Record
5. User Request
6. Sensing information for reply query (encrypted)
7. User service pricing

Figure 5: Two-factor user authentication protocol applied to wireless sensor network

environment, the sensor data may be tampered with, discarded, or even directly damaged by the sensor during the transmission process to gain access. Therefore, there must be an excellent key establishment mechanism to protect data transmission. In wireless sensor networks, some scholars divide key agreements into three categories, namely trusted third parties, public-key encryption, and key pre-distribution [10,24]:

1) Trusted server: SPINS [24] is one of the most famous security architectures. In this architecture, if any two sensors want to negotiate a key, they must communicate with a trusted third party in advance. Trusted third-party certification can ensure that the sensor can safely transmit data. Unfortunately, sensors with low energy capacity cannot bear the burden of this mechanism. Since long-distance data transmission consumes a lot of power, the battery life of the sensor was tested. Therefore, this protocol is usually only used in data centers for remote control of sensors. Since the packet size of the control data will not be too large at this time, it is well coordinated with the multicast.

2) Public key infrastructure: RSA encryption algorithm and elliptic curve cryptosystems are some of the most famous public key cryptosystems [5]. Although the security of such mechanisms has proven to be very robust, they are also widely used in general computer security. However, the complicated index calculation still makes this method unsuitable for wireless sensor networks.

3) Pre-distribution key [10]: This mechanism will preload a unique key into each sensor before the sensor is distributed to the environment. Therefore, a large amount of exponential calculation or remote transmission is not required to achieve the purpose of the key agreement. Many researchers have devoted themselves to research in this area to provide an excellent key protocol for using sensor networks or devices that require low energy consumption.

In wireless sensor networks, we pay special attention to an attack called a node capture attack. Under such an attack, a legitimate sensor will become a malicious offensive node and become the hacker's best assistant. Hackers already know the key, and various attacks can be carried out. For example, an attacker can modify the perceived data and send it to the data center to make wrong decisions because of the wrong data. Therefore, scholars began to develop related detection mechanisms. However, after malicious nodes are discovered, a lightweight and secure key update mechanism is still needed to allow legitimate sensors to exclude malicious nodes, thereby protecting the security of the entire wireless sensor network. Because the group rekeying mechanism proposed in the current research is still too complicated to be used in wireless sensor networks, this research proposes a solution that allows sensors to protect data without spending a lot of electricity.

Next, we introduce several well-known group rekeying mechanisms. In 1998, Blundo and other scholars proposed a symmetric polynomial method to establish a key [6]. A good security architecture can be established at a lower computational cost than the public key encryption mechanism through simple polynomial substitution. Blundo *et al.* proposed two methods of interactive key distribution and non-session key distribution [6]:

1) Noninteractive Polynomial-based Key Predistribution Scheme:

Key distribution phase: First, the data center will generate a symmetric multivariate t-order polynomial, which will be generated according to the following mathematical characteristics:

$$\begin{aligned} f(x_1, x_2, \cdots, x_n) &= \sum_{i_1, i_2, \cdots, i_n = 0}^{t} a_{i_1, i_2, \cdots, i_n} \\ &\quad \prod_{k=1}^{n} x_k^{i_k} a_{i_1, i_2, \cdots, i_n} \\ &\quad (0 \leq i_1, i_2, \cdots, i_n \leq t) \\ f(x_1, x_2, \cdots, x_n) &= f(x_1, \cdots, x_n, x_{n-1}) \\ &= \cdots \\ &= f(x_n, x_{n-1}, \cdots, x_1) \end{aligned}$$

The variable x denotes the number of sensors. This polynomial is symmetric, and modulus calculations are added to the calculations. Finally, these secret polynomials are individually loaded into the sensor to facilitate subsequent key establishment.

Group key agreement: When sensors want to establish a key, they first send their numbers to each other. Then load the other party's number into his/her multiple secret variables, and calculate a group key. Although this method is simple to calculate, the group key can be obtained. But whenever the key needs to be updated, the data center reassigns a new symmetric polynomial to the sensor. Therefore, it can cause a lot of security issues. In other words, the transmission cost will be too large, and the energy of the sensor will be exhausted.

2) Interactive Polynomial-based Key Predistribution Scheme:

Key distribution phase: The data center will randomly generate a symmetric bivariable t-degree polynomial. And the following characteristics will be observed during the production process:

$$\begin{aligned} f(x, y) &= \sum_{i,j=0}^{t} a_{i,j} x^i y^j \quad 0 \leq i, j \leq t \\ f(x, y) &= f(y, x). \end{aligned}$$

t is the degree. In other words, the security strength of this polynomial. Naturally, the higher, the more robust, but the relatively required computing resources and storage capacity will also increase. In addition, two-variable polynomials have symmetric properties. Therefore, modulus calculation (mod) will be added during operation to increase its security. The following example illustrates:

$$f(x, y) = 4x^2y^2 + x^3y^1 + x^1y^3$$

This is a 3-degree bivariate symmetric polynomial. Next, we will introduce how the sensor uses these polynomials to generate keys. First, the data center will calculate the secret polynomial for each sensor by substituting the sensor number into the variable x. The receiver stores these secret polynomials in the memory of each sensor in the form of a matrix for key creation or update.

Group key agreement: In the group rekeying phase, first, the largest sensor becomes the key creation initiator. Therefore, the sensor will first calculate two temporary keys. Then, the sensor randomly generates a group key. After the key is generated, two public values are calculated to protect the group key. Then the expected values are transmitted to other sensors respectively. Finally, when the sensor receives it, it can calculate the key.

The interactive key pre-distribution mechanism can solve the problem that the data center needs to reload the polynomial when updating the key in the non-interactive key pre-distribution mechanism, which causes a large transmission. However, this mechanism has the problem of unbalanced resource consumption because the generation and protection of keys are performed and led by a single sensor. When the number of sensors in the network reaches a certain level, they will be compensated for premature exhaustion of resources because they cannot be loaded. Therefore, we have developed a new pre-key distribution mechanism. In addition to reducing the computational complexity, it also increases the transmission cost and storage capacity and can solve the shortcomings of the above two methods simultaneously.

5.2 The Proposed Group Rekeying Mechanism

This study will use low-complexity polynomial operations to generate keys and update. And put forward the corresponding registration and layout system. The specific methods proposed in this study are as follows:

1) The key pre-distribution phase: The data center randomly generates $(t + 1)$ variable t safety polynomials $f(x_1, x_2, \cdots, x_{t+1})$ in the finite field $GF(q)$. After the polynomial is generated, the data center loads the identification code of each sensor into the polynomial to generate a unique pre-load polynomial $f(S_{ID}, \cdots, x_{t+1})$. The coefficients of the last unique identifier code and $f(S_{ID}, \cdots, x_{t+1})$ will be loaded into the memory of each sensor, respectively.

2) Direct key establishment phase: When a sensor wants to establish a group key, each sensor broadcasts its unique identification code to the network. After the sensors share each other's unique identification code, each sensor will calculate the group

key, $GK = f(SID, SID_{v_1}, \cdots, SID_{v_n}, timestamp)$. Where v stands for adjacent nodes and n stands for the number of adjacent nodes. And load the *timestamp* to update the key. This method can reduce the transmission cost and calculation cost caused by excluding malicious nodes.

3) Sensor leaving phase: The system will activate the group rekeying mechanism to exclude the malicious node when the sensor is detected as a malicious node. At this time, each sensor will uniquely pre-load the malicious node identification code in the polynomial, $f(SID_u, \cdots, x_{t-1}, compromisedID, timestamp)$ is replaced with the variable value $f(SID_u, \cdots, x_{t-1}, x_t, timestamp)$. And when the group key is established in the future, it will be replaced with a timestamp or a constant value.

4) Sensor joining phase: When sensors enter the network, the backward security function can be realized to prevent additional sensors from snooping on past sensory information. The system must also activate the mechanism to update the key. At this time, each sensor replaces the variable value part $f(SID_u, \cdots, x_{t-1}, x_t, timestamp)$ in the unique pre-load polynomial with the identification code $f(SID_u, \cdots, x_{t-1}, addingSensorID, timestamp)$. During the key establishment process, a synchronization mechanism is used to ensure the consistency of the timestamp.

6 Topic 5: Research on Avoiding Malicious Message Injection Attacks in WSNs

The data center transmits control information to other sensors as needed and reduces transmission costs through multicast. At this time, the creation and update of the group key play an important role. However, because sensors use wireless data transmission, they are vulnerable to data eavesdropping, malicious message modification, node intrusion attacks, and message discarding attacks. We will review these attacks in the following subsection.

6.1 Related Works on Avoiding Malicious Message Injection Attacks in WSNs

The wireless sensor network security mechanism is divided into three steps to explore to resist the above attacks. First, they are false report filtering [27,31], compromised node locator [32]), and group rekeying [6,29]:

False report filtering: Attackers can perform malicious perception data injection attacks because the attacker can invade the sensor to obtain its internal secret value and then modify the content of the sensing data. Therefore, false report filtering mechanisms are used to prevent such attacks from occurring. The most famous false report filtering mechanism was proposed by Ye *et al.* [27]. The method is briefly described as follows:

1) Key assignment: In the beginning, the data center prepared a key pool full of keys. And cut it into many non-repetitive pieces. Then each sensor will randomly select a sub-block and randomly select n keys as the key for each sensor.

2) Sensor report generation: When a sensing event occurs, the sensor uses its random key to encrypt the data and create a message authentication code.

3) En-route filtering: When the data is sent back to the data center, there is a certain probability that the intermediate nodes in the path will have the same key. Therefore, the sensor data can be verified at this time. Furthermore, when data fraud is found, data packets can be discarded in advance to reduce the burden of network transmission.

4) Sink verification: If there are no nodes in the path, the data center will thoroughly verify the sensor data. This method can effectively find malicious data packets and keys but cannot accurately detect malicious nodes, and there is a risk of Denial of Service attacks. In addition, the number of sensor keys will also affect the performance and security of the entire network. Therefore, how to further reduce the number of keys is one of the focuses of this research.

Compromised node locator: After the data center discovers the malicious packet, it further discovers the malicious node, updates the node group key, and protects the security of the sensor network. Zhang *et al.* proposed a compromised node locator [32]. This method is mainly based on the fact that the amount of data received by the sensor must be equal to the amount of data transmitted.

Group rekeying: When many sensors are hacked, the secret values are also exposed. In other words, when the attacker collects enough group keys, the attacker will have a remarkable ability to deceive the data center by modifying the sensor data. Therefore, the data center must perform the group rekeying. In addition, due to the severe conditions of the sensor's low calculation and limited power. Therefore, when proposing a method to update the key, a low-complexity mechanism must be proposed to make this mechanism suitable for wireless sensor networks. Zhang *et al.* proposed a collaborative key agreement framework [29]. The method is explained in three phases below:

1) The key pre-distribution phase: In the initial phase, the data center will assign a polynomial g(x) to the memory of each sensor. The univariate polynomial of degree t is as follows:

$$g(x) = \sum_{i=0}^{t} a_i x^i$$

2) Setup phase: At this phase, a binary (t, u) degree polynomial is generated to protect the unary secret in the sensor confidential polynomial. Therefore, each sensor will use the generated binary polynomial to add the original one-variable polynomial to get $g'(x)$. The formula is as follows:

$$\begin{aligned} e_{s_i}(x, y) &= \sum_{i=0}^{t}\sum_{j=0}^{u} a_{ij} x^i y^j \\ g'(x) &= g(x) + e_{s_i}(x, s_i). \end{aligned}$$

3) Rekeying Phase: When the group key needs to be updated, the sensor needs other sensors to transmit related values to reconstruct $e(c, y)$. Here c refers to the group key of the current version.

Zhang *et al.* method can create a key with a small amount of calculation. However, this method will cost the sensor too much transmission cost. And its safety is still insufficient. Therefore, this study will improve this part. In addition, this research will propose a filtering system with robust detection of malicious packets. The system has a suitable, effective, and safe mechanism. The improvement and cooperation of the three security mechanisms of false report filtering, compromised node locator, and group rekeying ensure the accuracy and confidentiality of perceived materials.

6.2 Effectively Avoiding Malicious Message Injection Attacks

Wireless and unattended environments have caused many attacks in wireless sensor networks, such as data eavesdropping, malicious message modification, node intrusion attacks, and message discarding. To resist these attacks, the wireless sensor network security mechanism is divided into three primary directions: malicious packet filtering, malicious node location, and group rekeying. In addition to continuously improving the group rekeying mechanism, this research also studies malicious packet filtering and group rekeying. The purpose is to improve the accuracy of detecting malicious packets. Therefore, we divide it into two steps.

Malicious node location: This research is divided into the following steps to locate malicious nodes:

1) Initial phase: At this phase, parameters need to be loaded between the sensor and the data center to ensure that the sensor can drive the encryption mechanism when an event occurs and that the sensed data can be secretly transmitted to the data center. Therefore, it is necessary to pay attention to the consistency of sensor data, data fusion, and other matters that must be considered in the design.

2) Positioning value establishment phase: The positioning value is established in each sensor through the statistical estimation, irreversible function, hardware positioning system assistance, signal fingerprint recognition, and the positioning value. This value is updated regularly and needs to have unique values such as non-repudiation and uniqueness to facilitate future queries.

3) Positioning value collection phase: When the data center finds a malicious data packet in the network, the data center will request the relevant sensor to send the location value back to the data center for future verification.

4) Stage of locating malicious nodes: Find the problem node by comparing the positioning value, and then confirm the problem node by asking the neighboring nodes. Finally, a malicious node is found so that we can implement the group rekeying mechanism.

Group Rekeyig We update the group key in the following steps:

1) Initial phase: At this phase, the improved polynomial is generated by the data center. Therefore, when making, you need to consider lightweight goals and robustness.

2) The key pre-distribution phase: After the mathematical model is generated, the unique characteristics of each sensor are added to the mathematical model to generate a secret polynomial. At this time, pay attention to whether the relevant attributes are compliant.

3) Group key generation phase: Each sensor will first find nodes in the same group and generate a shared group key according to its protocol.

4) Group key update phase: When a malicious node is found and confirmed, it uses relevant information to exchange information, allowing the still legitimate sensors to recreate a brand new group key.

Specific methods are proposed below for flat wireless sensor networks and hierarchical wireless sensor networks. First of all, in the planar wireless sensor network, this research uses the concept of a coprime identification code to establish a group key. Proceed as follows:

1) Calculate the coprime ID in the sensors phase: First, we make two reasonable assumptions. The first is that when two integers are relatively prime, the two numbers cannot be divisible. The second is that the prime factorization of prime numbers is unique.

   S1: Except 1 and 2, list all prime numbers in a range. For example, if the range is from 1 to 20, the prime numbers in the range are 3, 5, 7, 11, 13, 17, and 19.

   S2: According to the number of sensors, assign a coprime identification code to each sensor from the prime numbers listed above.

2) Key pre-distribution phase: First, the data center randomly generates a binary t-safe polynomial $f(x,y)$, x and y are variables, and p and q are prime numbers. When the prime number is large enough, we will guarantee the security of this mechanism. Its polynomials and properties are as follows:

$$\begin{aligned} f(x,y) &= \sum_{i=0}^{t} a_i(xy)^i \bmod p, \\ f(x,y) &= f(y,x) \text{ if } xy = yx. \end{aligned}$$

3) After the coprime polynomial is generated, a unique preload polynomial $f(SID, y)$ is established according to the identification code of each sensor. At this time, special attention should be paid to ensure that the unique identification code of the sensor is relatively prime before loading.

4) Group key establishment phase: Each sensor will share the other's coprime identification code. Then load the polynomial through the identification code can get the same key $GK = f(SID, \prod_{i=1}^{n} SID_{v_i})$ where v represents the node of the neighboring member. And $1 \leq i \leq n$ and n is the number of adjacent member nodes.

5) Sensor leaving the phase: The key update mechanism will activate when the sensor leaves the wireless sensor network due to power exhaustion, environmental disaster, or intrusion. This method is the same as the group key establishment phase, except the sensor does not include the identification code that leaves the sensor in the calculation during the establishment process.

6) Sensor joining phase: To achieve the high scalability of the network, new sensors need to be added according to the situation after establishing the sensor network. However, to ensure backward security, when a sensor newly enters the network, a key update action is also required. This method is the same as the group key establishment phase, except that the identification code of the new sensor is included in the calculation during the establishment process.

This method has three advantages: the first coprime identification code reduces the transmission cost; the second unique secret preload polynomial can be used for identification, and the last is the storage cost. Moreover, because only the binary polynomial is used, the memory consumption will be lower than few scholars in the front. Next, this research proposes a key update mechanism of the polynomial based on subgroup verification in the hierarchical structure as follows:

1) The key pre-distribution phase: Before development, the data center first loads some confidential values into the cluster head and general sensors as follows:

   $K_{a,S_{ij}}$: The pairing key used between the cluster head and the general sensor S_{ij}.

   $K_{a,BS}$: The pairing key is used between the Cluster Head and the data center.

   K_{a,SG_j}: The pairing key used between the cluster head and the subgroup SG_j.

   GK_a^t: The group key used between the cluster head and the general sensor.

   t: Refers to the group key version used in round t.

   u: Refers to the version of the subgroup key used in round u.

   Ω_a: refers to the collection of all intrusion sensors in the cluster head.

2) Group establishment phase: The number of sensors in the cluster head is N. The system divides it into m subgroups. Each subgroup has n sensors. That is, $N = n \times m$. As shown in Figure 6.

3) Group key generation phase: A subgroup that leaves or joins a sensor must perform an authenticated key update. First, the cluster head generates the t-th group key. The formula is as follows: Where m represents the m-th subgroup, therefore, when the key is updated, the cluster head does not need to send the new key to the m-th group. Then, if the subgroup leaves or joins the sensor, the key update method based on subgroup verification must be adopted. As shown in Figure 7, the following step-by-step instructions:

   S1: Cluster Head generates the following equation so that the new key can be protected.

   S2: After the verification equation is generated, Cluster Head encrypts the following data and sends it to the subgroup using this method to update the key.

   S3: Finally, the sensor decrypts with the previous group key and then loads the equation with its private key and the nonce value to obtain a new group key.

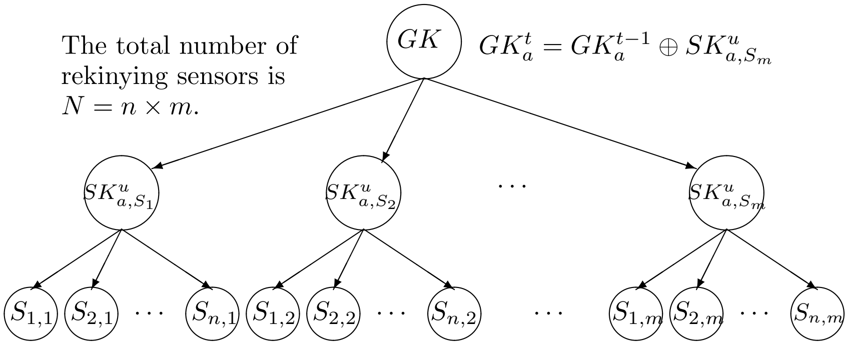


Figure 6: Sub-group architecture

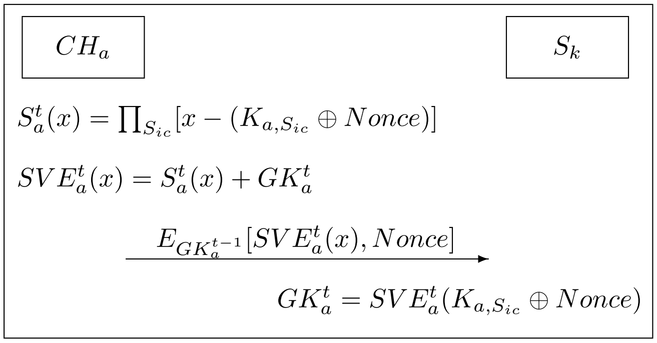

CH_a
S_k
$S_a^t(x) = \prod_{S_{ic}}[x - (K_{a,S_{ic}} \oplus Nonce)]$
$SVE_a^t(x) = S_a^t(x) + GK_a^t$
$E_{GK_a^{t-1}}[SVE_a^t(x), Nonce]$
$GK_a^t = SVE_a^t(K_{a,S_{ic}} \oplus Nonce)$

Figure 7: The key update method based on group verification

After completing the above two studies, the packet filtering system, sensor detection system, and key update system will be improved.

7 Conclusions

This research aims to develop a safe and efficient wireless sensor network system. Technologies such as secure routing protocols and image authentication are used in wireless sensor networks. Therefore, make the wireless sensor network system safe, reliable, and practical. This research also solved the key management problem and developed a secure pricing mechanism in the wireless sensor network.

This research has the following three contributions:

1) This topic studies the multi-user broadcast authentication mechanism and key management mechanism of heterogeneous wireless sensor networks. Because public-key cryptography has good scalability and convenience. We use public-key cryptography to solve multi-user broadcast authentication. When a new user joins or leaves, the BS only needs to recalculate the authentication information required by each sensor node and broadcast it. Our research does not require that every sensor node must hold its public key and private key. It only needs to save a piece of authentication information.

2) Analyze and design the smart card information that must be calculated/stored during the registration process to prevent secret information from being calculated and derived by illegal users.

3) We propose a practical and feasible polynomial-based group key update architecture. Optimize the mathematical model and create a new structure, and propose an effective group key agreement. And effectively avoid malicious message injection attacks on wireless sensor networks. Improve the three aspects of malicious packet filtering, malicious node location, and group key update. Establish a cooperative security mechanism and improve its implementation to realize the goal of the ubiquitous network.

Acknowledgments

The Ministry of Science and Technology partially supported this research, Taiwan (ROC), under contract no.: MOST 109-2221-E-468-011-MY3, MOST 108-2410-H-468-023, and MOST 108-2622-8-468-001-TM1.

References

[1] S. Ali, A. Humaria, M. S. Ramzan, *et al.*, "An efficient cryptographic technique using modified Diffie–Hellman in wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. 16, 2020.

[2] W. Bechkit, Y. Challal, and A. Bouabdallah, "A new class of hash-chain based key pre-distribution schemes for WSN," *Computer Communications*, vol. 36, no. 3, pp. 243-255, 2013.

[3] Z. Benenson, N. Gedicke, and O. Raivio, "Realizing robust user authentication in sensor networks," *Real-World Wireless Sensor Networks*, Stockholm, 2005.

[4] C. Benzaid, K. Lounis, A. Al-Nemrat, N. Badache, M. Alazad, "Fast authentication in wireless sensor networks," *Futur Generation Computer Systems*, vol. 55, pp. 362–375, 2016.

[5] I. Blake, G. Seroussi, and N. Smart, *Elliptic Curves in Cryptography*, Cambridge University Press, London Mathematical Society 265, Tech. Rep., 1999.

[6] C. Blundo, A. D. Santis, A. Herzberg, S. Kutten, U. Vaccaro, and M. Yung, "Perfectly-secure key distribution for dynamic conferences," *Lecture Notes in Computer Science*, vol. 20, no. 1, pp. 57-64, 1993.

[7] H. Chan and A. Perrig, "Security and privacy in sensor networks," *Computer*, vol. 36, no. 10, pp. 103-105, 2003.

[8] A. K. Das, "Improving identity-based random key establishment scheme for large-scale hierarchical wireless sensor networks," *International Journal of Network Security*, vol. 14, no. 1, pp. 1-21, 2012.

[9] W. Du, J. Deng, Y.S. Han, and P.K. Varshney, "A pairwise key predistribution scheme for wireless sensor networks," in *Proceedings of the 10th ACM Conference on Computer and Communications*, pp. 42–51, 2003.

[10] D. He, L. Cui, H. Huang, and M. Ma, "Design and verification of enhanced secure localization scheme in wireless sensor networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 20, no. 7, pp. 1050-1058, 2009.

[11] H. Huang, Q. Huang, F. Xiao, W. Wang, Q. Li, T. Dai, "An improved broadcast authentication protocol for wireless sensor networks based on the self-reinitializable hash chains," *Security and Communication Networks*, vol. 2020, 2020.

[12] C. Jiang, B. Li, H. Xu, C. Jiang, B. Li, and H. Xu, "An efficient scheme for user authentication in wireless sensor networks," in *Proceedings of the AINAW*, pp. 438–442, 2007.

[13] P. Kasyoka, M. Kimwele, S. M. Angolo, "Multi-user broadcast authentication scheme for wireless sensor network based on elliptic curve cryptography," *Engineering Reports*, vol. 2, no. 7, 2020.

[14] B. Kheradmand, "Enhancing energy efficiency in wireless sensor networks via improving elliptic curve digital signature algorithm," *World Apply Science Journal*, vol. 21, no. 11, pp. 1616–1620, 2013.

[15] V. Kumar, N. Malik, G. Dhiman, T. K. Lohani, "Scalable and storage efficient dynamic key management scheme for wireless sensor network," *Wireless Communications and Mobile Computing*, vol. 2021, pp. 1–11, 2021.

[16] D. Liu and P. Ning, "Multilevel μTESLA: Broadcast authentication for distributed sensor networks," *ACM Transactions on Embedded Computing Systems*, vol. 3, no. 4, pp. 800-836, 2004.

[17] D. Liu, P. Ning, and R. Li, "Establishing pairwise keys in distributed sensor networks," *ACM Transactions on Information and System Security*, vol. 8, no. 1, pp. 41-77, 2005.

[18] S. R. Maidhili, G. M. Karthik, "Energy efficient and secure multi-user broadcast authentication scheme in wireless sensor networks," in *Proceedings of the International Conference on Computer Communication and Informatics (ICCCI'18)*, 2018.

[19] G. Mehmood, M. Z. Khan, H. U. Rahman, and S. Abbas, "An efficient and secure session key establishment scheme for health-care applications in wireless body area networks," *Journal of Engineering and Applied Sciences*, vol. 37, 2018.

[20] G. Mehmood, M. S. Khan, A. Waheed, *et al.*, "An efficient and secure session key management scheme in wireless sensor network," *Complexity*, vol. 2021, pp. 1-10, 202021.

[21] M. L. Messai and H. Seba, "EAHKM+: Energy-aware secure clustering scheme in wireless sensor networks," *International Journal of High Performance Computing and Networking*, vol. 11, no. 2, pp. 145–155, 2018.

[22] H. H. Ou, M. S. Hwang, J. K. Jan, "A simple mobile communication billing system among charged parties," *Applied Mathematics and Computation*, vol. 192, no. 2, pp. 487-495, Sept. 2007.

[23] H. H. Ou, M. S. Hwang, J. K. Jan, "A provable billing protocol on the current UMTS," *Wireless Personal Communications*, vol. 55, no. 4, pp. 551-556, 2010.

[24] A. Perrig, R. Szewczyk, V. Wen, D. Culler, and J. D. Tygar, "Spins: Security protocols for sensor networks," in *Proceedings of the 7th Annual International Conference on Mobile Computing and Networking*, pp. 189-199, July 2001.

[25] K. Ren, W. Lou, K. Zeng, and P. J. Moran, "On broadcast authentication in wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. 6, no. 11, pp. 4136–4144, 2007.

[26] P. T. Sharavan, D. Sridharan, R. Kumar, "A privacy preservation secure cross layer protocol design for IoT based wireless body area networks using ECDSA framework," *Journal of Medical Systems*, vol. 42, no. 10, pp. 196, 2018.

[27] F. Ye, H. Luo, S. Lu, and L. Zhang, "Statistical en-route filtering of injected false data in sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 4, pp. 839-850, 2005.

[28] H. L. Yeh, T. H. Chen, P. C. Liu, T. H. Kim, H. W. Wei, "A secured authentication protocol for wireless sensor networks using elliptic curves cryptography," *Sensors*, vol. 11, no. 5, pp. 4767-4779, 2011.

[29] W. Zhang, S. Zhu, and G. Cao, "Predistribution and local collaboration-based group rekeying for wireless sensor networks," *Ad Hoc Networks*, vol. 7, no. 6, pp. 1229-1242, 2009.

[30] X. Zhang and J. Wang, "An efficient key management scheme in hierarchical wireless sensor networks," in *International Conference on Computing, Communication and Security*, pp. 1-7, 2015.

[31] Y. Zhang, W. Liu, Y. Fang, and D. Wu, "Secure localization and authentication in ultra-wideband sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 4, pp. 829–835, 2006.

[32] Y. Zhang, J. Yang, W. Li, L. Wang, and L. Jin, "An authentication scheme for locating compromised sen-

sor nodes in WSNs," *Journal of Network and Computer Applications*, vol. 33, no. 1, pp. 50-62, 2010.

[33] H. Zhong, R. Zhao, J. Cui, X. Jiang, J. Gao, "An improved ECDSA scheme for wireless sensor networks," *International Journal of Future Generation Communication Netwoeks*, vol. 9, no. 2, pp. 73-82, 2016.

[34] Y. Zhou, and Y. Fang, "A two-layer key establishment scheme for wireless sensor networks," *IEEE Transactions on Mobile Computing*, vol. 6, no. 9, pp. 1009–1020, Sept. 2007.

[35] S. Zhu, S. Setia and S. Jajodia, "LEAP: Efficient security mechanisms for large-scale distributed sensor networks," *ACM Transactions on Sensor Networks*, vol. 2, no. 4, pp. 500–528, 2006.

Biography

Eko Fajar Cahyadi is a lecturer in the Faculty of Telecommunication and Electrical Engineering, Institut Teknologi Telkom Purwokerto, Indonesia. He is currently pursuing a Ph.D. degree in the Department of Computer Science and Information Engineering, Asia University, Taichung, Taiwan, under the supervision of Prof. Min-Shiang Hwang. He receives the B. Eng. and M. Sc. degree in electrical engineering from Institut Sains dan Teknologi Akprind Yogyakarta in 2009, and Institut Teknologi Bandung in 2013, respectively. His research interest includes information security, VANETs, and WLANs.

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

Nan-I Wu received a Ph.D. degree in the Institute of Computer Science and Engineering from Nation Chung Hsing University (NCHU), Taichung, Taiwan, in 2009. From 2010 to 2011, he was a post-doctoral research fellow at the Academia Sinica Institute of information science. He was an assistant professor at the Department of Animation and Game Design, TOKO University (Taiwan), during 2011-2018 and an associate professor during 2018-2019. Now he is an associate professor at the Department of Digital Multimedia, Lee-Ming Institute of Technology (Taiwan) since 2019 and also the Director of the eSports Training Centre since 2020. His current research interests include game design, eSports training/magagement, multimedia processing, multimedia security, data hiding, and privacy-preserving. He published more than 10 international journal papers (SCI) and conference papers.

Min-Shiang Hwang received M.S. in industrial engineering from National Tsing Hua University, Taiwan in 1988, and a Ph.D. degree in computer and information science from National Chiao Tung University, Taiwan, in 1995. He was a distinguished professor and Chairman of the Department of Management Information Systems, NCHU, during 2003-2011. He obtained 1997, 1998, 1999, 2000, and 2001 Excellent Research Awards from the National Science Council (Taiwan). Dr. Hwang was a dean of the College of Computer Science, Asia University (AU), Taichung, Taiwan. He is currently a chair professor with the Department of Computer Science and Information Engineering, AU. His current research interests include information security, electronic commerce, database, data security, cryptography, image compression, and mobile computing. Dr. Hwang has published over 300+ articles on the above research fields in international journals.

ID-Authentication Based on PTPM and Certificateless Public-Key Cryptography in Cloud

Hui Xia[1] and Weiji Yang[2]

(Corresponding author: Weiji Yang)

School of Software, Shenyang Normal University[1]
School of Life Science, Zhejiang Chinese Medical University[2]
548 Binwen Rd, Binjiang, Hangzhou, Zhejiang, China
(Email: yangweiji@163.com)

(Received Sept. 3, 2019; Revised and Accepted Mar. 28, 2020; First Online Oct. 18, 2021)

Abstract

To solve the security problems and deficiencies in the current authentication between user and cloud, this paper applies Portable TPM chip and certificateless public key cryptography for the first time to solve the issues in the cloud environment, and a scheme for bidirectional ID authentication between user and cloud is proposed. The proposed scheme has several advantages compared with previous authentication schemes. First, based on the unique identity of user and cloud by the identity management mechanism, portable TPM can not only achieve a secure and trusted terminal platform, which ensures the authentication result between user and cloud is correct and valid but also supports the objectives of ID authentication between user and cloud in the user's any terminal device. Furthermore, Dual-factor ID authentication (password + key) is implemented with a new scheme's certificateless public-key signature algorithm. Finally, security proof and performance analysis show that this proposed scheme has the security level of EUF-CMA. The computation overhead of ID authentication between user and cloud is significantly improved.

Keywords: Certificateless Public Key Cryptography; Cloud Computing; EUF-CMA; Portable TPM

1 Introduction

Cloud computing is a new service model based on the Internet to provide resources such as storage and computing.With cloud services, enterprises, organizations and individual users can quickly and easily carry out massive data computing and data storage and sharing operations. Otherwise, users who do not apply for registration or purchase of cloud services can use the cloud services, which on the one hand to bring a huge CSP service response burden and serious economic losses, at the same time legitimate users may not receive timely response to the results of the service and the loss of stored information. At the same time, the application for the use of cloud services users need to CSP identity authentication, or hackers or malicious organizations can obtain counterfeit CSP user account and privacy and other important information, to the user serious economic losses and information disclosure threat. Therefore, the need for CSP and the use of cloud services, the identity of the user's safety certification, to ensure the legitimacy of both identity and correctness.

At the same time, cloud computing based on a variety of deployment models and service models for the large number of users can provide a variety of different types of services, which may come from different management domain, if the service-based identity authentication mechanism will inevitably lead to cumbersome authentication process [15]; in addition, the user will be in different work areas (such as enterprise internal work and external cloud work domain) at any time switch status, if each work domain to establish cloud user identity management mechanism, user identity will appear multiple, user identity will appear multiple, thus making user authentication and access very complex [10]. Therefore, compared with the traditional calculation mode, cloud environment identity authentication also need to consider the issue of cloud user identity management, through the establishment of identity management mechanism to achieve the identity of different users within the domain of information uniqueness, so as to improve the user experience and solve the problem of user identity synchronization in different domains.

In cloud environments, enterprises, organizations, and individual users can access cloud services using terminal devices, including PCs (personal computers), PDAs (personal digitalassistants), laptops and mobile phones, therefore, authentication not only involves the secure con-

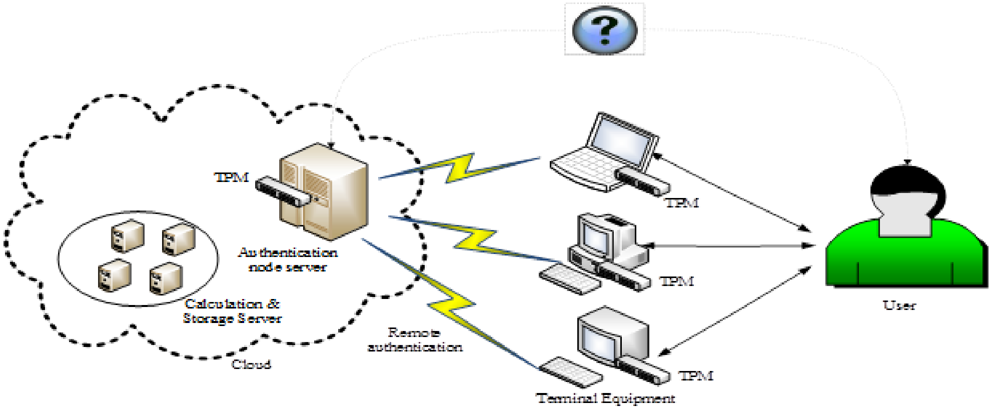


Figure 1: ID authentication based on TPM in cloud environment

nection between the cloud and the terminal device, also need to consider the user and the cloud between the security connection. This is because the user is the ultimate CSP service object, the terminal device is only the user's tool and service platform. As shown in Figure 1, both the server and the user terminal in the cloud are authenticated by the TPM (trustedplatform module) security chip to complete the remote authentication process. Although the use of TPM chip in the server and the terminal device to establish a trusted connection, but if used to implement the user authentication process, there will be security problems. This is because if the user terminal device using malicious software, then the attacker can fool the user by tampering with the authentication result, that is, the trusted path connection can not be safely extended from the terminal device to the user. In addition, Maybe I didn't make it clear. I mean that If the user try to access the encrypted data in other terminal equipment by TPM or other platform, they have to migrate all the data to the terminal. This will give users a complex operating process and even cause the user's privacy leak. Therefore, to achieve the identity between the cloud and the user authentication on the one hand the need to ensure the authenticity of the certification results. On the other hand, it is necessary to support the user to use any terminal device to complete the authentication process.

The rest of this paper is organized as follows. Section 2 of this paper describes the basic knowledge used in the program. Section 3 describes in detail the identity authentication scheme presented in this paper. Section 4 gives the security proof of the scheme. Section 5 presents a comparative analysis of the existing work and the program in this paper.

2 Related Work

In recent years, domestic and foreign scholars have done a lot of research on identity authentication under cloud environment. Document [2, 14] uses the certificate-based public-key cryptosystem to solve the problem of identity authentication between users and the cloud. Although the use of public key certificate can correctly realize the identity authentication process in the cloud environment, but the management and maintenance of public key certificates will consume huge computing resources [11]; second, the user's terminal device security is not guaranteed; in addition, cloud user identity management issues have not been effectively addressed. Document [3, 5, 8, 16] identity-based cryptography [12] presents an authentication scheme between the user and the cloud. Compared with the public key cryptosystem, the ID-based cryptosystem does not need the public key certificate, thereby solving the problem of certificate management. At the same time, Document [3,5,16] respectively through the establishment of identity management mechanism to solve the cloud environment, the uniqueness of the user's identity problems.But due to the introduction of third-party PKG (private key generator), the above scheme produces the key escrow problem [11]; at the same time, if the PKG malicious behavior, it can use any user's private key forged signatures, to deceive the purpose of the verification side; in addition, as in the document [3,5,8,16] has a problem that the terminal apparatus can not be secured and trusted.

Al-Riyami and Peterson in 2003 proposed a certificate-free public key cryptosystem [1]. It not only avoids the traditional certificate management of public key cryptosystem, but also solves the key escrow problem of identity-based cryptosystem. Therefore, compared with the traditional public key cryptosystem and the identity-based cryptosystem, the certificate-free public key cryptosystem has the advantages of high efficiency and strong security. Considering the characteristics of cloud sharing, such as sharing resources and supporting multiple access modes, while the use of a large number of users of cloud services, certificate-free public key cryptosystem is more suitable to solve the user and the cloud between the identity authentication. Document [4, 9] proposes an anonymous identity authentication scheme between the user and the cloud based on the certificateless cryptosystem. The proposed authentication process is achieved by verifying whether the hash value generated by the identity ID and the public key value of the correspondent party is correct, attackers can easily break through the middleman attack authentication process. In addition, the scheme is based on the single-factor authentication process; But also did not consider the terminal equipment platform security issues, and therefore can not guarantee the authenticity of the certification results.

In addition, according to the investigation made by Fujitsu Research Institute, 88% of users worry that their data stored in the cloud will be unauthorized access. In order to ensure the identity of the user who accesses the data, a more secure user authentication mechanism needs to be established. Therefore, as described above, aiming at the existing problems and shortcomings in realizing the authentication between user and cloud, this paper is based on PTPM (portableTPM) and certificateless public key signature algorithm, proposed a way to support the two-way identity between the cloud and the user authentication program. The specific contribution is as follows:

1) For the first time, PTPM and non-certificate public key signature algorithm are combined to solve the problem of identity authentication between users and cloud in cloud environment.

2) Based on hierarchical ID tree structure, the identity management mechanism including user and cloud is established, and the identity of any communication entity is achieved.

3) The use of PTPM platform to ensure the safety and credibility of the terminal and the cloud between the authenticity of the authenticity of the correct.

4) To achieve the cloud and the user between the "password + key" two-factor authentication process.

5) To support users to use any terminal device to complete the two-way authentication process with the cloud.

3 Related Basics

3.1 Security Theory Hypothesis

The security of this scheme is based on the difficulty of CDH (computational Diffie-Hellman), the relevant definitions are as follows.

Definition 1. *CDH problem. It is known that $a, b \xleftarrow{R} Z_g^*$. g is a generator. Given (g, g^a, g^b), calculate g^{ab}. Here $a, b \xleftarrow{R} Z_g^*$ indicates that elements a and b are selected from Z_q^* that conform to a uniform distribution.*

Definition 2. *CDH hypothesis. The probability of algorithm B solving CDH problem in probability polynomial time is $Adv_{CDH}(B) = Pr[g^{ab} \leftarrow B(g, g^a, g^b)]$. If $Adv_{CDH}(B)$ is negligible, the CDH problem is said to be difficult.*

3.2 Security Model

The identity authentication scheme proposed in this paper is based on the idea of non-certificate public key signature algorithm. Therefore, according to the security attack model defined in [1], the security of this program need to consider the following two types of adversaries.

External rival A_I: A_I represents an ordinary third-party attacker, A_I does not have a system master key, but you can replace the user's public key with an arbitrary value;

Internal rival A_{II}: A_{II} on behalf of malicious KGC (key generating center), A_{II} has the system master key, but does not allow you to replace the user's public key.

3.3 PTPM

Document mentioned the Intel Corporation in 2002, first proposed the portable TPM (portable TPM) concept. As with the TPM, PTPM also has features such as secure storage, key generation, and data signing. According to the description in [13, 17], as PTPM through the USB interface or PC card interface to communicate with the terminal device, the trust foundation of the trusted computing platform can be transferred from the platform itself to the users themselves.Each user can have their own identity PTPM, and can be used on one or more terminal devices. The PTPM hardware module implemented in [17] also has a miniature liquid crystal window, so that the calculation process can guarantee the authenticity of the results and correctness. Therefore, the use of PTPM on the one hand can build a trusted platform for the terminal chain, to achieve the integrity of the terminal platform measurement; on the other hand, the user can securely store important data such as a key in the PTPM, to achieve the purpose of the user to use any terminal device to complete the authentication.

It should be noted that, as the focus of this paper is the issue of identity authentication between cloud and user, therefore, TPM and PTPM how to ensure cloud authentication node server and user terminal platform, the credibility of the security will no longer be discussed. Authentication scheme in the back of the introduction, you can think of the cloud and the user identity authentication, the TPM and PTPM-based terminal platform integrity measurement processes have been implemented using [13, 17].

4 Identity Authentication Scheme Design

4.1 Overall Structure

As shown in Figure 2, the user holds the PTPM hardware module, cloud authentication node server embedded TPM security chip. The two-way authentication process between user and cloud consists of two phases: user registration shown in Figure 2(a) and login authentication shown in Figure 2(b).

In the registration phase, user u_i first enters the password pw, and identity ID_i and then obtains the registration request information Reg_{req} by using the PTPM calculation; After the authentication node server receives the user registration request information Reg_{req}, first, according to the identity of ID_i query user u_i is registered, and then enter the public key of KGC and u_i and use TPM calculation to verify whether the signature value generated by information such as pw and ID_i is correct. After verification is correct, the authentication node server stores the registration information of user u_i, and sends the corresponding registration response information Reg_{req} to u_i; upon receipt of the registration response message Reg_{req}, u_i. First, enter the public key of the KGC and the authentication node server and use PTPM to verify that the signature value of the authentication node server is correct. If it is correct, it outputs the registration success flag and stores information such as identity ID_{auth} and secret value of the authentication node server.

In the authentication phase, the user u_i first sends the authentication request information $Auth_{req}$ including ID_i, $H_2(ID_i||pw_i)$ and g^{r_i} to the authentication node server; after verifying that the received $H_2(ID_i||pw_i)$ value is correct, the authentication node server first calculates $HMAC_k(g^{r_i})$, where in the key k used by the HMAC operation depends on the secret information values generated by the user and the authentication node server during the registration phase, then the authentication node server sends authentication response information $Auth_{res}$ such as ID_{auth}, $HMAC_k(g^{r_i})$ and g^{r_i} to u_i. After calculating the correctness of the received $HMAC_k(g^{r_i})$ value, the authentication of the authentication node server is completed, at the same time also need to calculate $HMAC_k(g^{r_i})$ as the response information; and the authentication node server completes the authentication of the user u_i identity by verifying whether the value of $HMAC_k(g^{r_i})$ sent by u_i is correct or not, at the same time in order to allow u_i confirmed by certification, also need to send $HMAC_k(g^{r_i})^{r_j}||ID_{auth})$ value to u_i again, the HMAC value is calculated based on the random number and ID_{auth} generated by both parties before; End-user u_i in the use of PTPM verify the correctness of $HMAC_k(g^{r_i})^{r_j}||ID_{auth})$ value, output the authentication success flag to the PTPM display window. As the cloud environment, users can use any terminal device to access the use of cloud services. Thus, the single user shown in Figure 3(a) uses a plurality of terminal devices and the multi-user shown in Figure 3(b) to complete the authentication process between the user and the cloud using a terminal device.

4.2 Algorithm Description

Given the safety parameters k, a large prime number p of k bits is selected. Assume that G_1 and G_2 are the multiplicative cyclic groups of order p, g is the generator of G_1. Bilinear mapping $e : G_1 \times G_2 \rightarrow G_2$. Select the anti-collision hash function $H_1 : \{0,1\}^* \rightarrow G_1$ and $H_2 : \{0,1\}^* \rightarrow G_1$. The system exposes the global parameter $params$ to $(G_1, G_2, e, p, g, H_1, H_2)$.

4.2.1 Identity ID Generation

Based on the hierarchical ID tree structure proposed in [7], this paper defines the users in the cloud environment, cloud servers, and other role of the identity ID value. The entire hierarchical structure consists of two layers, the root node is KGC. A third party key generation center which generates a user's private key; the leaf node indicates the end user registered in the cloud and the cloud authentication node server. Obviously, all nodes in the hierarchical ID tree structure have unique names, thus achieving the user and the cloud server identity uniqueness of the goal. Suppose the identity $ID_i = DN_0||DN_i$ of user u_i, the cloud authentication node server's identity $ID_{auth} = DN_0||DN_{server}$, wherein DN_0, DN_i, DN_{server} represents KGC, u_i and $server_{auth}$ in the hierarchical ID tree structure defined in the name, $||$ represents the concatenation of strings.

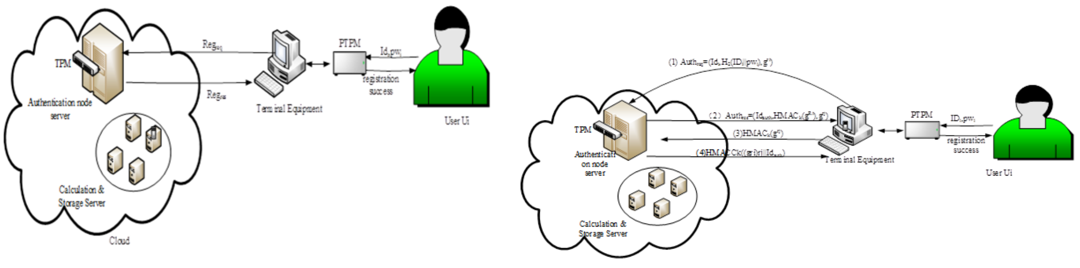
4.2.2 Key Generation

According to the idea of certificate-free public key cryptosystem. The key generation process of the leaf node $node_i$ in the scheme is as follows:

1) $node_i$ selects $x_i \overset{R}{\leftarrow} Z_p^*$ as the secret value, calculates and publicizes the public key $pk_i = g^{x_i}$.

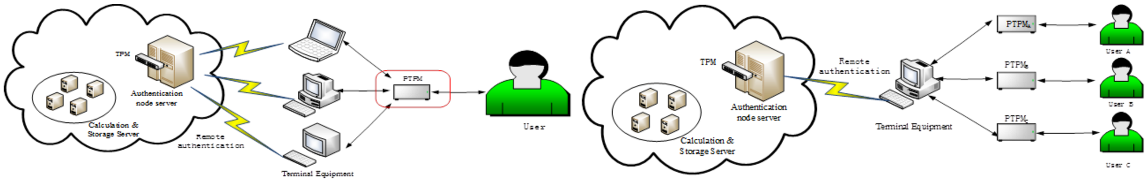
2) KGC select $S_0 \overset{R}{\leftarrow} Z_p^*$, S_0 is the master key of KGC, and public key $pk_{KGC} = g^{S_0}$ is calculated and publicized. Given each leaf node $node_i$ in the hierarchical ID structure, KGC first obtains the corresponding identity ID_i and pk_i value of $node_i$, then calculate $Q_i = H_i(ID_i)$, finally KGC uses Q_i to compute and send $(g^{x_i})^{S_0}Q^{S_0}$ to $node_i$.

3) $node_i$ first obtains the partial private key $Q_i^{S_0}$ generated by KGC by calculating $\frac{(g^{x_i})^{S_0}Q^{S_0}}{(g^{S_0})^{x_i}}$, and then obtain the identity ID_i by querying the hierarchical ID tree structure, and then obtain the identity ID_i by querying the hierarchical ID tree structure, calculate $Q_i = H_i(ID_i)$ and verify the correctness of $Q_i^{S_0} : e(Q_i^{S_0}, g) = e(Q_i, pk_{KGC})$, if they are equal, the private key $sk_i = (Q_i^{S_0}, x_i)$ is generated.


TPM
Authentication node server
Calculation & Storage Server
Cloud
Reg_req
Reg_res
Terminal Equipment
PTPM
Id_i,pw_i
registration success
User Ui
(1) Auth_req=(Id_i,H_2(ID_i||pw_i),g^r)
(2) Auth_res=(Id_auth,HMAC_k(g^ri),g^r)
(3)HMAC_k(g^r)
(4)HMACCk((gri)ri||Id_auth)
Authentication node server
Calculation & Storage Server
TPM
Terminal Equipment
PTPM
ID_i,pw_i
registration success
User Ui

(a) (b)

Figure 2: Bidirectional ID authentication between user and cloud: (a)User registration phase, (b) Login authentication phase.



(a) (b)

Figure 3: ID authentication between user and cloud in any terminal device: (a)Single user uses multiple terminal devices to complete authentication, (b) Multi-users use a terminal device to complete the certification.

In the following description, the public-private key pair of user u_i is denoted as $sk_i = (Q_i^{S_0}, x_i)$, $pk_i = g^{x_i}$; and the public-private key pair of the cloud authentication node server $server_{auth}$ is denoted by $sk_{server} = (Q_j^{S_0}, x_j)$, $pk_{server} = g^{x_j}$.

4.2.3 User Registration

1) User u_i inputs ID_i and password value pw_i, calculate $H_2(ID_i||pw_i)$ by PTPM; then select $S_i \overset{R}{\longleftarrow} Z_p^*$ to calculate g^{s_i} and $V = H_2(ID_i||pw_i)g^{S_i}$. Finally, u_i sends the registration request $Reg_{req} = (ID_i, g^{S_i}, V^{x_i}Q_i^{S_0}, H_2(ID_i||pw_i))$ to the cloud authentication node server $server_{auth}$.

2) After $server_{auth}$ receives Reg_{req}, first, according to the registered user information table $T_{register}$ to check whether ID_i exists, if not, then $Q_i' = H_i(ID_i)$, then verify the correctness of $V_{x_i}Q_i^{S_0}$ value: $e(V^{x_i}Q_i^{S_0}, g) \overset{?}{=} e(H_2(ID_i||pw_i)g^{S_i}, g^{x_i}) \cdot e(Q_i', pk_{KGC})$, if equal, $server_{auth}$ selects $S_i \overset{R}{\longleftarrow} Z_p^*$ and calculates g^{S_i} using TPM, store ID_i, S_j, g^{S_i} and $H_2(ID_i||pw_i)$ to $T_{registaer}$, $W = H_2(ID_{auth}||pw_j)g^{S_j}$ is then calculated using the TPM and the registration response message $Reg_{res} = (ID_{auth}, g^{S_j}, W^{x_j}Q_j^{S_0})$ is sent to u_i. Otherwise it returns failure flag register to u_i. If $T_{registar}$ has stored the ID_i value, it is returned to the registered mark u_i.

3) After u_i receives the registration response information Reg_{req}, firstly PTPM calculate $Q_J' = H_1(ID_{auth}$ and $H_2(ID_{auth}||pk_j)$, then $H_2(ID_{auth}||pk_j)g^{S_j}$ is calculated and the correctness of the $W^{x_j}Q_j^{S_0}$ value is verified by determining whether or not the equation $e(w^{X_j}Q_j^{S_0}, g) \overset{?}{=} e(H_2(ID_{auth}||pw_j)g^{S_j}, pk_j) \cdot e(Q_j', pk_{KGC})$ is satisfied. If equal signifies that u_i successfully registers at $server_{auth}$, PTPM outputs the registration success flag to the display window, While u_i stores ID_{auth}, S_i and g^{S_j}; otherwise, the registration failure flag is output.

4.2.4 Login Authentication

1) User u_i first inputs ID_i and PW_i, and use PTPM to calculate $H_2(ID_i||PW_i)$, at the same time select $r_i \overset{R}{\longleftarrow} Z_p^*$ and calculate g^{r_i}, the PTPM then sends the login authentication request $Auth_{req} = (ID_i, H_2(ID_i||pw_i), g^{r_i})$ to $server_{auth}$.

2) After $server_{auth}$ receives message $Auth_{req}$, first of all, according to ID_i query $T_{register}$ store the $H_2(ID_i||pw_i)$ value and the received is equal, if not, $server_{auth}$ returns the password error message to u_i; Otherwise $server_{auth}$ first obtains the corresponding S_j and g^{S_i} values of ID_i, calculate $(g^{S_i})^{S_j}$ with TPM; then select $r_i \overset{R}{\longleftarrow} Z_p^*$ and use TPM to calculate $D_i = HMAC_k(g^{r_j})$ and g^{r_j}, finally, send $Auth_{req} = (ID_{auth}, g^{r_j}, D_i)$ to u_i.

3) u_i received $server_{auth}$ returned information $Auth_{res}$, first, according to the received ID_{auth} query to obtain the corresponding S_i and g^{S_j}, then use PTPM to calculate $k' = (g^{S_j})^{S_i}$ and $D_1' = HMAC_{k'}(g^{r_t})$ respectively, verify that D_1 and D_1' are equal. If equal

signifies that u_i has completed the authentication of identity $server_{auth}$, then $D_2 = HMAC_{k'}(g^{r_j})$ is calculated using PTPM and sent to $server_{auth}$; otherwise the authentication $server_{auth}$ identity fails, u_i terminates the verification process.

4) $server_{auth}$ uses the TPM to compute $D'_2 = HMAC_k(g^{r_j})$ and compare it with D_2. If equal signifies that $server_{auth}$ has completed the authentication of identity u_i, then $server_{auth}$ computes $D_3 = HMAC_k((g^{r_i})^{r_j}||ID_{auth})$ using the TPM and sends it to u_i; otherwise the authentication u_i identity fails, $server_{auth}$ terminates the verification process.

5) u_i uses PTPM to compute $D'_3 = HMAC_{k'}((g^{r_j})^{r_i}||ID_{auth})$ and compare it with the received D_3. If they are equal, PTPM output validation success flag to the display window; otherwise, the authentication failure flag is output.

Upon completion of the above authentication process, u_i and $server_{auth}$ can use the session key $sk_{auth\leftrightarrow i} = g^{S_iS_j}g^{r_ir_j}$ to follow up the information exchange process.

4.2.5 Password Update

Suppose u_i wants to update the original password pw_i to pw'_i, then u_i first use of PTPM were calculated $H_2(ID_i||pw'_i)$, and $New_{pw} = sk_{auth\leftrightarrow i} \times H_2(ID_i||pw'_i)$, and $(g^{r_j})^{S_i}$ then sends a password update request $Update_{pw} = (ID_i, New_{pw}, (g^{r_j})^{S_i})$ to $server_{auth}$, among them, × represents the multiplication on group G_1. When $server_{auth}$ receives $Update_{pw}$, first, according to ID_i query stored g^{S_i} value and use TPM to calculate $(g^{S_i})^{r_j}$, determine whether $(g^{r_j})^{S_i}$ and $(g^{S_i})^{r_j}$ are equal. If not equal to terminate the password update process; otherwise use TPM to calculate $\frac{New_{pw}}{sk_{auth\leftrightarrow i}}$ to get $H_2(ID_i||pw'_i)$, replace $H_2(ID_i||pw_i)$ with $H_2(ID_i||pw'_i)$ by querying ID_i.

4.3 Program Features

The authentication scheme proposed in this paper solves the problem of identity authentication between cloud and user in cloud environment.Program features are mainly reflected in:

1) In the key generation phase, it is not necessary for the user to establish a secure channel between the user and the KGC, more in line with the cloud environment, the practical application of open communication requirements;

2) All data calculation process is completed in the TPM or PTPM, the security of the hardware ensures the correctness of calculation and storage security.

3) The user uses PTPM to store the key and so on the information, as the PTPM has the characteristics of easy to carry, the user can use any terminal device to complete the registration and login authentication process;

4) Based on HMAC algorithm to achieve the identity authentication process, while ensuring the correctness of the certification results, which greatly improves the computing efficiency of both sides.

5 Proof of Safety

In this paper, the proposed authentication scheme is based on the certificate-less public key signature algorithm, at the same time, according to Section 3.2 of the program algorithm, since the security of the login authentication phase depends on the key value $k = g^{S_iS_j}$ of the HMAC algorithm, the HMAC security of the algorithm has been demonstrated in the literature [6]. Therefore, as long as the attacker in the previous stage of user registration can be calculated to obtain k, then it can be considered to break the authentication scheme proposed in this paper. The security of k value depends on the user and the cloud between the use of certificate-free public key signature algorithm to complete the process of user registration. As described in Section 2.2, the attack model of the certificateless public key signature algorithm includes two types of adversaries. Therefore, we need to attack the two types of adversaries to give the program to prove the security process.

Theorem 1. *Assuming that CDH is true for group G_1, for attacking adversary A_I, the scheme of this paper is based on the assumption that Random Oracle model has unforgeability (EUF-CMA) under adaptive selective message attack. That is, if any external adversary A_I is in time t_I. The advantage ϵ_I to the hash function H_1, H_2, secret value generation, public key generation, KGC generate part of the private key, public key replacement and signature generation, such as Oracle up to q_{H_1} times, q_{H_2} times, q_{sv} times, q_{pk} times, q_{part} times, q_{pkp} and q_s inquiries can forge a signature, then there is algorithm B_1, can be in time t_I to dominate the ϵ_I group G_1 on the CDH problem.among them,*

$$\begin{aligned}\epsilon_I &\geq \frac{(q_{part}+q_s)^{q_{part}+q_s}\epsilon_1}{q_{H_1}(q_{part}+q_s+1)^{q_{part}+q_s+1}}\\ t_1 &\leq t_I + (q_{H_1}+q_{part}+q_s)T_{G_1} + q_{H_2} + q_{sv} + q_{pk}.\end{aligned}$$

T_{G_1} represents the exponential time of one exponent in group G_1.

Proof. With algorithm B_I as the challenger, select $a, b \overset{R}{\longleftarrow} Z_p^*$. Given (g, g^a, g^b), B_I and the opponent A_I the following EUF-CMA attack game to get g^{ab}, where a, b is unknown to B_I.

1) System establishment. B_I sends the open system parameter $(G_1, G_2, e, p, g, H_1, H_2, pk_{KGC})$ to A_I, among them, $pk_{KGC} = g^a$. B_I controls Random Oracle H_1 and H_2, while maintaining the initially empty

hash value lists H_1^{list} and H_2^{list}, for Oracle hash inquiry response process A_I as follows:

H_1 **inquiry.** A_I request for identity ID_i of the H_1 value query. Assuming that $Y_i \in \{0,1\}$, among them, $pr[Y_i = 1] = \alpha$. A_I selects $r \xleftarrow{R} Z_p^*$ for each inquiry $(ID_i, Y_i), B_I$. If $Y_i = 1$ defines $H_1(ID_i) = g^r$; otherwise $H_1(ID_i) = (g^b)^r$ is defined. Finally, add $(ID_i, Y_i, H_1(ID_i))$ to list H_1^{list}. And $H_1(ID_i)$ as the result of response.

H_2 **inquiries.** A_I requests the H_2 value query for identity ID_i and public key pk_i. If tuple (ID_i, pk_i, β) exists in list H_2^{list}, the predefined output value is returned as the result of the query. Otherwise, select $y \xleftarrow{R} G_1$, Add (ID_i, y) to list H_2^{list}, and y as the result of response.

2) Stage 1. A_I initiates a series of inquiries. B_I maintains the public key list pk^{list} with the initial state empty, the response is as follows:

   a. Public key query A select B. If pk^{list} exists in tuple (ID_i, pk_i), returns pk_i as the result of the response; otherwise select $x_i \xleftarrow{R} Z_p^*$, calculate the public key $pk_i = g^{x_i}$ and return to A_{II}, and add (ID_i, x_i, pk_i) to list pk^{list}.

   b. Secret value generation query i. A_{II} select ID_i', B_{II} submits ID_i to the public key query Oracle. And returns x_i as the result of the response.

   c. Part of the private key to generate inquiries i. A_{II} requests partial private key value inquiries for identity ID_i. A_{II} select ID_i, B_{II} evaluates Q_i^{so} and returns to A_{II}.

   d. Signature generation query i. A_{II} select ID_i, if in the list H_2^{list} query ID_i corresponding to the $Y_i = 1$, then return to $(g^a)^r Q_i^{so}$. Here, a represents the private key value of the user; otherwise return $\perp$.

3) Forged. A_{II} the end of Phase 1 of the inquiry. Output Target ID_s and Fake Signature δ_s. makes the following response:

   a. Get from ID_i to generate public inquiry public key value g^a;

   b. Obtaining the secret value a of ID_s from the secret value generating inquiry;

   c. From the H_2 query to get the H_2 value of $ID_s(g^b)^r$;

   d. Output of forged signatures $\delta_s = (g^{ab})^r Q_i^{so}$.

So that B_{II} can get g^{ab} by calculating $g^{ab} = (\delta/Q_i^{so})^{r^{-1}}$, since δ_s, Q_i^{so}, and r are known for B_{II}, if A_{II} wins the EUFCMA attack game. There are: $e((g^{ab})^r Q_i^{so}, g) = e(Q_i, g^{so}) \cdot e(g^{br}, g^a)$, then B_{II} can break the group G_1 on the CDH problem. Event $\neg sv_{abort}$ indicates that B_{II} has not stopped A_{II} inquiring about the generation of the secret value, event $\neg sv_{abort}$ indicates that B_{II} did not stop A_{II} 's query for signature generation, the event $sigErr$ indicates that the forged signature δ_s of the object ID_s is generated, event Err_{id} represents List H_2^{list} stores ID_s, event Err_{y_i} represents, $Succeed$ indicates that B_{II} breaks the CDH problem on group G_1. According to the simulation process description, event $Succeed$ can be represented as $\neg sv_{abort} \wedge \neg sign_{abort} \wedge signErr \wedge Err_{id} \wedge Err_{y_i}$.

When $Y_i = 0$, B_{II} will stop A_{II} from generating a query for the secret value. Since A_{II} carries out q_{sv} times of secret value generation inquiries at most, So $pr[\neg sv_{abort}] \geq (1-\alpha)^{q_{sv}}$.

When $Y_i = 0$, B_{II} stop A_{II} signature generated inquiries. Since A_{II} performs q_s-times signature generation queries, so $pr[\neg sign_{abort}] \geq (1-\alpha)^{q_{sv}}$.

If $Err = \neg sv_{abort} \wedge \neg sign_{abor} \wedge Err_{y_i}$ occurs, then A_{II} would think that simulation attacks and real environment indistinguishable. As A_{II} break the advantages of the program ϵ_{II}, so $pr[signErr|Err] \geq \epsilon_{II}$.

Since A_{II} performs q_{H_2}-times H_2-queries at most, So $pr[Err_{id}] \geq 1/q_{H_2}$, which is:

$$\begin{aligned}
pr[Succeed] &= pr[\neg sv_{abort} \wedge \neg sign_{abort} \\
&\qquad \wedge signErr \wedge Err_{id} \wedge Err_{y_i}] \\
&= pr[\neg sv_{abort} \wedge \neg sign_{abort} \wedge Err_{y_i} \\
&\qquad \wedge signErr] pr[Err_{id}] \\
&= pr[Err] pr[signErr|Err] pr[Err_{id}] \\
&= pr[\neg sv_{abort}] pr[\neg sign_{abort}] pr[Err_{y_i}] \\
&\qquad pr[signErr|Err] pr[Err_{id}] \\
&\geq \frac{(1-\alpha)^{q_{sv}}(1-\alpha)^{q_s}\alpha\epsilon_{II}}{q_{H_2}} \\
&\geq \frac{(1-\alpha)^{q_{sv}+q_s}\alpha\epsilon_{II}}{q_{H_2}}
\end{aligned}$$

When $\alpha = \frac{1}{q_{sv}+q_s+1}$, $\frac{(1-\alpha)^{q_{sv}+q_s}\alpha\epsilon_{II}}{q_{H_2}}$ has a maximum value. So B_{II} breaks group G_1 on the CDH problem probability is

$$\epsilon_2 \geq \frac{(q_{sv}+q_s)^{q_{sv}+q_s}\epsilon_{II}}{q_{H_2}(q_{sv}+q_s+1)^{q_{sv}+q_s+1}}$$

According to the simulation process description, for each H_2 and signature generation and other inquiries, B_{II} needs to carry out an additional exponential operation on group G_1, so the running time of algorithm B_{II} is $t_{II} + (q_{H_2} + q_s)T_{G_1} + q_{H_1} + q_{sv} + q_{pk}$. □

6 Program Analysis

6.1 Efficiency Analysis

Since both the literature [4,9] and the proposed scheme adopt the idea of certificateless public key cryptosystem to solve the problem of identity authentication between cloud and user. This section therefore presents the performance analysis of users and clouds in terms of computational and communication overhead for these three scenarios. For convenience of description, EXP_{G_1} is defined here to denote an exponential operation on group G_1, EXP_{G_2} represents the exponential operation on group G_2, $Pairing$ represents a bilinear pair operation, H_{G_1} represents the hash operation on group G_1, H denotes that the hash value space in [4, 9] is not a hash operation of group G_1, M_{G_1} represents a multiplication (division) operation on group G_1, M_{G_2} represents a multiplication operation on group G_2, H_{mac} represents HMAC calculation. It should be added that the [4, 9] scheme also involves XOR operations, which are negligible in terms of computational overhead due to their very low computational cost.

6.1.1 Computing Overhead

According to [9] the program. In the key generation phase, the user and the cloud to generate public-private key pairs need to be 3 times M_{G_1} operation, at the same time in order to verify the PKG generated part of the private key value of the correctness,need to carry out 2 times $Pairing$ operation; in the registration phase, the cloud according to the identity of the user ID to send to determine whether the authorized user, so as to decide whether to register the user ID, does not involve any computing operation process (although efficient, but there are serious security vulnerabilities); in the certification phase, the user performs $4Pairing + M_{G_1} + M_{G_2} + 2EXP_{G_2} + 6H$ operation, the cloud performs $4Pairing + M_{G_1} + M_{G_2} + 2EXP_{G_2} + 6H + H_{G_1}$ operations.

For the program [4] proposed. In the key generation phase, the user and the cloud to generate public-private key pairs are required for 1 times M_{G_1} operation, at the same time in order to verify the PKG generated part of the private key value of the correctness,need to carry out 2 times $Pairing$ operation; in the registration phase with the literature [17] the same, the process does not involve any calculation operation; in the certification phase, the user performs $4Pairing + 6M_{G_1} + 2EXP_{G_2} + 6H$ operation, the cloud performs $3Pairing + 5M_{G_1} + EXP_{G_2} + 6H + 2H_{G_1}$ operations.

The following section focuses on the computational overhead of the text scheme: in the key generation phase, users and the cloud first need to EXP_{G_1} operation to generate a public key, then carry out $EXP_{G_1} + M_{G_1}$ operation to get part of the private key sent by KGC. Finally, the correctness of the received partial private key is verified by $H_{G_1} + 2Pairing$ operation. In the user registration phase, the user first performs $H_{G_1} + 2(M_{G_1} + EXP_{G_1})$ operation to generate the registration information; and then the cloud for $3Pairing + M_{G_1} + M_{G_2} + H_{G_1}$ operations on the registration information to complete the verification, and performing $H_{G_1} + 2(M_{G_1} + EXP_{G_1})$-time operation to generate the return information. Finally, the user performs $3Pairing + M_{G_1} + M_{G_2} + 2H_{G_1}$ operation to verify the cloud's return information. In the login authentication phase, the user first performs $EXP_{G_1} + H_{G_1}$ operation to generate authentication request information; then the cloud carries on $2EXP_{G_1} + H_{mac}$ operation to generate authentication information; then the user performs $EXP_{G_1} + 2H_{mac}$ operation to verify the identity of the cloud server; the cloud again $EXP_{G_1} + 2H_{mac}$ operations to verify the user's identity; the end user performs the $EXP_{G_1} + H_{mac}$ operation to determine whether the cloud verifies that the user's identity was successful.When the user needs to update the password, only need to carry out $EXP_{G_1} + 2H_{G_1}$ operations can be sent to the cloud password update information, and the cloud only need to be $EXP_{G_1} + H_{G_1}$ operation can complete the password update process.

6.1.2 Communication Overhead

Definition $|K|$ represents the length of the security parameter K, $|p|$ represents the length of the element in Z_p, $|id|$ represents the length of the user identity ID, $|hmac|$ denotes the information length of the HMAC algorithm. According to the scheme described in [9], in the registration phase, the user only needs to send the identity ID value, so the communication overhead is $|id|$, the cloud does not need to return any messages to the user, so there is no communication overhead; in the authentication phase, the length of the message sent by the user to the cloud is $3|p| + |id| + 2|K|$, and the length of the response message returned by the cloud to the user is $2|p| + |id| + |K|$. For the scheme proposed in [4], the communication overhead in the registration phase is the same as that in [9]; in the authentication phase, the length of the message sent by the user to the cloud is $2|p| + |id| + |K|$, and the length of the response message returned by the cloud to the user is $2|p| + |id|$.

For this program, the user registration phase, the registration information sent by the user includes ID_i, g^{S_i}, A $V^{x_i} Q_i^{s_0}$ and total length of $3|p| + |id|$; the information returned by the cloud includes ID_{auth}, g^{S_j} and $W^{x_j} Q_j^{s_0}$, the length is $2|p| + |id|$. In the login authentication phase, the login authentication request sent by the user includes ID_i, $H_2(ID_i||pw_i)$ and g^{r_i}, its length is $2|p| + |id|$; the server-side information returned includes ID_{auth}, g^{r_j} and D_1, length $|p| + |id| + |hmac|$; the user then sends $|hma|$-length information to the server;eventually, the server sends an $|hmac|$-length message to the user to determine whether the cloud verifies that the user's identity was successful.When the user needs to update the password, the user sends a password update request information length $2|p| + |id|$.

The following is the comparison of the communication cost between the scheme described in [4, 9] and the proposed scheme in the user and the cloud. As can be seen from Table 1, In the user registration phase, compared with the literature [4, 9] program. Our scheme users and the cloud need to give the other party to send additional information to complete the user registration process, but because the user registration is a one-time, so the communication process of the registration process is acceptable. In the authentication phase, since the value of $|hmac|$ can not be greater than $|K|$, so compared with the liter-

Table 1: Comparison between Ref.[12,13] and our scheme in communication overhead

Program	User registration phase		Login authentication phase	
	User	Cloud	User	Cloud
Cloud [9]	$\|id\|$	0	$3\|p\| + \|id\| + 2\|K\|$	$2\|p\| + \|id\| + \|k\|$
Cloud [4]	$\|id\|$	0	$2\|p\| + \|id\| + \|K\|$	$2\|p\| + \|id\|$
Our	$3\|p\| + \|id\|$	$2\|p\| + \|id\|$	$2\|p\| + \|id\| + \|hmac\|$	$\|p\| + \|id\| + 2\|hmac\|$

ature [4, 9] program, this paper program communication overhead in the certification process at least remain unchanged. In summary, for the user and the cloud, the communication overhead costs of our scheme produced is acceptable.

7 Conclusion

In this paper, an authentication scheme based on PTPM and certificateless public key signature algorithm is proposed. In the realization of the user and cloud identity on the basis of uniqueness. On the one hand, using PTPM to achieve the security of the terminal platform and the authenticity of the cloud and the results of the authenticity of the correct target, while allowing users to use any terminal device to complete with the cloud two-way identity authentication process; on the other hand, certificate-free public key signature algorithm solves the problem of certificate management of traditional public key cryptosystem and key escrow problem of identity-based cryptosystem. Finally, the security of the program was proved theoretically. At the same time, compared with the efficiency and other performance of the existing schemes, the proposed scheme not only significantly improves the efficiency of identity authentication between users and cloud, but also adapts to the practical application requirements of open communication in cloud environment. The next step is to consider cross-domain identity authentication and to prevent user identity and privacy issues such as leakage.

Acknowledgments

This work is partially supported by Scientific Study Project for Institutes of Higher Learning, Ministry of Education, Liaoning Province (LQN201720), and Natural Science Foundation of LaioNing Province, China(20170540819). The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

References

[1] S. S. Al-Riyami, K. G. Paterson, "Certificateless public key cryptography," in *Proceedings of the Advances in Cryptology (ASIACRYPT'03)*, pp. 452–473, 2003.

[2] S. Binu, M. Misbahuddin, P. Raj, "A strong single sign-on user authentication scheme using mobile token without verifier table for cloud based services," in *Computer and Network Security Essentials*, pp. 237–261, 2018. DOI: 10.1007/978-3-319-58424-9_14.

[3] C. L. Cao, R. Zhang, M. Y. Zhang, Y. X. Yang, "IBC-Based entity authentication protocols for federated cloud systems," *KSII Transactions on Internet & Information Systems*, vol. 7, no. 5, pp. 1291–1312, 2013.

[4] Z. M. Dong, L. Zhang, J. T. Li, "Security enhanced anonymous remote user authentication and key agreement for cloud computing," in *Proceedings of the 17th International Conference on Computational Science and Engineering (CSE'14)*, pp. 1746–1751, 2014.

[5] H. A. Elbaz, M. H. Abdelaziz, T. Nazmy, "Trusting identity based authentication on hybrid cloud computing," in *Proceedings of Cloud Computing*, pp. 179–188, 2014.

[6] L. Han, J. Q. Liu, D. W. Zhang, Z. Han, X. Y. Wei, "A portable TPM scheme for general-purpose trusted computing based on EFI," in *Proceedings of International Conference on Multimedia Information Networking and Security (MINES'09)*, pp. 140–143, 2009.

[7] H. W. Li, Y. S. Dai, L. Tian, H. M. Yang, "Identity-based authentication for cloud computing," in *Proceedings of the 1st International Conference on Cloud Computing*, pp. 157–166, 2009.

[8] X. H. Li, B. Yang, "Efficient identity-based signature authentication scheme in cloud service," *International Journal of Advancements in Computing Technology*, vol. 5, no. 5, pp. 867–876, 2013.

[9] R. Mishra, "Anonymous remote user authentication and key agreement for cloud computing," in *Proceedings of the 3rd International Conference on Soft Computing for Problem Solving*, pp. 899–913, 2014.

[10] W. Nie, X. Xiao, Z. Wu, *et al.*, "The research of information security for the education cloud platform based on AppScan technology," in *The 5th IEEE International Conference on Cyber Security and Cloud Computing (CSCloud'18) and the 4th IEEE International Conference on Edge Computing and Scalable Cloud (EdgeCom'18)*, 2018. DOI: 10.1109/CSCloud/EdgeCom.2018.00040.

[11] Y. X. Sang, *Study on Some Topics of Certificateless Public-key Cryptography*, Ph.D. Thesis, Xiamen University, 2009.

[12] A. Shamir, "Identity based cryptosystems and signature schemes," in *Proceedings of the Advances in Cryptology (CRYPTO'84)*, pp. 47–53, 1985.

[13] X. W. Wu, *Research and Implementation of Key Technology on Portable TPM*, MS. Thesis, Beijing Jiaotong University, 2010.

[14] Z. Xia, X. Wang, L. Zhang, *et al.*, "A privacy-preserving and copy-deterrence content-based image retrieval scheme in cloud computing," *IEEE Transactions on Information Forensics & Security*, vol. 11, no. 11, pp. 2594–2608, 2017.

[15] X. Xiao-Tao, C. Zhe, J. Fei, *et al.*, "Research on service-oriented cloud computing information security mechanism," in *IEEE International Conference on Computer & Communications*, 2016. DOI: 10.1109/CompComm.2016.7925188.

[16] L. Yan, C. M. Rong, G. S. Zhao, "Strengthen cloud computing security with federal identity management using hierarchical identity-based cryptography," in *Proceedings of the Cloud Computing*, pp. 167–177, 2009.

[17] D. W. Zhang, Z. Han, G. W. Yan, "A portable TPM based on USB key," in *Proceedings of the 17th ACM Conference on Computer and Communications Security*, pp. 750–752, 2010.

Biography

Hui Xia received his B. S. and M. S. degree from Xidian University, China, in 2003, 2006, respectively. He is currently a associate professor with Shenyang Normal University. His research interests include cloud computing, cryptography and information security.

WeiJi Yang received his B. S. degree from Zhejiang Chinese Medical University, China, in 2005, M. S. degrees from Beijing University of Posts and Telecommunications, China, in 2009. He is currently a Research Associate with ZheJiang Chinese Medical University. His research interests include Information Security and Traditional Chinese Medicine Informatics.

High Security Speech BioHashing Authentication Algorithm Based on Multi-feature Fusion

Yi-Bo Huang[1], Hao Li[1], Yong Wang[1], and Qiu-Yu Zhang[2]
(Corresponding author: Yi-Bo Huang)

College of Physics and Electronic Engineering[1]
Northwest Normal University, Anning District, Lan Zhou, China
School of Computer and Communication[2]
Lanzhou University of Technology, Qilihe District, Lan Zhou, China
Email: huang_yibo@foxmail.com

(Received July 16, 2020; Revised and Accepted Apr. 24, 2021; First Online Oct. 19, 2021)

Abstract

Aiming at the problems of low discrimination and weak security in existing speech authentication algorithms, the paper proposes a high-security speech BioHashing authentication algorithm based on multi-feature fusion. Firstly, in the frequency domain, the cepstrum and short-term average amplitudes are fused by convolution and then fused with the amplitudes. Next, the improved pseudo-random function is used to transform the fused features to generate a Biosafety template. Then, K-means and KNN are used to divide the similar features into K classes, and the BioHashing sequence is constructed for the features in each class. Finally, Anrold chaotic systems of different secret keys encrypt the BioHashing sequence in each class. Experimental results show that the algorithm can classify the extracted features and has good discrimination, robustness, and security. Meanwhile, the algorithm can realize the small range of tamper detection and location.

Keywords: BioHashing; Improved Pseudo-Random Matrix; K-means; KNN; Speech Authentication

1 Introduction

In recent years, the explosive growth of Internet-based digital speech and the increasingly prominent speech communication network security issues, true and complete speech authentication is particularly important for the security of Internet applications. BioHashing defines an orthogonal pseudo-random function using a specific secret key, and uses this function to transform the extracted features. Therefore, BioHashing protects user speech information without reducing authentication performance [9, 11, 16].

BioHashing is also called adding salt method. It is mainly used for the identification, authentication and retrieval of iris, fingerprints, palmprints, voiceprints, face and other multimedium. In terms of speech, BioHashing authentication mainly includes the construction of Biosafety template, BioHashing structure and matching, in which Biosafety template mainly includes speech feature extraction and pseudo-random matrix construction.

At present, the features extracted from speech clips include short-term energy, short-term zero-crossing rate, short-term correlation [22], frequency band variance [7], resonance hump, spectrum entropy, spectrogram [8], discrete wavelet transform (DWT) [1, 21], linear prediction coefficient (LPC) [4, 20], linear spectrum frequence (LSF), Mel-frequency cepstrum coefficient (MFCC) [2, 14] and perception linear prediction (PLP) [15], *etc.* Chen *et al.* [3] proposed perceptual audio hashing algorithm based on Zernike moment and maximum-likelihood watermark detection, the algorithm is not only robust to content preserving operations, but also has good anti-collision ability. Li *et al.* [12] extracted voiceprint Mel-frequency cepstral coefficients of segmented speech as perceptual features and construct perceptual hashing sequence. The experiment shows that the robustness is better, especially in resampling and MP3 compression, but the discrimination is relatively poor.

In [13], Radon Transform (RT) was applied in DWT domain, and then dimensionality was reduced by Discrete Cosine Transform (DCT), which has higher efficiency but relatively low randomness. Zhang *et al.* [18] used spectral subtraction to extract energy and entropy value of speech clips after noise reduction processing as perceptual features and construct perceptual hashing sequence. The algorithm has good robustness and efficiency, but the discrimination still need to be further improved. Zhang *et al.* [19] used a part of the spectrogram to represent the low-frequency information of speech clips, then obtained features through LBP and construct perceptual hashing. The algorithm is not easily affected by the addition of MP3 compression, noise and volume regulation, show in that it has good robustness. Huang *et al.* [6] obtained features by transforming the variance of frequency domain subbands through Constant Q Transform (CQT) and Tensor Decomposition (TD). This experiment verifies its good discrimination by comparing perceptual hashing sequences of different lengths, but it does not protect the security of perceptual hashing sequence in the cloud. Yao *et al.* [17] proposed the lattice-based remote biometric authentication scheme (RRBAS) for multi-server environments with good security and efficiency than existing solutions.

From the above analysis, it can be seen that hashing algorithm does not protect the security and privacy of speech features, and most speech authentication algorithms are independently optimized for recognition and robustness. To solve the above problems, this paper proposes the high security speech BioHashing authentication algorithm based on multi-feature fusion, an improved pseudo-random function is used to transform speech features into another domain, the constructed BioHashing not only improve security, but also balance overall performance.

2 Related Theory

2.1 Improved Pseudo-Random Matrix

By combining two one-dimensional mapping functions, an improved mapping function with strong randomness and sensitivity is constructed, and an improved pseudo-random matrix with higher performance is constructed.

The improved mapping function χ_{n+1} is constructed of the iterative sequence ξ_{n+1} generated by the Logistic mapping function as the initial value ζ_n of the Tend mapping function, as shown in Equation (1):

$$\chi_{n+1} = 1 - 4\chi_n(1-\chi_n) \quad 0 \le \chi_n \le 1 \tag{1}$$

Where, the Logistic mapping function and the Tend mapping function are shown in Equations (2) and (3) respectively. Especially, when parameters α and β are 2, the two mapping functions are highly sensitive to initial values, have obvious chaotic characteristics, and have wider ergodicity of mapping iteration.

$$\xi_{n+1} = \alpha\xi_n(1-\xi_n) \rightarrow \ 2\xi_n(1-\xi_n) \quad 0 \le \xi_n \le 1 \tag{2}$$

$$\zeta_{n+1} = \beta - 1 - \beta|\zeta_n| \rightarrow 1 - \ 2|\zeta_n| \quad -1 \le \zeta_n \le 1 \tag{3}$$

The improved pseudo-random matrix is constructed of the improved mapping function, and the specific steps are as follows:

Step 1: A pseudo-random sequence $a = \{a_1, a_2, ..., a_L\}$ with length $L \ (L >> l)$ is constructed of Equation (1), where the initial value $a_0 = 0.4$.

Step 2: A new pseudo-random sequence $A = \{A_1, A_{1+d}, A_{1+\ 2d}, ..., A_l\}$ is constructed by selecting a numbers from sequence, where the sampling interval $d = 10$.

Step 3: The pseudo-random matrix B is constructed of the new sequence A, as shown in Equation (4), and then schmidt ortho-gonalization of B is performed to obtain the final pseudo-random matrix Φ.

$$B = \begin{bmatrix} A_1 & \cdots & A_{1+MNd} \\ A_{1+d} & \cdots & A_{1+\ (MN+1)d} \\ \vdots & \vdots & \vdots \\ A_{1+\ (N-1)d} & \cdots & A_l \end{bmatrix} \tag{4}$$

2.2 K-Means

K-means is a unsupervised machine learning algorithm [5], whose core idea is to classify the input sample set Ω into K clusters. In this paper, the K-means cluster result is used as the label for the classification sample.

According to Equation (5), the distance I between the remain object and K-th initial cluster center point is calculated.

$$I(K) = ||y_{j(i)} - y_{\bar{K}(i)}||^2 \quad j + K = S \tag{5}$$

Where, $y_{j(i)}$ is the i-th sampling value of the remain object j, $y_{\bar{K}(i)}$ is the i-th sampling value of the K-th initial cluster center point, $|| \bullet ||$ is the summation formula.

After clustering the input sample set Ω is completed, the new center point in each cluster is recalculated according to Equation (6):

$$y^{\bar{\prime}}_K = \frac{\sum_{1 \le i \le n} y^i}{n} \tag{6}$$

Where, $y^{\bar{\prime}}_K$ is the new cluster center point in the K-th cluster.

2.3 KNN

KNN is a supervised machine learning algorithm [10], its core idea is to input training set $T1$ and test set $T2$, and determine the data type of test object according to the data type of K samples closest to $T1$.

According to Equation (7), the weight ϑ between the test object and K neighbors is calculated.

$$\vartheta\ (\bar{z_c}, C_K) = \sum dist(\bar{z_c}, \bar{z_x})\partial\ (\bar{z_x}, C_K) \tag{7}$$

Where, $\bar{z_c}$ is the feature vector of the test object, $\bar{z_x}$ is the feature vector of the training object, C_K is the sample set of class K, $\sum dist(\bar{z_c}, \bar{z_x})$ is the similarity calculation formula between samples, $\partial\ (\bar{z_x}, C_K)$ is the class attribute function.

According to Equation (8), the ϑ between classes is compared, and the test object is divided into the class with the largest ϑ.

$$\partial\ (\bar{z_x}, C_k) = \begin{cases} 1, z_x \in C_K \\ 0, z_x \notin C_K \end{cases} \tag{8}$$

Output the label of the sample with the largest ϑ in the training sample.

3 The Proposed Algorithm

3.1 BioHashing Structure

The block diagram of constructing feature vector is shown in Figure 1.

Step 1: Pre-processing Pre-processing includes pre-emphasis, framing and windowing. The speech clip $x'(t)$ is obtained by pre-emphasis the input speech clip $x(t)$, then the processed $x'(t)$ is framed and windowed, where in the window function selects a Hamming window to smooth the edge of the frame. The speech clip $x'(t)$ is divided into N frame, and $x(n) = \{x_i(n)|n = 1, 2..., M; i = 1, 2, ..., N\}$, where, $x(n)$ is the n-th sampling value of i-th frame.

Step 2: Feature extraction FFT is applied to each frame of speech clip to transform the speech clip from time domain data to frequency domain data. Then, amplitude features (F) are extracted from the upper half of each frame and the matrix $\Im$ is constructed, cepstrum features (c) are extracted from real in the lower half of each frame and the matrix $\Re$ is constructed, short-time average amplitude features (E) are extracted from imag in the lower half of each frame and the matrix $\aleph$ is constructed. The extracted features by fusion are named The F-cE algorithm.

$$1\ c = FT^{-1}[\ln|x_i(\omega|] \qquad \omega = 1, 2, ..., w \tag{9}$$

$$E = \sum_{\omega=1}^{M-w} |x_i(\omega+1) - x_i(\omega| \tag{10}$$

Where, $ln|\bullet|$ is the formula for taking the real, $x_i(\omega)$ is the ω-th sampling value of the i-th frame of the frequency domain data, w is the length of the lower half.

Step 3: Improved pseudo-random matrix According to Equation (1) and Equation (4), the pseudo-random matrix is constructed as shown in Equation (11):

$$\Phi = \begin{bmatrix} \phi_1 & \cdots & \phi_{1+MNd} \\ \phi_{1+d} & \cdots & \phi_{1+\ (MN+1)d} \\ \vdots & \vdots & \vdots \\ \phi_{1+\ (N-1)d} & \cdots & \phi_l \end{bmatrix} \tag{11}$$

Where, $\phi_1, ..., \phi_l$ are schmidt ortho-gonalization values.

Step 4: Feature vector According to Equation (12), the matrix $\Re$ convolves with the matrix $\aleph$, and then splices with the matrix $\Im$, then it iterates with matrix Φ to form matrix $H(N, N)$, finally, N - dimensional matrix H is generated into feature vector $H(1, N)$. The generation process of feature vector is shown in Equation (13):

$$C(v, g) = \sum_{\iota=0}^{\iota r-1} \sum_{\kappa=0}^{\iota c-1} \Re\ (\iota, \kappa\aleph\ (v-\iota, g-\kappa \tag{12}$$

Where, $0 \le v < \iota r + \kappa r - 1, 0 \le g < \iota c + \kappa c - 1$, ιr and ιc are the rows and columns of matrix $\Re$, κr and κc are the rows and columns of matrix $\aleph$.

$$[\Phi] \cdot \begin{bmatrix} \Im \\ \Re * \aleph \end{bmatrix} \rightarrow H(N, N) \rightarrow H(1, N) \tag{13}$$

Where, $[\bullet] \cdot [\bullet]$ is the iterative formula, $*$ is the convolution formula.

Step 5: K-means According to the ratio of training set and test set 7:3, feature vectors are randomly selected from feature vector library as training set $T1$. Then, according to the specific steps of K-means, the K class is clustered.

Step 6: KNN The remaining feature vectors in the feature vector library are used as test set $T2$ and classified according to the specific steps of KNN.

The process of K-means-KNN is shown in Figure **??**. Where, $K = 3$. Purple circles, green pentagons and blue squares respectively represent a class with similar features, and red triangles represent the center point of clustering.

Step 7: BioHashing structure As shown on the Figure 3, the BioHahing sequence is constructed for the feature vectors in each class, and the speech data is converted into a short binary string. The generated BioHashing sequence is shown in Equation (14):

$$h = [h(1) \quad h(2) \quad \cdots \quad h(N)] \tag{14}$$

In Figure 3, According to K-means, K class labels are generated by T1, which serve as the classification basis of T2. After the classification is completed, the BioHahing sequence is constructed by using the same method for the feature vectors in each class.

Constructing method: Find the mean value ϖ of the feature vector H. If the i-th sampling point value of the feature vector H is greater than ϖ, the i-th sampling point value of the BioHashing sequence h is 1, otherwise, it is 0.

$$h(i) = \begin{cases} 1, H(i) > \varpi \\ 0, else \end{cases} \tag{15}$$

Step 8: Scrambling encryption Firstly, the Arnold chaotic mapping is used to generate the pseudo-random sequence $S = [s_1, s_2, ..., s_N]$ with the same length as the BioHashing sequence h. Then, arrange S in descending order to obtain S'. Finally, the encrypted BioHashing sequence h' was obtained according to the one-to-one mapping relationship between BioHashing sequence h and S'. As shown in Figure 3, As shown in Figure 3, different secret keys ℓ are applied to each class. Arnold function is shown in Equation (16):

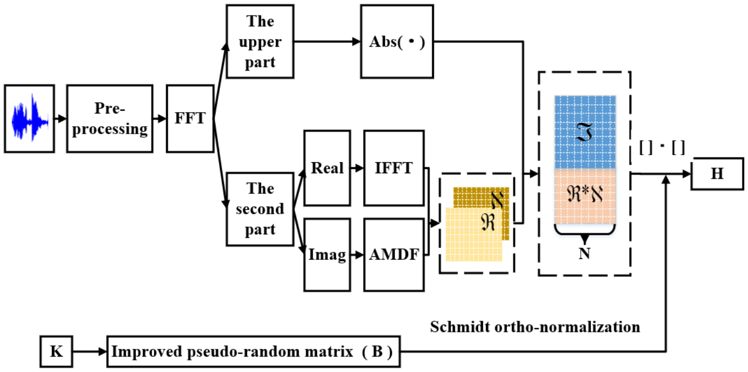


Figure 1: Block diagram of constructing feature vector

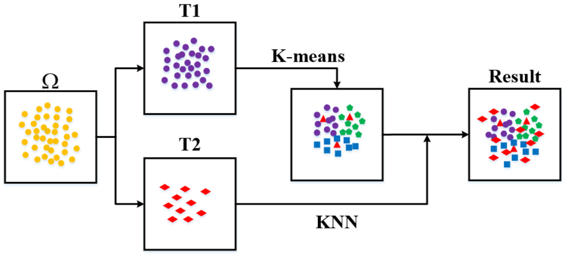


Figure 2: K-means-KNN

$$\begin{bmatrix} \ell_{\theta+1} \\ S_{\theta+1} \end{bmatrix} = \begin{bmatrix} 1 & \varphi \\ \psi & \psi\varphi+1 \end{bmatrix} \begin{bmatrix} \ell_\theta \\ S_\theta \end{bmatrix} \bmod (N) \quad (16)$$

Where, ℓ is the secret key for mapping encryption, ψ and φ are parameters, θ is the number of iterations, mod $(\bullet)$ is the remainder formula.

3.2 BioHashing Matching

The block diagram of BioHashing matching is shown in Figure 4.

Matching method: According to the BioHashing structure, the paper obtain the BioHashing sequences h_1 and h_2 of the speech clips x_1 and x_2, and $D(:,:)$ is denoated as the normalized Hamming distance of h_1 and h_2, also known as bit error rate (BER). The formula is as shown in Equation (17):

$$\begin{aligned} D(h_1,h_2) &= \frac{1}{N}\sum_{i=1}^{N}|h_1(i)-h_2(i)| \\ &= \frac{1}{N}\sum_{i=1}^{N}(h_1(i)\oplus h_2(i)) \end{aligned} \quad (17)$$

Where, $h_1(i)$ and $h_2(i)$ are the i-th sampling values of the BioHashing sequences constructed by speech clips x_1 and x_2 respectively.

The paper use the hypothesis test of BER is used to describe the BioHashing matching.

Δ_0: If the perceptual contents of the two speech clips x_1 and x_2 are the same: $D < \tau$.

Δ_1: If the perceptual contents of the two speech clips x_1 and x_2 are not the same: $D \geq \tau$.

Where, τ represents the authentication threshold. When the bit error rate between two BioHashing sequences is less than the set threshold τ, the authentication passes, otherwise, the authentication fails.

4 Experimental Results and Analysis

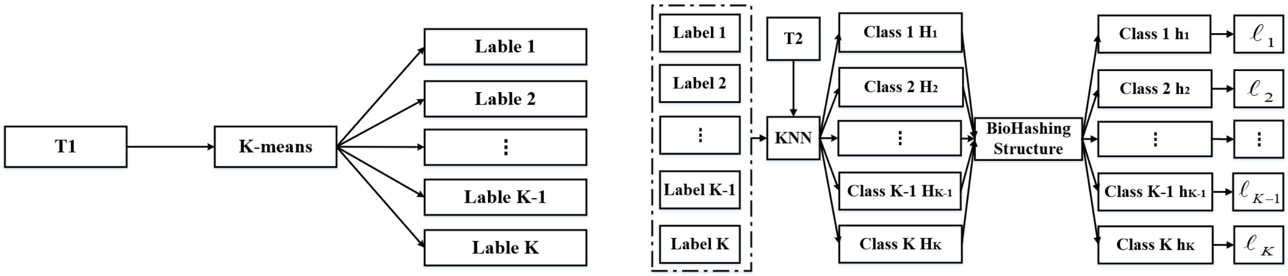
The experimental speech data come from TIMIT speech library and TTS speech library. There are different 1200 speech clips in the original speech library. The format of each speech clip is WAV, the length is 4 s, the sampling frequency is 16 kHz, and the bit rate is 256 kbps.

The operating experimental hardware platform is Intel(R) Core(TM) i5-7500 CPU, 3.40 GHz, with computer memories of 4G. The operating software environment is MATLAB R2018b of Windows 7 system. In this study, the main parameters of the experiment are set as follows: the number of frames N=802 bits, the length of a frame M=200 ms, and the number of classes K=18.

4.1 Discrimination Test and Analysis

Discrimination is used to evaluate the reliability of the algorithm by distinguishing different or the same speech clips. And the similarity between them is determined by calculating the BER between two pairs of BioHashing sequences. Speech clips of different contents generate different BioHashing sequences, and the BER between them basically follows the normal distribution. In this paper, a total of 719,400 BER can be obtained, as shown on the Figure 5.

Figure 5 describes BER among different BioHashing sequences basically following the normal distribution. Therefore, the algorithm can good distinguish speech clips of different contents.



(a) K-means (b) KNN

Figure 3: Block diagram of BioHashing structure and encryption

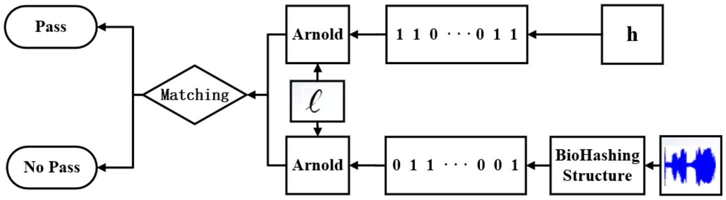


Figure 4: Block diagram of BioHashing matching

Table 1: Normally distributed parameters of different BioHashing sequence lengths

BioHashing length	Theoretical value		Experimental value	
	μ	σ	μ	σ
802 bits (Proposed)	0.50	0.0177	0.4952	0.0183
639 bits	0.50	0.0197	0.4948	0.0206
532 bits	0.50	0.0216	0.4945	0.0226
401 bits	0.50	0.0250	0.4937	0.0261

According to the De Moivre-Laplace central limit theorem, the hamming distance is approximate obeying the normal distribution $\mu = p$, $\sigma = \sqrt{p(1-p)/N}$, p represents the probability of 0 or 1, N is the number of bits in a hashing sequence. In this paper, the length of the BioHashing sequences is 802 bits, and the mean value and standard deviation of the theoretical normal distribution parameters are $\mu = 0.50$, $\sigma = 0.0177$. The normal distribution parameters of the experiment are shown in Table 1.

It can be seen from Table 1 that the experimental values of normal distribution obtained by the algorithm are very close to the theoretical values. Therefore, the BioHashing sequence obtained by the algorithm has good randomness.

In order to measure the discrimination of the algorithm under different thresholds, Fales Accept Rate (FAR) is introduced. FAR in this paper refers to the speech that should not be matched as the matched speech. FAR is shown in Equation (18):

$$\begin{aligned} FAR &= P_{FAR}\,(x|L<\tau) \\ &= \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\tau} e(\frac{-(x-\mu)^2}{2\sigma^2})dx \qquad (18) \end{aligned}$$

Where, μ is the expected value, σ is the standard deviation. Equation (18) shows that the smaller τ is, the smaller FAR is, *i.e.*, the better discrimination is.

As can be seen from Figure 6, FAR curve obtained by the experiment almost coincides with the theoretical FAR curve, *i.e.*, the FAR of the algorithm is very close to the theoretical FAR. Therefore, the BioHashing sequence obtained by the algorithm has good anti-collision ability.

Further, FAR is used to illustrate the discrimination of the algorithm in this paper, the FAR of different algorithms under different thresholds τ is calculated respectively, as per the table shown.

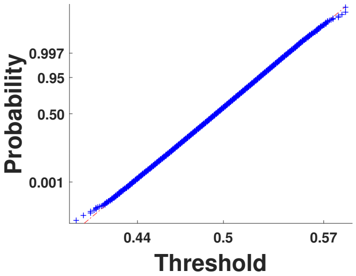


Figure 5: BER normal distribution of the algorithm

Table 2 shows with the same threshold τ, the FAR of the algorithm in this paper is smaller than that of other algorithms, and the smaller τ is, the smaller FAR is. Therefore, the algorithm in this paper has good discrimination. As τ increase, the number of speech clips error judgment will increase correspondingly. However,

Table 2: Comparison of FAR of different algorithms

τ	Proposed	[21]	[12]	[19]
0.10	$\mathbf{2.2611e^{-103}}$	$3.6542e^{-42}$	$3.0310e^{-38}$	$1.7668e^{-28}$
0.20	$\mathbf{1.2129e^{-58}}$	$1.4050e^{-24}$	2.6890^{e-22}	$1.9604e^{-16}$
0.25	$\mathbf{4.1748e^{-41}}$	$1.2151e^{-17}$	$5.1740e^{-16}$	$9.8689e^{-12}$
0.30	$\mathbf{8.8188e^{-27}}$	$6.1663e^{-12}$	$7.5420e^{-11}$	$6.6409e^{-08}$
0.35	$\mathbf{1.1701e^{-15}}$	$1.8744e^{-7}$	$8.4800e^{-7}$	$6.1048e^{-05}$

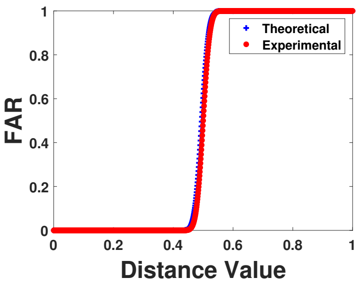


Figure 6: FAR curve of the algorithm

$\tau = 0.35$, it means that the number of false judgments in each 1×10^{15} speech clips is only 1.1701.

It can be obtained from Table 1 and Table 3 that FAR is easily affected by the length of BioHashing sequence in the same algorithm, so in the experiment, it is not enough to use FAR to measure the discrimination of the algorithm. In statistics, Entropy Rate (ER) is used as a measure of the uncertainty of random events. In this paper, ER is introduced as a measure of discrimination to better illustrate the discrimination of the algorithm. ER is shown in Equation (19):

$$ER = -[c\log_2 c + \ (1-c)\log_2(1-c)] \tag{19}$$

Where, $c = \frac{1}{2}(\sqrt{\frac{|\sigma^2 - {\sigma_1}^2|}{\sigma^2 + {\sigma_1}^2}} + 1)$.

As shown in Table 4, the ER of the algorithm in this paper is the largest, and the greater ER is, the better discrimination is. Therefore, it proves that the algorithm has good discrimination.

Table 4: Comparison of ER of different algorithms

Algorithm	ER
Proposed	**0.9758**
[21]	0.9432
[12]	0.8874
[19]	0.9126

In order to measure the influence of K-means-KNN on discrimination, the FAR of the two algorithms in Table 5 under different thresholds was calculated respectively.

Table 5 shows with the same threshold, the FAR of the algorithm proposed in this paper is less than the FAR of the F-CE algorithm, which shows that the algorithm proposed in this paper not only extracts excellent features, but also improves the discrimination to a certain extent by using K-means-KNN.

In a word, the algorithm has good discrimination, and it has no conflicts between the generated BioHashing sequences.

4.2 Robustness Test and Analysis

In order to evaluate the robustness of the proposed algorithm, BER of the original speech library after various content preserving operations is calculated. According to the environment of speech transmission, the content preserving operations are performed on each speech clip in the original speech library. A speech library of 13200 different content preserving operations was established, including 11 types of content preserving operations, such as echo, noise, low pass filter, resampling and MP3 compression. Various content preserving operations is summarized in Table 6.

The BER mean and maximum values after 11 content preserving operations were calculated respectively, it is illustrated as follow.

As can be seen from Table 7, compared with other algorithms, the BER of the proposed algorithm in this paper is small, with the mean value below 0.1 and the maximum value of the mean is 0.0851. Therefore, the proposed algorithm has better robustness. Especially in volume adjustment and MP3 compression, the algorithm has better performance than other algorithms. BioHashing sequence has not great changed in volume adjustment, so the BER is lower. In MP3 compression, the MP3 format uses a large compression ratio in the high frequency part and a small compression ratio in the low frequency part, which has a great impact on the speech features. However, the BER mean and maximum of the algorithm in this paper are both small.

Fales Rejection Rate (FRR) is introduced to further measure the robustness of the algorithm. FRR in this paper refers to the speech that should be matched as not be matched speech. FRR is shown in Equation (20):

$$\begin{aligned} FRR &= P_{FRR}\,(x|L \geq \tau) \\ &= 1 - \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\tau} e(\frac{-(x-\mu)^2}{2\sigma^2})dx \end{aligned} \tag{20}$$

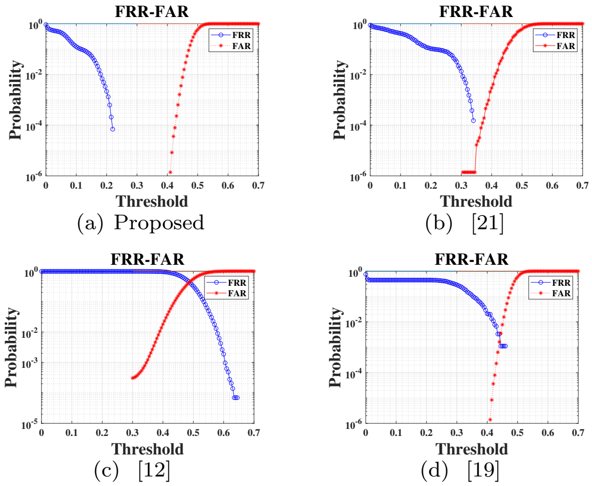
Table 3: Comparison of FAR of different BioHashing sequence lengths

τ	802 bits (Proposed)	639 bits	532 bits	401 bits
0.10	$\mathbf{2.2611e^{-103}}$	$1.5240e^{-82}$	$8.0575e^{-69}$	$9.6736e^{-52}$
0.20	$\mathbf{1.2129e^{-58}}$	$5.7655e^{-47}$	2.8234^{e-39}	$1.0845e^{-29}$
0.25	$\mathbf{4.1748e^{-41}}$	$5.1413e^{-33}$	$1.0901e^{-27}$	$4.8375e^{-21}$
0.30	$\mathbf{8.8188e^{-27}}$	$1.2844e^{-21}$	$3.2132e^{-18}$	$5.7117e^{-14}$
0.35	$\mathbf{1.1701e^{-15}}$	$9.1963e^{-13}$	$7.3942e^{-11}$	$1.8243e^{-08}$

Table 5: FAR of the proposed algorithm

τ	The F-cE algorithm	The F-cE algorithm+ K-means-KNN
0.10	$4.4479e^{-102}$	$\mathbf{2.2611e^{-103}}$
0.20	$7.3062e^{-58}$	$\mathbf{1.2129e^{-58}}$
0.25	$1.5532e^{-40}$	$\mathbf{4.1748e^{-41}}$
0.30	$2.1804e^{-26}$	$\mathbf{8.8188e^{-27}}$
0.35	$2.0691e^{-15}$	$\mathbf{1.1701e^{-15}}$

Equation (20) shows that the smaller τ is, the smaller FRR is, *i.e.*, the better robustness is.



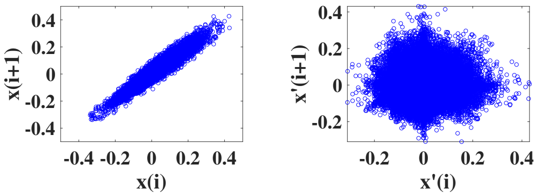
(a) Proposed (b) [21]

(c) [12] (d) [19]

Figure 7: FRR-FAR curves of diffierent algorithms

It can be seen from Figure 7 that, the FRR-FAR curves of the proposed algorithm are not crossed, and the two curves have a large interval, which indicates that the algorithm not only has both good discrimination and robustness, but also can accurately identify content preserving operations. Compared with the algorithm in this paper, the FRR-FAR curves of [21] is very close, indicating that it does not be good solve the problem of discrimination and robustness conflicts. The FRR-FAR curves of [12] and [19] intersect each other, indicating that the extracted features has poor robustness, and the randomness of BioHashing sequence is poor.

To sum up, the algorithm not only has good anti-interference ability, but also has good robustness.



(a) Before encryption (b) After encryption

Figure 8: Scatter plot before and after encryption

4.3 Security Analysis

To solve the problem of high algorithm transparency and improve the security of the BioHashing sequence in the cloud, this paper proposes an Arnold chaotic mapping algorithm that uses different secret keys for different classes.

In order to verify the correlation of speech before and after encryption, a speech clip is randomly selected from the original speech library, and then its continuous 32000 sample points are randomly selected as the sampling points. Taking $x(i)$ as the horizontal ordinate and $x(i+1)$ as the ordinate, the scatter diagram before and after encryption is shown in Figure 8.

In order to verify the security of this algorithm, a speech clip is randomly selected from the original speech library, speech encryption and decryption waveform is shown in Figure 9.

As can be seen from Figure 8 and Figure 9, the algorithm proposed in this paper has good encryption performance. The correlation of the original speech clip is completely disrupted in the algorithm, which makes the waveform of the speech clip disordered and difficult to be cracked in practice.

The security of this algorithm is futher tested through Spearman's correlation coefficient (ρ). The formula is shown in Equation (21):

$$\rho = \frac{\sum_i (x_i - \bar{x})(S_i - \bar{S})}{\sqrt{\sum_i (x_i - \bar{x})^2 (S_i - \bar{S})^2}} \tag{21}$$

Where, x_i is the i-th sampling value of the speech clip before encryption, S_i is the i-th sampling value of the speech clip after encryption, $barx$ and $barS$ are the mean values of speech clip before and after encryption.

Figure 10 shows, ρ after encryption and ρ after error decryption is between (-0.1,0.1) , indicating that the data

Table 6: Content preserving operations

Operation means	Operation method	Abbreviation
Volume Adjustment 1	Volume up 50%	V.1
Volume Adjustment 2	Volume down 50%	V.2
Low-pass Filtering 1	12 order FIR low-pass filtering, Cutoff frequency of 3.4 kHz	F.I.R
Low-pass Filtering 2	12 order Butterworth low-pass filtering, Cutoff frequency of 3.4 kHz	B.W
Resampling 1	Sampling frequency decreased to 8 kHz, and then increased to 16 kHz	R.8→16
Resampling 2	Sampling frequency decreased to 32 kHz, and then increased to 16 kHz	R.32→16
Echo Addition	Superimposed attenuation 30%, delay 100 ms	E
Narrowband Noise 1	SNR=30 dB narrowband Gaussian noise, center frequency distribution in 0 4 kHz	G.N1
Narrowband Noise 2	SNR=50 dB narrowband Gaussian noise, center frequency distribution in 0 4 kHz	G.N2
MP3 Compression 1	Re-encoded as MP3, and then decoding recovery, the rate is 32 k	M.32
MP3 Compression 2	Re-encoded as MP3, and then decoding recovery, the rate is 128 k	M.128

Table 7: The BER mean and maximum value of different algorithms after content preserving operations

CPO	Proposed		[12]		[19]		[6]	
	Mean	Max	Mean	Max	Mean	Max	Mean	Max
V.1	**0.0013**	**0.0175**	0.0070	0.0333	0.0479	0.1624	0.0064	0.0583
V.2	**0.0034**	**0.0313**	0.0090	0.0630	0.0179	0.0940	0.0069	0.0573
F.I.R	0.0751	0.1103	**0.0413**	**0.1000**	0.0813	0.1851	0.0892	0.1673
B.W	0.0851	**0.1441**	0.0884	0.1593	**0.0784**	0.1667	0.0974	0.2096
R.8→16	0.0328	0.1040	0.0529	0.1185	**0.0263**	**0.0470**	0.0379	0.1945
R.32→16	**0.0033**	**0.0326**	0.0048	0.0444	0.0094	0.0149	0.0034	0.0367
E	0.0561	0.0890	**0.0315**	**0.0778**	0.1026	0.2051	0.0650	0.1024
G.N1	**0.0830**	**0.1855**	0.1108	0.2333	0.1528	0.1992	0.0909	0.2096
G.N2	0.0128	**0.0526**	0.0391	0.1259	0.0709	0.1005	**0.0112**	0.0677
M.32	**0.0458**	**0.0977**	0.2535	0.3370	0.1097	0.1838	0.0512	0.1372
M.128	**0.0027**	**0.0238**	0.2315	0.3148	0.0248	0.0855	0.0032	0.0367

is not correlated. Therefore, the proposed algorithm has good security. By comparing Figure 10(b) and 10(c), ρ after correct decryption is constant 1, and ρ after error decryption is completely different from 10(a), indicating that the proposed algorithm can improve the encryption performance before and after chaotic mapping.

In conclusion, the algorithm makes the BioHashing sequence disordered, thereby improving the security of the BioHashing sequence in the cloud.

4.4 Tampering Detection and Location

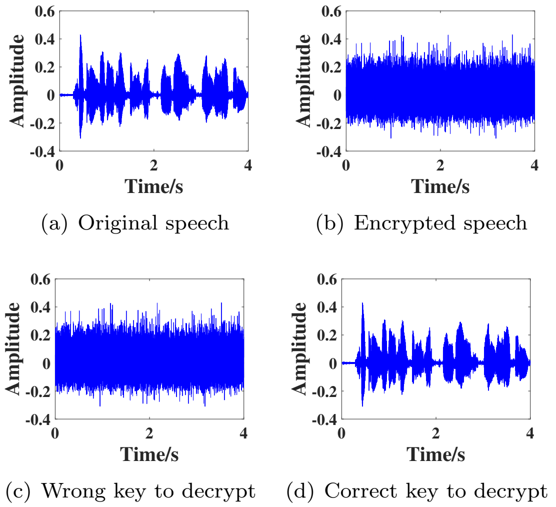
For small-scale tampering attacks, it is generally to tamper with the local part of speech, with small tampering range and low bit error rate. To solve the above problems, this paper proposes a tamper detection and localization algorithm based on Hamming code minimum code distance (MCD), which is defined as Equation (22):

$$MCD(i) = \begin{cases} 1, h(i) \neq h'(i) \\ 0, h(i) = h'(i) \end{cases} \tag{22}$$

Where, $h'(i)$ is the i-th sampling value of the BioHahing sequence of the tampered speech, and MCD matrix form is as shown in Equation (23):

$$MCD(i) = [MCD(1) \quad MCD(2) \quad \cdots \quad MCD(i)] \tag{23}$$

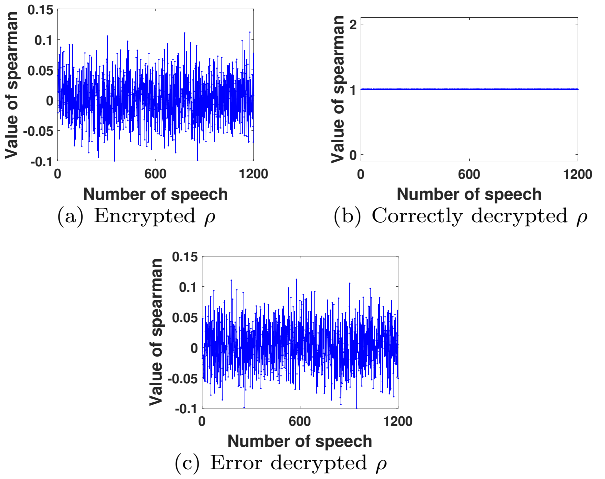
In this paper, the duration of the speech clip is 4s. If it is a malicious attack of 1 %, 5 % and 10 %, the duration of the malicious attack is 0.04s, 0.2s and 0.4s respectively. A speech clip is randomly selected from the original speech library. In order to measure the detection and location



(a) Original speech (b) Encrypted speech

(c) Wrong key to decrypt (d) Correct key to decrypt

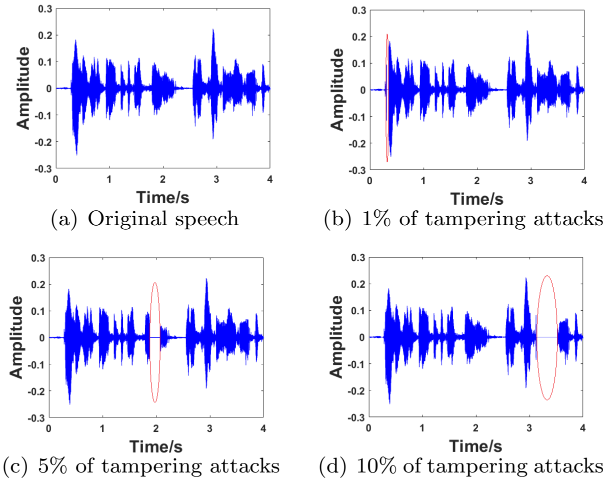
Figure 9: Waveform plot before and after encryption



(a) Encrypted ρ (b) Correctly decrypted ρ

(c) Error decrypted ρ

Figure 10: Spearman's correlation coefficient



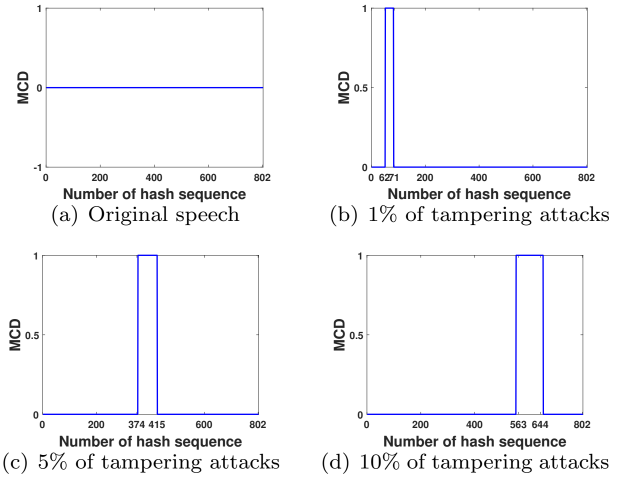
(a) Original speech (b) 1% of tampering attacks

(c) 5% of tampering attacks (d) 10% of tampering attacks

Figure 11: Tamper detection and location of speech clip

ability of MCD algorithm for small range tampering, the speech clip is maliciously attacked. As shown in Figure 11, the speech clip can be better detected the location and range of tampering attacks.

As shown in Figure 12, the BioHahing sequence constructed by this speech clip can also be better detected the location and range of tampering attacks.



(a) Original speech (b) 1% of tampering attacks

(c) 5% of tampering attacks (d) 10% of tampering attacks

Figure 12: Tamper detection and location of BioHahing sequence

In order to better measure the overall detection performance of the algorithm, Figure 11 and Figure 12 are combined as shown in Table 8.

In Table 8, the unit of speech clipt tampering is sampling point, and the unit of BioHashing sequence tampering is bit. Table 8 describes that the experimental value and theoretical value of the algorithm are equal, which indicates that the overall detection performance of the algorithm is well. The range of Biohashing sequence tampering is same as the scale marked in Figure 10, and the range and length of BioHashing sequence tampering corresponds to the range and length of speech clip tampering.

On the whole, the proposed algorithm can not only realize the tamper detection of the speech clip in a small range, but also accurately locate the location and range of the BioHashing sequence attacked by tampering attacks.

4.5 Efficiency Analysis

Algorithm efficiency refers to the execution time of the algorithm. The shorter the time, the simpler the structure and the lower the algorithm complexity. To measure the efficiency of the proposed algorithm, 200 speech clips were randomly extracted from the original speech library and the average running time was calculated as shown in Table 9.

Table 8: Tamper detection and location different lengths and ranges

Speech length	Speech range	BioHasing length	BioHasing range	Theoretical value	Experimental value
640	4999-5640	8	63-70	1%	1%
3200	29999-33200	40	375-414	5%	5%
6400	49999-56400	80	564-643	10%	10%

Table 9: Efficiency of the algorithm

Algorithm	Average time/s
The F-cE algorithm	0.0370s
The F-cE algorithm +K-means	0.2874s
The F-cE algorithm+ K-means-KNN	0.4474s

In Table 9, $K = 3$ and K-means cluster 160 speech clips randomly. It can be obtained from Table 9 that the average running time of the proposed algorithm is relatively long, but the feature extraction structure is simple and the calculation amount is small.

To further measure the efficiency of the proposed algorithm, 200 speech clips were randomly extracted from the original speech library and the average running time of different algorithms at the same main frequency was calculated, as shown in Table 10.

Table 10: Efficiency of different algorithms

Algorithm	Main frequency /GHz	Average time /s
The F-cE algorithm+ K-means-KNN	3.4GHz	0.4474s
[7]	3.4GHz	0.0439s
[21]	3.4GHz	0.0848s
[19]	3.4GHz	0.1825s
[6]	3.4GHz	0.0788s

As can be seen from Table 10, compared with other algorithms, the proposed algorithm efficiency is lower. Complexity increased because of the machine learning algorithm calculates the similarity, according to the extracted features, resulting in the increase of the computation amount and the average running time. However, the algorithm meets the efficiency of real-time speech communication. Therefore, our proposed algorithm meets the expected improvement requirements.

5 Conclusions and Future Work

In this paper, the proposed algorithm can effectively classify the extracted features and has excellent overall performance. The main advantages are summarized as follows: **A**. The algorithm has good discrimination and robustness than other algorithms. **B**. The algorithm uses different secret keys to encrypt the results of K-means and KNN classification, *i.e.*, it has good security. **C**. The algorithm can realize small range of tamper detection and location.

The robustness of low pass filtering is poor. Therefore, as a future work, we plan to study some solutions to improve the robustness after pass low pass filtering content preserving operations.

Acknowledgments

This work is supported by the National Natural Science Foundation of China(No. 61862041), Youth Science and Technology Fund of Gansu Province of China(No. 1606RJYA274).

References

[1] C. Albin, D. Narayan, R. Varu, and V. Thanikaiselvan, ”DWT based audio encryption scheme,” in *2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA'18)*, pp. 920–924, 2018.

[2] A. Awais, S. Kun, Y. Yu, S. Hayat, A. Ahmed, and T. Tu, ”Speaker recognition using mel frequency cepstral coefficient and locality sensitive hashing,” in *International Conference on Artificial Intelligence and Big Data (ICAIBD'18)*, pp. 271–276, 2018.

[3] N. Chen and H. Xiao, ”Perceptual audio hashing algorithm based on zernike moment and maximum-likelihood watermark detection,” *Digital Signal Processing*, vol. 23, no. 4, pp. 1216–1227, 2013.

[4] A. Chowdhury and A. Ross, ”Fusing MFCC and LPC features using 1D triplet cnn for speaker recognition in severely degraded audio signals,” *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 1616–1629, 2019.

[5] J. Han, J. Xu, F. Nie, and X. Li, ”Multi-view k-means clustering with adaptive sparse memberships and weight allocation,” *IEEE Transactions on Knowledge and Data Engineering*, pp. 1-1, 2020. DOI: 10.1109/TKDE.2020.2986201.

[6] Y. Huang, H. Hou, Y. Wang, Y. Zhang, and M. Fan, ”A long sequence speech perceptual hashing authentication algorithm based on constant q transform

and tensor decomposition," *IEEE Access*, vol. 8, pp. 34140–34152, 2020.

[7] Y. Huang and Y. Wang, "Multi-format speech perception hashing based on time-frequency parameter fusion of energy zero ratio and frequency band variance," in *The 3rd International Conference on Electronic Information Technology and Computer Engineering (EITCE'19)*, pp. 243–251, 2019.

[8] Y. Huang, Y. Wang, Q. Zhang, W. Zhang, and M. Fan, "Multi-format speech biohashing based on spectrogram," *Multimedia Tools and Applications*, vol. 79, pp. 24889—24909, 2020.

[9] Y. B. Huang, Y. Wang, Q. Y. Zhang, and H. X. Hou, "Multi-format speech perception hashing algorithm based on short-time logarithmic energy and improved mel energy parameter fusions," *International Journal of Network Security*, vol. 22, no. 6, pp. 1043–1053, 2020.

[10] B. Jiang and D. Yuan, "Improved knn rss fingerprint localization based on spherical directional antenna array," in *IEEE 7th International Conference on Computer Science and Network Technology (ICCSNT'19)*, pp. 427–431, 2019.

[11] X. Li, J. Peng, M. S. Obaidat, F. Wu, M. K. Khan, and C. Chen, "A secure three-factor user authentication protocol with forward secrecy for wireless medical sensor network systems," *IEEE Systems Journal*, vol. 14, no. 1, pp. 39–50, 2020.

[12] J. Li, W. Tao, and H. Wang, "Perceptual hashing based on correlation coefficient of MFCC for speech authentication," *Beijing University of Posts and Telecommunications*, vol. 38, no. 2, pp. 89–93, 2015.

[13] J. Li and T. Wu, "Perceptual audio hashing using rt and dct in wavelet domain," in *The 11th International Conference on Computational Intelligence and Security (CIS'15)*, pp. 363–366, 2015.

[14] Q. Li, Y. Yang, T. Lan, H. Zhu, Q. Wei, F. Qiao, X. Liu, and H. Yang, "MSP-MFCC: Energy-efficient MFCC feature extraction method with mixed-signal processing architecture for wearable speech recognition applications," *IEEE Access*, vol. 8, no. 48720–48730, 2020.

[15] C. Shi, X. Li, and H. Wang, "A novel integrity authentication algorithm based on perceptual speech hash and learned dictionaries," *IEEE Access*, vol. 8, pp. 22249–22265, 2020.

[16] L. Wu, Y. Ma, Z. Peng, and W. Zheng, "Review of biometric template protection," *Chinese Journal of Scientific Instrument*, vol 37, no. 11, pp. 2407–2420, 2016.

[17] H. Yao, C. Wang, X. Fu, C. Liu, B. Wu, and F. Li, "A privacy-preserving rlwe-based remote biometric authentication scheme for single and multi-server environments," *IEEE Access*, vol. 7, pp. 109597–109611, 2019.

[18] Q. Zhang, W. Hu, S. Qiao, and Y. Huang, "Speech perceptual hashing authentication algorithm based on spectral subtraction and energy to entropy ratio," *International Journal Network Security*, vol. 19, no. 5, pp. 752–760, 2017.

[19] Q. Zhang, T. Zhang, S. Qiao, and D. Wu, "Spectrogram-based efficient perceptual hashing scheme for speech identification," *International Journal Network Security*, vol. 21, no. 2, pp. 259–268, 2019.

[20] Q. Zhang, W. Hu, Y. Huang, and S. Qiao, "An efficient perceptual hashing based on improved spectral entropy for speech authentication," *Multimedia Tools and Applications*, vol. 77, no. 2, pp. 1555–1581, 2018.

[21] Q. Zhang, S. Qiao, Y. Huang, and T. Zhang, "A high-performance speech perceptual hashing authentication algorithm based on discrete wavelet transform and measurement matrix," *Multimedia Tools and Applications*, vol. 77, no. 16, pp. 21653–21669, 2018.

[22] Q. Zhang, L. Zhou, T. Zhang, and D. Zhang, "A retrieval algorithm of encrypted speech based on short-term cross-correlation and perceptual hashing," *Multimedia Tools and Applications*, vol. 78, no. 13, pp. 17825–17846, 2019.

Biography

Yi-Bo Huang received the Ph.D candidate degree form Lanzhou University of Technology in 2015, and now working as an Associate Professor in the college of physics and electronic engineering in Northwest Normal University. He main research interests include Multimedia information processing, information security, speech recognition.

Hao Li received the BS degrees in Changzhou University Huaide College, Changzhou, China, in 2019. His research interests include speech signal processing and application, BioHashing authentication techniques.

Yong Wang received the BS degrees in Henan Institute of Science and Technology, Henan, China, in 2017. His research interests include audio signal processing and application, multimedia authentication techniques.

Security Situation Assessment Algorithm for Industrial Control Network Nodes Based on Improved Text SimHash

Rui-Hong Dong, Chuang Shu, and Qiu-Yu Zhang
(*Corresponding author:Qiu-Yu Zhang*)

School of Computer and Communication,Lanzhou University of Technology
No.287,Lan-Gong-Ping Road,Lanzhou 730050,China
Email: zhangqylz@163.com

(Received July 22, 2020; Revised and Accepted Apr. 24, 2021; First Online Oct. 19, 2021)

Abstract

Aiming at the problems that the existing network security situation assessment methods are weak in dealing with the complex and diverse attack types and difficult to adapt to the big data industrial control network environment, a security situation assessment algorithm for industrial control network nodes based on improved text SimHash was proposed. Firstly, the algorithm constructs the pre-attack and post-attack text according to the industrial control network security data obtained by node attack detection. Secondly, the improved SimHash algorithm is applied to calculate the similarity of the constructed text. Finally, the text-similarity is used to quantify the security situation value of industrial control network nodes. Experimental results show that the proposed method can assess network security situations more accurately. In addition, compared with existing methods, the proposed algorithm can improve the ability to deal with complex network attack types and can effectively adapt to the big data industrial control network environment.

Keywords: Improved SimHash Algorithm; Industrial Control Network; Node Attack Detection; Security Situation Assessment; Text Similarity

1 Introduction

With the rise of concepts and technologies such as industry 4.0, ”Made in China 2025” and cyber-physical fusion system, industrial control system (ICS) has been widely applied in nuclear and thermal power plants, water treatment facilities, power generation, heavy industry and power distribution systems, and is developing towards the direction of digitalization, networking and intelligence [11]. Although ICS has been isolated from the Internet for a long time, significant commercial interests are promoting the integration of traditional physically isolated industrial control networks with operating systems , communication protocols and other technologies in the IT network field, leading to the huge security threats to the industrial control network [2]. In recent years, new network attack technologies have emerged in an endless stream, network vulnerability issues have become increasingly prominent, and network information security is facing severe challenges. Accurate and efficient assessment of the target network security state is of great significance to the stable and safe operation of the network, and has become a research focus in the field of industrial control network security [12].

The successive ICS network security incidents in recent years have shown that critical industrial infrastructure can no longer be protected from targeted IT attacks [1]. Due to the inherent differences between industrial control networks and IT networks, the security situation assessment standard of industrial control network cannot copy the IT network completely. IT network security situation assessment research is relatively mature, while the research on industrial control network security situation assessment has not formed the architecture yet. Therefore, the research of industrial network security situation assessment method can be used for reference from IT network security situation assessment technology. Information fusion is the core of network security situation assessment. Through comprehensive analysis and understanding of network security-related issues, the state of network security can be mastered. There are many existing assessment methods, such as DS evidence theory, Hidden Markov model, Bayesian technology, neural network, *etc.* [5].

Lu *et al.* [17] proposed a security situation awareness framework based on particle filter, which has a good state estimation capability for nonlinear systems and can accurately evaluate system situation. Zhu *et al.* [27] introduced three kinds of network security situational awareness models: network security situational awareness model based on neural network, network security situational awareness model based on random forest and network security situational awareness model based on star structure. Each model is composed of different machine learning algorithms to achieve different functions. Zhao *et al.* [25] studied a network security situation assessment model based on D-S evidence theory for the problems that the assessment information source is too single and the accuracy deviation is too large in the situation awareness system. Wang *et al.* [21] brought up a network security situation assessment model and quantitative method based on AHP, aiming at the problems existing in the hierarchical network security situation assessment model, such as large subjective index weight factor, large evaluation index system, large calculation amount and low efficiency. Lin *et al.* [15] proposed an assessment model based on SimHash in the big data environment, aiming at the high complexity of the existing network security situation assessment methods and poor effect in the big data environment. In order to improve the objectivity and comprehensibility of the evaluation results, based on fuzzy theory, particle swarm optimization and RBF neural network, Yi *et al.* [24] proposed a network security risk assessment model based on fuzzy theory. Dong *et al.* [5] combined cuckoo search algorithm and BP neural network, proposed a quantitative evaluation method of adaptive learning, which solved the problems of low efficiency and poor reliability of existing network security situation assessment methods.

In [7], BP neural network optimized by improved adaptive genetic algorithm was established for APT attacks to predict the security risks of network nodes, and APT attack paths and maximum possible attacks were calculated. To accurately perceive the current network security situation and timely detect the attack behavior, Lv *et al.* [18] proposed a multi-scale risk assessment model for network security based on Long Short Term Memory neural network (LSTM). Aiming at the problem that the current network security situation assessment method only focuses on attack behavior, a network security situation assessment method based on Markov decision process and game theory is proposed in [13]. Qiang *et al.* [19] proposed a D-S evidence theoretical situation assessment model based on an optimized CS-BP neural network. This model improves the local search capability of cuckoo algorithm by conjugate gradient calculation, and introduces it into BP neural network, which improves the convergence speed of training and overcomes the local minimum problem. In [4], a network security situation assessment model (GSA-SVM) based on GSA optimization was proposed to solve the problem of insufficient parameter selection accuracy of support vector machine (SVM) in situation assessment.

At present, the domestic and foreign research on network security situation is mainly divided into two aspects:

1) Instantiation of the model and method of situation assessment in the field of network security situation perception, and continuous testing and optimization in practice, mainly based on the improvement of Endsley model [6], JDL model [9] and other typical models;

2) Breakthrough in model algorithm optimization and application level, especially in the aspect of quantitative situation calculation and perception [14].

With the development of new technologies such as cloud computing and machine learning, network security assessment technology is developing towards the direction of intelligence, integration and scale [5]. The latest research direction is to combine network security situation assessment with artificial intelligence to evaluate network security status through multi-layer nonlinear fitting. At present, the research in this field is in its infancy. The high-dimensional and non-linear data in the offensive and defensive process is abstracted, and the network security situation value is calculated through intelligent methods, both in feasibility and optimality, need to be tested and improved in practice [14].

By analyzing the above research, the existing research provides a series of feasible solutions for network security situation assessment, but there are still some problems. For instance, it can dynamically and comprehensively display the current network security situation as a whole, but the assessment is highly subjective, the assessment process is complicated, the assessment is poor in operability, the ability to deal with complex network attacks is weak, and it is difficult to accurately assess the security situation when the network scale is large. To solve these problems, this paper proposes a security situation assessment algorithm for industrial control network nodes based on improved text SimHash, which can deal with complex and diverse network attack types and adapt to big data network environment effectively. The main contributions of this paper can be summarized as follows:

1) Using Louvain algorithm to detect community structure in complex networks to divide the large-scale network into modules.

2) Intrusion detection equipment, traffic monitoring equipment and vulnerability scanning system are exploited to obtain security data and vulnerability information on each node of industrial control network.

3) The improved SimHash algorithm is adopted to process the obtained security data and evaluate the security situation of nodes quickly and efficiently.

The remaining part of this paper is organized as follows. Section 2 introduces the network security situation assessment index system. Section 3 describes in detail the improvement of security situation assessment algorithm and its SimHash algorithm. In Section 4, the security situation assessment algorithm is experimentally verified and compared with the existing methods. Section 5 concludes the presented work and raises several issues of future work.

2 Construction of Network Security Situation Assessment Index System

2.1 Indicators of Situation Assessment

In order to achieve an accurate assessment of the network security situation and to facilitate the description of the assessment results, based on the vulnerability information of each node in the network, in view of the complexity and heterogeneity of network operation data, as well as the relevance of various impact indicators, a situation index assessment factor is proposed. Including: attack quantity, attack frequency, attack threat, vulnerability quantity, vulnerability threat, vulnerability probability of exploitation, and other six factors, and give the relevant definition.

Definition 1. *Attack Quantity Factor: Defined as the number of attack types detected within a certain period of time, denoted as N.*

Definition 2. *Attack Frequency Factor: Defined as the attack frequency of each type of attack in a certain period of time, denoted as C_i (i represents different attack types).*

Definition 3. *Attack Threat Factor: Different attack types have different degrees of impact on network security operations, denoted X_i (i represents different attack types).*

Definition 4. *Vulnerability Quantity Factor: Defined as the number of vulnerability types scanned on a network node, denoted as $M_i d_v$.*

Definition 5. *Vulnerability Threat Factor: Defined as the impact of each vulnerability on the normal operation of the network, denoted as $impact_v$.*

Definition 6. *Probability Factor: Defined as the probability of each vulnerability being successfully exploited, denoted as p_{suc}.*

2.2 Assessment of Indicator Factor Measurement

For each type of attack, the threat index of the assessment indicator at time t is:

$$F_i(t) = f(N, C_i(t), X_i(t)) = \frac{C_i(t)}{N} \times 10^{X_i(t)}$$

where $F_i(t)$ represents the threat index of a certain attack type in the time period t.

2.3 Classification of Network Security Assessment Levels

This paper refers to the basic situation index of network security of The National Internet Emergency Response Center, and combines the characteristics of network threats, vulnerabilities and other elements, divides the network security situation into four levels, namely, severe risk, moderate risk, mild risk and security. In order to facilitate the intuitive analysis of the results of network security situation assessment, the 0-1 numerical method is used to quantitatively represent each security level. As show in Table 1.

3 The Proposed Security Situation Assessment Algorithm

3.1 Data Preprocessing

The data obtained from various data sources (such as logs and network traffic) have different formats, fast generation speed, and contain dirty data, which all lead to inefficient data exchange and sharing. At the same time, today's sophisticated network-attacks that occur across multiple dimensions and stages, and traditional platforms will have no opportunity to defend a network. Aiming at these problems, the idea of complex networks [22] is introduced into the network security situation assessment.

In this paper, Louvain algorithm [3] is used to classify a large-scale network. Louvain is a graph algorithm model based on Modularity. Different from the ordinary graph based on Modularity and Modularity gain, this algorithm has a high speed and has a particularly obvious clustering effect for graphs with multiple vertices and few edges [8]. The definition of Modularity is the value Q that describes the degree of closeness within the community, which is defined as Equation (1):

$$Q = \frac{1}{2m}\sum_{i,j}[A_{ij} - \frac{k_i k_j}{2m}]\delta(c_i, c_j) \tag{1}$$

where

$$\delta(c_i, c_j)=\begin{cases} 1 & c_i = c_j \\ 0 & \text{otherwise} \end{cases},$$

A_{ij} represents the weight of edge between node i and node j. When the network is not weighted graph, the weight of all edges can be regarded as 1. $k_i = \sum_j A_{ij}$ represents the sum of the weights of all edges connected to node i; c_i represents the community to which the node belongs; $m = \frac{1}{2}\sum_{i,j} A_{ij}$ represents the sum of the weights of all the edges (number of edges). In the equation, $A_{ij} - \frac{k_i k_j}{2m} = A_{ij} - k_i \frac{k_j}{2m}$, the probability that node j is connected to any node is $\frac{k_j}{2m}$, and now node i has the degree of k_i, so the edge between node i and j is $k_i \frac{k_j}{2m}$ in the random case. The algorithm in this paper uses the positive and negative values of the Modularity gain ΔQ to determine whether the i-th node joins the module to which the j-th node belongs. The definition of ΔQ is shown in Equation (2):

$$\begin{aligned} \Delta Q \quad = \quad & [\frac{\sum_{in} + 2k_{i,in}}{2m} - (\frac{\sum_{tot} + k_i}{2m})^2] \\ & -[\frac{\sum_{in}}{2m} - (\frac{\sum_{tot}}{2m})^2 - (\frac{k_i}{2m})^2] \end{aligned} \tag{2}$$

Table 1: Classification of network security levels

Security index	Security level	Network operation
0∼0.2	Safe	Network operation is normal
0.2∼0.5	Mild risk	Network operation was slightly affected
0.5∼0.8	Moderate risk	Network operation was greatly damaged
0.8∼1	Severe risk	Network operation suffered a serious security incident

In order to greatly reduce the complexity, Equation (2) can be reduced to the form of Equation (3), but it should be noted that the simplified calculation results are relative gain rather than absolute gain.

$$\Delta Q^{'} = k_{i,in} - \frac{\sum_{tot} \times k_i}{m} \quad (3)$$

where $k_{i,in}$ represents the sum of the weights of incident cluster C by node i, $\sum_{tot}$ represents the total weight of incident cluster C, and k_i represents the total weight of incident node i.

Figure 1 shows an example of the network partition by Louvain algorithm. The specific process of Louvain algorithm is as follows:

Step 1: Initially treat each vertex in the graph as a community, with the same number of communities as the vertices.

Step 2: Merging each vertex with its adjacent vertex in turn, calculating whether their Modularity gain is greater than 0, if so, put the node into the community where the adjacent node is located.

Step 3: Iterating the second step until the algorithm is stable, that is, all vertices belong to communities that do not change.

Step 4: Compressing all nodes in each community into a single node. The weight of nodes in the community is converted into the weight of new node ring, and the weight of communities is converted into the weight of new node edge.

Step 5: Repeating **Steps 1-3** until the modularity of the entire figure no longer changes.

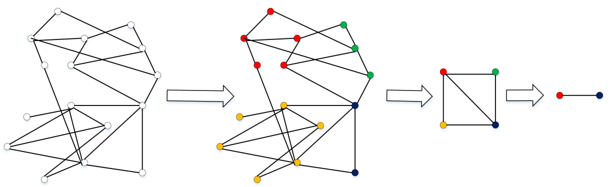


Figure 1: An example of complex network partitioning

As can be seen from Figure 1, assuming that there are 16 nodes in the network. Initially, each network node is regarded as a community, and the weight of the edges between the nodes is determined according to the degree of closeness of communication between the nodes. A random network node is selected as the initial node, and it is attempted to merge the node and adjacent nodes in a community in turn until all vertices belong to a community that does not change. It can be seen from the second stage of Figure 1 that after the first round of node classification, four large communities of red, yellow, blue and green are formed. The four large communities are then compressed into a single node, as shown in the third stage of Figure 1. Finally, repeat the above steps to further compress the community structure to form the red and blue communities shown in stage 4 of Figure 1. At this point, a large-scale network is divided into smaller network structures by the Louvain algorithm.

The multi-source heterogeneous data generated within each partition module cannot be directly used for network security situation assessment, and these data need to be integrated. The purpose of data integration is to organize the data in each independent system into a whole according to certain rules by certain technical means, so that other systems or users can effectively access the data. After the multi-source heterogeneous data is uploaded to the distributed file system, the format of multi-source heterogeneous data is unified, a large number of noisy data irrelevant to network security situation assessment is eliminated, and redundant attribute data is merged. Finally, these pre-processed data are stored in the database as data that can be directly used for network security status assessment.

3.2 Node Security Situation Assessment Model based on Improved Text SimHash

ICS communication networks usually contain a large number of hosts, network devices, and various detection systems that monitor the network from different perspectives and generate logs and alerts. The traditional network security situation assessment method has weak ability to deal with complex and diverse industrial control network attack methods and attack types, and it is difficult to adapt to the big data industrial control network environment. In addition, the existing evaluation methods tend to use more complex algorithm, directly affects the evaluation of timeliness, delayed the network administrator take measures of best time. In response to these problems, this paper proposes a security situation assessment algorithm for industrial control network nodes based on improved text SimHash. Figure 2 shows the framework of the assessment algorithm.

It can be seen from Figure 2 that the model mainly includes the following assessment steps:

Step 1: Using the algorithm (Louvain algorithm) to detect community structure in the complex network to divide the large-scale industrial control network and obtain the module and its weight.

Step 2: Collecting network security situation elements, and then upload to distributed file system for storage.

Step 3: Attack types and attack times are obtained by preprocessing the elements of network security situation.

Step 4: Inside the module, the vulnerability of nodes is scanned and the probability of success of each attack type is calculated.

Step 5: For each node in the module, the severity of the attack is calculated using an improved SimHash algorithm based on the type and number of attacks.

Step 6: Attack threat value (the same as the attack threat factor in Section 2.1) refers to the threat degree of different types of attacks to the network, that is, the depth at which the attack enters the target network.

Step 7: Calculating the security situation value of the node according to the attack threat value, attack severity and probability of success.

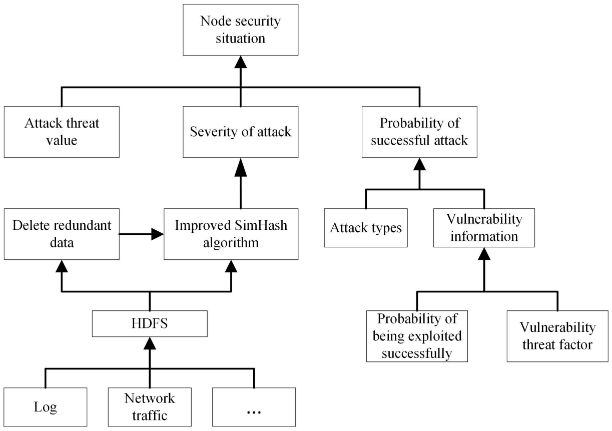

Node security situation
Attack threat value
Severity of attack
Probability of successful attack
Delete redundant data
Improved SimHash algorithm
Attack types
Vulnerability information
HDFS
Probability of being exploited successfully
Vulnerability threat factor
Log
Network traffic
...

Figure 2: Node security situation assessment model based on improved SimHash

3.3 Improved Text SimHash Node Security Situation Assessment Algorithm

3.3.1 SimHash Algorithm

SimHash [20] is a fingerprint extraction algorithm proposed by Google in 2007. It is essentially a dimensionality reduction method that maps high-dimensional vectors to smaller fingerprints, so that the original vector features can also be retained. Generally, the signature fingerprint generated by the SimHash algorithm can be 32-bit, 64-bit, 128-bit, *etc.* [16]. SimHash is a local sensitive hash (LSH), which means that each bit of the hash value reflects the local change of the input source data. A 64-bit SimHash value can remember 64 features of the text, and the hamming distance between two SimHash hash values can tell the similarity of the contents of the two texts [23]. It can be seen that SimHash algorithm is an efficient algorithm for calculating text similarity.

The concrete processing steps of SimHash algorithm are as follows:

Step 1: Inputing a text and then generate an N-dimensional text feature vector V, with each feature having a certain weight.

Step 2: Initializing a C-dimensional vector Q with initial values of 0 and C bits binary signature S of 0.

Step 3: For each feature of vector V, use the hash algorithm to calculate a C bits hash value H.

Step 4: For any $i \in [1, C]$, if the i-th bit of H is 1, then the weight of the feature is added to the i-th of Q, otherwise, it is reduced.

Step 5: If the i-th dimensional element of final Q is greater than 0, the i-th dimensional element of S is 1, otherwise it is 0.

Step 6: The final C-dimensional binary signature S is the binary signature of the text.

3.3.2 Text Processing

Before using the SimHash algorithm for node security assessment, text needs to be constructed as input to the algorithm. First, the text is randomly generated. The words that make up the text will not repeat each other, and the number of words that make up the text is related to the total number of attacks during that time. Then assign different words to different types of attacks. These words do not overlap with words in the randomly generated text. Finally, extract the attack information over a period of time and calculate the number of attacks for different types of attacks. If there is a type of attack, a copy of the original text will be generated. For each type of attack, according to the number of attacks, replace some words in the copy with the specified words.

If multiple attacks occur over a period of time, several revised texts will be obtained. The SimHash algorithm is used to obtain multiple hash values corresponding to the modified text, and these hash values are compared with the hamming distance generated by the original text. Hamming distance is the number of different bits between two m-bit hash values, which can be used to evaluate the similarity between two vectors. The larger the Hamming distance is, the smaller the similarity between the two vectors will be. This feature can be used to quantify the attack severity of a certain type of attack on this node over a period of time. Depending on the vulnerability of nodes, different types of attacks have different probability of success. If there are many types of attacks over a period of time, we will get several Hamming distances. In addition, the attack threat values of different types of attacks are also different, that is, different types of attacks have different degrees of impact on network security operation. This paper combines the text similarity, the corresponding attack success rate probability and the attack threat value to get the final node security situation value.

3.3.3 Improved Text SimHash Assessment Algorithm

In general, SimHash algorithm is used by search engines such as Baidu and Google for webpage deduplication and document similarity detection, which greatly reduces computation. Due to the massive data generated by various security devices in large networks, an efficient network security situation assessment algorithm is urgently needed to enable network administrators to quickly understand the current security state of the network. However, the actual evaluation process of the existing assessment algorithm is relatively complicated, and it takes a lot of time to calculate. Therefore, there are certain limitations in the application in large-scale network environments [15]. In order to solve these problems, the SimHash algorithm is introduced into the network security situation assessment.

In the traditional SimHash de-duplication algorithm, hamming distance is mainly used to determine whether the data is repeated. However, due to its own characteristics, even if two pieces of data are completely unrelated, hamming distance will be relatively small. If the two pieces of data are similar, there must be a considerable degree of repetition in the text content [26]. This paper first defines the repetition rate of two pieces of data:

$$dup(A,B) = \frac{cnt(tok(A) \bigcap tok(B))}{cnt(tok(A) \bigcup tok(B))}$$

where $tok(A)$ and $tok(B)$ respectively represent the set of data A and B after word segmentation, and the repetition rate is defined as the ratio of the number of repeated phrases between the two sets and the total number of phrases included. Repetition of a phrase in one of the sets itself is not recorded as repetition.

Considering the Hamming distance and repetition rate, the similarity of the two data is defined as follows:

$$Sim(A,B) = \frac{1}{Ham(A,B)+1} + K \cdot dup(A,B)$$

where K is a parameter, generally between 2 and 8.

With the definition of similarity, also need to set a threshold to determine whether the data is similar. Due to the different types of data sets, the threshold should be defined according to the actual effect of the test data, but the general recommendation is not less than 1+0.2 K, that is, even if the Hamming distance is 0, a repetition rate of 20% is still required to confirm that the two texts are similar, and when the Hamming distance is larger, a higher repetition rate is required. It does not matter even if the Hamming distance is large when the repetition rate is high [26].

Based on the improved SimHash algorithm, the node security situation assessment algorithm is proposed. First, use the text processing described above to generate pre-attack and post-attack text. Then, these texts are used to quantify the severity of attacks, and finally, the security status of nodes is quantified. The algorithm is shown in Algorithm 1.

Algorithm 1 Node security situation assessment algorithm based on improved text SimHash

```
Begin
$b \Leftarrow$ The length of the hash value.
$d \Leftarrow 0$
$T_1 \leftarrow$ Randomly generate n words of equal length, each word is not repeated.
$F_1(t) \Leftarrow$ The eigenvector on $T_1$.
$H_1 \Leftarrow SimHash(T_1, b)$
for $a \in A$ and $i = 0$ to $n$ do
    $word \Leftarrow$ Randomly generate a word.
    $T_2 \Leftarrow replace(T_1, i, word)$
end for
$F_2(t) \Leftarrow$ The eigenvector on $T_2$
$H_2 \Leftarrow SimHash(T_2, b)$
for $i = 1$ to $b$ do
    if $H_1[i] \neq H_2 i$ then
        $d \Leftarrow d + 1$
    end if
end for
$Sim(T_1, T_2) = \frac{1}{d} + K \cdot dup(T_1, T_2)$
for $a \in A$ do
    if $id \in V$ then
        $NSA \Leftarrow NSA + impact_v \cdot \frac{1}{Sim(T_1,T_2)} \cdot p_s uc$
    else
        $NSA \Leftarrow NSA + 0$
    end if
end for
return $NSA$
End
```

In the actual network, due to its own loopholes, even if it is not attacked, the network still faces security risks. Scan the target network with scanning tools such as Nessus to obtain vulnerability information. Based on the vulnerability information, the probability of success of various attacks is obtained. However, there are many attacks against vulnerabilities, so in the statistical process of attack

$$NSA = \sum_{i=1}^{n} (\frac{1}{Sim_i} \cdot X_i \cdot p_i).$$

A network node may be subjected to multiple types of attacks at the same time, $\frac{1}{Sim_i}$ represents the severity of the i-th attack on this node, and p_i represents the probability of success of the i-th attack based on the node's vulnerability information, with the value of 0 or 1.

4 Experimental Results and Analysis

4.1 Simulation Experiment

4.1.1 Establishment of Experimental Environment

In order to verify the effectiveness and efficiency of the proposed algorithm, an experimental environment as shown in Figure 3 is set up for verification.

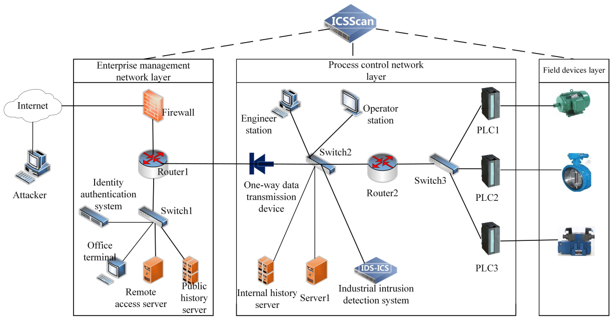


Figure 3: Experimental network topology

It can be seen from Figure 3 that the process control network is the core of the industrial control system, mainly for the user to monitor and control the entire system's industrial production process through the human-computer interaction interface. The operator through the use of the supervisory control and data acquisition system can run at the scene of the production equipment for data collection and reporting, has an extremely high reliability. In this experiment, the simulated attack machine will launch an attack on the industrial control network. The industrial control intrusion detection system will be used to collect the attack information of the switches in the process control network layer. At the same time, the vulnerability scanning system is used to scan the enterprise management network layer and process control network layer of the industrial control system successively to obtain the vulnerability information in the industrial control network.

In order to test the criticality of computer vulnerability information in network security situation assessment, node vulnerability information is matched with attack-dependent vulnerability information to analyze the feasibility of the attack. In the simulation experiment, it is assumed that there are four typical vulnerabilities in the network nodes. The vulnerability information is shown in Table 2.

Because the industrial control network environment involves many network nodes, it is not clear enough to draw a complete network topology. Therefore, suppose the network topology shown in Figure 3 is one of the communities after clustering a large industrial control network by the Louvain algorithm. All the nodes in the community are compressed into a node, and the weight of the nodes in the community is converted into a new node. The weight of the ring and the weight between the communities are converted into the weight of the new node, so the entire community can be regarded as a new node.

Table 2: Vulnerability information

Vulnerability type	$impact_v$	p_{suc}
SYN Flood	0.9	0.7
Sadmind Buffer Overflow	0.8	0.8
RPC	0.2	0.8
Remote Login	0.1	0.6

4.1.2 Experimental Dataset

The NSL-KDD dataset solves the inherent problems in the KDD99 dataset. The setting of NSL-KDD training set and test set is reasonable, and the evaluation results of different research work will be consistent and comparable [10].

Each sample in the NSL-KDD dataset consists of 41-dimensional features and 1-dimensional tags. The dataset contains a number of attacks that fall into four categories: Denial of Service (DoS), surveillance and probing (Probe), unauthorized access from a remote machine to a local machine (R2L) and unauthorized access to local superuser privileges by a local unprivileged user (U2R). NSL-KDD includes the training dataset KDDTrain+_20Percent and the test dataset KDDTest-21. The training dataset consists of 21 attack types, and 17 new attack types are added to the test set KDDTest-21. The specific data distribution of NSL-KDD dataset is shown in Table 3.

The complete NSL-KDD dataset contains more than 140,000 records in total. If all of them are used as the verification dataset, it is too large. Therefore, 20% of the NSL-KDD dataset is selected as the sample data for experimental verification in this paper.

Each record in the NSL-KDD dataset is composed of 42 fields, of which the first 41 bits are characteristic fields and the last bit is the network behavior labeling field. The classification of network attack types in this dataset is shown in Table 4, where X_i is the attack threat factor defined in Section 2.1, different attacks have different threat levels according to the system permissions obtained by the attack and the degree of impact on the network operation. As the threat level increases, different attack types affect network security and the impact of operational threats has increased.

Table 3: NSL-KDD dataset distribution

Dataset name	**Type distribution**					
	Normal	DoS	Probe	U2R	R2L	Total
KDDTrain+	67345	45926	11655	52	995	125973
KDDTrain+_20Percent	13499	9234	2289	11	209	25192
KDDTest+	9711	7458	2421	200	2754	22544

Table 4: NSL-KDD dataset attack types

Threat level	**Attack types**	X_i
low	DoS	0.3
↓	U2R	0.5
	R2L	0.6
high	Probe	0.8

Table 5: Test samples

Sample	C_i				**Total**
	DoS	U2R	R2L	Probe	
1	202	0	7	47	256
2	274	0	9	75	358
3	137	0	3	32	172
4	76	0	2	26	104
5	290	0	6	63	359
6	57	0	100	12	69
7	329	1	6	78	414
8	224	0	2	61	287
9	155	0	6	36	197
10	130	0	4	35	169
11	222	1	3	55	281
12	88	1	0	17	106
13	303	0	9	69	381
14	88	0	2	23	113

4.2 Analysis of Experimental Results

The network situation assessment has a strong periodicity, set the assessment period to T, extract the NSL-KDD sample set data in time T to simulate the intrusion detection operation, and utilize Snort to obtain it. Using Python 3.6 and MATLAB for simulation and obtaining the situation assessment factor, the proposed algorithm based on improved text SimHash is used to perform node security situation assessment. Assuming that the evaluation period is 1 day, obtain the attack information within the evaluation period, count the types of attacks within 1 day, and the frequency of attacks of various types of attacks. Exploiting the text processing method in Section 3.3.2, several modified texts are obtained. The SimHash algorithm is adopted to generate multiple hash values corresponding to the modified text, and the hamming distance between these hash values and the hash values generated by the original text is calculated. The larger the Hamming distance, the smaller the similarity between the two vectors. This feature can be used to quantify the severity of an attack of this type on a certain period of time.

The impact of the attack on the network nodes is formed by the combination of external attack information, internal vulnerability information of the host node, and threat information of the attack itself. This paper combines these Hamming distances, corresponding probability of attack success rate and attack threat value to get the final node security situation value. At the same time, randomly selected 14 consecutive test samples are shown in Table 5.

The experiment simulates a vulnerability-based attack. Refer to the Common Vulnerability Scoring System(CVSS) for scoring, combined with the threat index of the evaluation index factor and the CVSS basic score of the vulnerability in Table 2, it was taken as the actual network security situation values to verify the feasibility and efficiency of the method presented in this paper.

The node security situation assessment algorithm based on improved text SimHash is used to evaluate the node's security status. The selected sample set data is analyzed to eliminate duplicate data, and finally the attack information A is obtained, and the vulnerability information V is constructed according to Table 2. A and V are taken as the input data of the **Algorithm 1** to generate the security situation value of each node. In order to verify the effectiveness of the proposed method, the improved text SimHash algorithm, the improved BP neural network [5] and the SimHash algorithm [15] were respectively used to conduct the situation assessment on the test samples. Normalized the assessment results, and combined with the grading of the security assessment level in Table 1 of Section 2.3, the final security situation levels are shown in Table 6.

In order to present the assessment results more clearly, Figure 4 shows a comparison chart of the security situation assessment results.

As can be seen from Figure 4, except for sample 12, the situation assessment results of the three methods are very close. It can be seen from Tables 5 and 6 that the total number of attacks in sample 4 is smaller than the number of attacks in sample 12, but the situation assessment results obtained by the proposed method and the improved BP neural network assessment method are that the former is greater than the latter. This is because the number of Probe attacks in sample 4 is greater than that in sample 12, and the attack threat factor of Probe attacks is greater than that of other three types of attacks; in addition, U2R and R2L attacks are relatively obviously dependent on specific vulnerabilities, but R2L relies on vulnerabilities with larger threat factors, so it is reasonable that the situation assessment results obtained

Table 6: Results of security situation assessment

Sample	**Assessment results**							
	Proposed	Level	References [15]	Level	References [5]	Level	Actual value	Level
1	0.532427	Moderate	0.538937	Moderate	0.544587	Moderate	0.530567	Moderate
2	0.749866	Moderate	0.775649	Moderate	0.766993	Moderate	0.674694	Moderate
3	0.389887	Mild	0.427361	Mild	0.380982	Mild	0.355811	Mild
4	0.356893	Mild	0.311806	Mild	0.348741	Mild	0.264081	Mild
5	0.737127	Moderate	0.747219	Moderate	0.753963	Moderate	0.643435	Moderate
6	0.187817	Safe	0.185255	Safe	0.183521	Safe	0.186335	Safe
7	0.9771607	Severe	0.963862	Severe	0.954842	Severe	0.971032	Severe
8	0.6342627	Moderate	0.599506	Moderate	0.648749	Moderate	0.623566	Moderate
9	0.5119247	Moderate	0.535847	Moderate	0.500232	Moderate	0.509436	Moderate
10	0.4095747	Mild	0.485785	Mild	0.400220	Mild	0.417648	Mild
11	0.565087	Moderate	0.595179	Moderate	0.552180	Moderate	0.585914	Moderate
12	0.230607	Mild	0.378245	Mild	0.225340	Mild	0.259097	Mild
13	0.9210917	Severe	0.889988	Severe	0.942129	Severe	0.792165	Moderate
14	0.3634067	Mild	0.415328	Mild	0.355106	Mild	0.266893	Mild

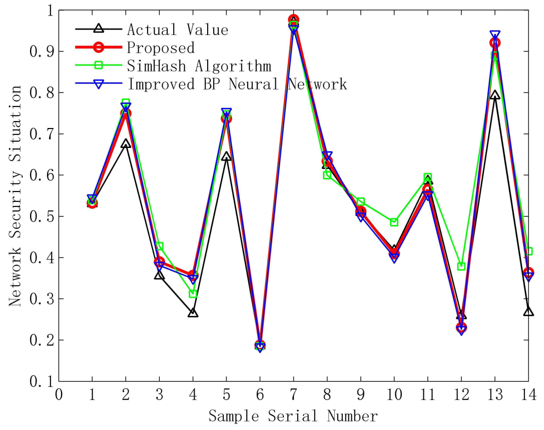


Figure 4: Situation assessment results

in sample 4 are larger than that obtained in sample 12.

The mean absolute percentage error(MAPE) of the assessment results is defined as follows:

$$MAPE = \frac{100\%}{n} \sum_{i=1}^{n} | \frac{\hat{y}_i - y_i}{y_i} |$$

where $\hat{y}_i$ and y_i are network situation assessment values and actual network situation values; n is the number of samples.

By calculation, the MAPE of the assessment results by the proposed method, the SimHash algorithm and the improved BP neural network are 10.24%, 15.22% and 11.19%, respectively. The MAPE of the proposed method is 4.98 and 0.96 percentage points lower than that of SimHash algorithm and improved BP neural network, respectively. The accuracy of the proposed method is better, which can more accurately assess the network security situation. Comprehensive analysis shows that the method proposed in this paper is reliable and feasible.

4.3 Performance Comparison

In view of the existing network security situation assessment methods to deal with the problem of weak ability of complex and diverse network attack methods and attack types, the proposed method only needs to obtain the type of attack suffered by the network node, the frequency of attacks and the vulnerability information of the node itself on the basis of IDS and vulnerability scanning tools to assess the security status of the network. In the proposed method, after obtaining network security data such as network traffic, through a certain preprocessing, only the hash values corresponding to the initial text and the modified text need to be stored separately. Compared with storing network traffic and logs directly, a lot of space is saved because these hashes are stored in HDFS in a big data environment. Using the Snappy compression algorithm to compress the data can further save the disk space occupied by the data, speed up the data transmission speed on the disk and the network, and thus increase the processing speed of the system. Table 7 shows a comparison of several compression algorithms.

In the context of big data, the traditional situation assessment algorithm is more complex and the calculation cost is higher. Figure 5 shows the running time comparison results of the proposed method with the improved BP neural network and SimHash algorithm.

It can be seen from Figure 5 that in order to calculate the node security situation value, the running time of the improved BP neural network, SimHash algorithm and the proposed method are 1896.2693s, 0.68773s and 0.67847s in sequence. The cost of the proposed method is much lower than that of the improved BP neural network, and the calculation speed is also slightly higher than that of SimHash algorithm. With the method of this paper, when the network node data increases rapidly, the network security status can be timely and effectively fed back to the network administrator.

The proposed method can be applied to large networks, because the Louvain algorithm is used to cluster large networks into multiple communities, and each community

Table 7: Performance comparison of common compression algorithms

Compression algorithm	**Uncompressed file size(byte)**	**Compressed file size(byte)**	**Compression time(ms)**	**Decompression time(ms)**	**Peak CPU**
DEFLATE	35984	8677	11591	2362	29.5
GZIP	35984	8804	2179	389	26.5
BZIP2	35984	9704	680	344	20.5
LZO	35984	13069	581	230	22
LZ4	35984	16355	327	147	12.6
Snappy	35984	13602	424	88	11

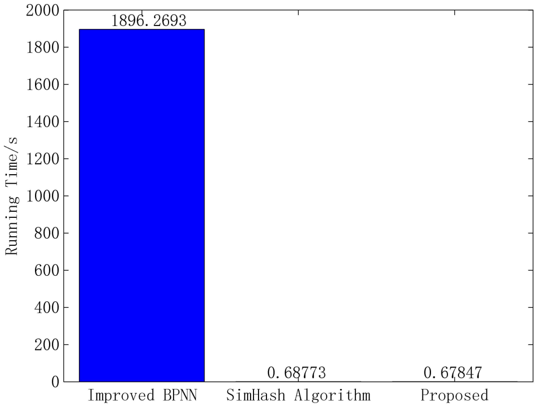


Figure 5: Running time comparison

can be regarded as a local network node. As long as there is a topology, the topology of the entire network can be determined. Network traffic, alarms, and other relevant network security data can be collected and uploaded to HDFS for storage and analysis at the time of generation, so this process can be performed in parallel with data generation.

5 Conclusions

A security situation assessment algorithm for industrial control network nodes based on improved text SimHash is proposed. The proposed method solves the problems of the existing industrial control network situation assessment method, such as weak ability to deal with complex and diverse network attack types, complicated assessment process, and difficulty in adapting to the big data industrial control network environment. Firstly, the Louvain algorithm for detecting community structure in a complex network is used to cluster the large-scale industrial control network. Secondly, intrusion detection equipment, traffic monitoring equipment and vulnerability scanning systems are exploited to obtain security data and vulnerability information on each node in the network. Finally, the text is constructed according to the obtained security data, and the improved SimHash algorithm is adopted to calculate the similarity of the constructed text to quickly and efficiently evaluate the security situation of the industrial control network nodes.

The shortcoming of this paper is that when assessing the security situation of industrial control network nodes, the performance indicators of specific services on the nodes are not considered. In future work, the attack information, node vulnerability information and node service performance information will be considered comprehensively to assess the security situation of industrial control network nodes more reasonably.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (No. 61862041, 61363078). The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

References

[1] M. R. Asghar, Q. W. Hu, and S. Zeadally, “Cybersecurity in industrial control systems: Issues, technologies, and challenges,” *Computer Networks*, vol. 165, p. 106946, 2019.

[2] D. Bhamare, M. Zolanvari, A. Erbad, R. Jain, K. Khan, and N. Meskin, “Cybersecurity for industrial control systems: A survey,” *Computers & Security*, vol. 89, p. 101677, 2020.

[3] V. D. Blondel, J. L. Guillaume, R. Lambiotte, and E. Lefebvre, “Fast unfolding of communities in large networks,” *Journal of Statistical Mechanics Theory and Experiment*, vol. 2008, no. 10, pp. 440–442, 2008.

[4] Y. X. Chen, X. C. Yin, and A. Sun, “Network security situation assessment model based on GSA-SVM,” *DEStech Transactions on Computer Science and Engineering*, vol. 291, pp. 414–420, 2018.

[5] G. S. Dong, W. C. Li, S. W. Wang, X. Y. Zhang, J. Z. Lu, and X. Li, “The assessment method of network security situation based on improved BP neural network,” in *International Conference on Computer Engineering and Networks*, pp. 67–76, Shanghai, China, August 2018.

[6] M. R. Endsley, “Design and evaluation for situation awareness enhancement,” *DEStech Transactions on Computer Science and Engineering*, vol. 32, no. 2, pp. 97–101, 1988.

[7] T. Fu, Y. Q. Lu, and W. Zhen, “APT attack situation assessment model based on optimized BP neural

network," in *2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, pp. 2108–2111, Chengdu, China, March 2019.

[8] S. Ghosh, M. Halappanavar, A. Tumeo, A. Kalyanaraman, H. Lu, D. Chavarri-Miranda, A. Khan, and A. Gebremedhin, "Distributed louvain algorithm for graph community detection," in *2018 IEEE International Parallel and Distributed Processing Symposium (IPDPS)*, pp. 885–895, Vancouver, BC, Canada, May 2018.

[9] N. A. Giacobe, "Application of the JDL data fusion process model for cyber security," in *Multisensor, Multisource Information Fusion: Architectures, Algorithms, and Applications 2010*, p. 77100R, Orlando, America, April 2010.

[10] Y. Hamid, V. R. Balasaraswathi, L. Journaux, and M. Sugumaran, "Benchmark datasets for network intrusion detection: A review," *International Journal of Network Security*, vol. 20, no. 4, pp. 645–654, 2018.

[11] L. Hu, H. L. Li, Z. H. Wei, S. Q. Dong, and Z. Zhang, "Summary of research on it network and industrial control network security assessment," in *2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, pp. 1203–1210, Chengdu, China, March 2019.

[12] M. Husák, T. Jirsík, and S. J. Yang, "SoK: contemporary issues and challenges to enable cyber situational awareness for network security," in *Proceedings of the 15th International Conference on Availability, Reliability and Security*, pp. 1–10, Ireland, August 2020.

[13] X. Li, Y. Lu, S. Liu, , and W. Nie, "Network security situation assessment method based on markov game model," *TIIS*, vol. 12, no. 5, pp. 2414–2428, 2018.

[14] Y. Li, G. Q. Huang, C. Z. Wang, and Y. C. Li, "Analysis framework of network security situational awareness and comparison of implementation methods," *EURASIP Journal on Wireless Communications and Networking*, vol. 2019, no. 1, pp. 1–32, 2019.

[15] P. W. Lin and Y. H. Chen, "Network security situation assessment based on text simhash in big data environment," *International Journal of Network Security*, vol. 21, no. 4, pp. 699–708, 2019.

[16] J. Liu, T. Jin, K. Pan, Y. Yang, Y. Wu, and X. Wang, "An improved knn text classification algorithm based on simhash," in *2017 IEEE 16th International Conference on Cognitive Informatics Cognitive Computing (ICCI*CC)*, pp. 92–95, Oxford, UK, July 2017.

[17] G. H. Lu and D. Q. Feng, "Network security situation awareness for industrial control system under integrity attacks," in *2018 21st International Conference on Information Fusion (FUSION)*, pp. 1808–1815, Cambridge, UK, July 2018.

[18] Y. F. Lv, H. R. Ren, X. F. Gao T. Sun, H. P. Zhang, , and X. Y. Guo, "Multi-scale risk assessment model of network security based on LSTM," in *International Conference on Verification and Evaluation of Computer and Communication Systems*, pp. 257–267, Xi'an, China, October 2020.

[19] J. Qiang, F. Wang, and X. L. Dang, "Network security based on DS evidence theory optimizing CS-BP neural network situation assessment," in *2018 5th IEEE International Conference on Cyber Security and Cloud Computing (CSCloud)/2018 4th IEEE International Conference on Edge Computing and Scalable Cloud (EdgeCom)*, pp. 153–159, Shanghai, China, June 2018.

[20] C. Sadowski and G. Levin, "Simhash: Hash-based similarity detection," *Technical report, Google*, 2007.

[21] H. Wang, Z. F. Chen, X. Feng, X. Q. Di, D. Liu, J. P. Zhao, and X. Sui, "Research on network security situation assessment and quantification method based on analytic hierarchy process," *Wireless Personal Communications*, vol. 102, no. 2, pp. 1401–1420, 2018.

[22] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.

[23] S. K. Yang and C. Chou, "A hybrid methodology of effective text-similarity evaluation," in *International Computer Symposium*, pp. 227–237, Yunlin, Taiwan, December 2018.

[24] B. Yi, Y. P. Cao, and Y. Song, "Network security risk assessment model based on fuzzy theory," *Journal of Intelligent & Fuzzy Systems*, vol. 38, no. 4, pp. 3921–3928, 2020.

[25] Z. W. Zhao, T. T. Zhou, and H. Wang, "Quantitative evaluation model of network security situation based on ds evidence theory," in *2019 6th International Conference on Dependable Systems and Their Applications (DSA)*, pp. 371–376, Harbin, China, January 2020.

[26] C. H. Zhou, "An improved algorithm for big data deduplication based on simhash," *Computer and Modernization*, vol. 000, no. 7, pp. 38–41, 2017.

[27] B. W. Zhu, Y. H. Chen, and Y. Q. Cai, "Three kinds of network security situation awareness model based on big data," *International Journal of Network Security*, vol. 21, no. 1, pp. 115–121, 2019.

Biography

Rui-hong Dong. 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.

Chuang Shu. Received the BS degree in network engineering from Hubei University of Technology, Wuhan, China, in 2018. Currently, he is studying for a master's degree in Lanzhou University of Technology. His research interests include network and information security, industrial control network security.

Zhang Qiu-yu. Researcher/Ph.D. 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.

Intrusion Detection Model Based on Feature Selection and Random Forest

Rui-Hong Dong, Yong-Li Shui, and Qiu-Yu Zhang

(Corresponding author: Qiu-Yu Zhang)

School of Computer and Communication, Lanzhou University of Technology
No. 287, Lan-Gong-Ping Road, Lanzhou 730050, China
Email: zhangqylz@163.com

(Received July 22, 2020; Revised and Accepted Mar. 6, 2021; First Online Oct. 19, 2021)

Abstract

Given the problems of massive high-dimensional data, redundancy, and uncertainty of data features in the industrial control networks intrusion detection system (IDS), an intrusion detection model based on feature selection and random forest (RF) is proposed. Firstly, the proposed model selects the relevant attributes through the information gain (IG) feature selection method. Secondly, the principal component analysis (PCA) feature extraction method selects the optimal feature subset from the relevant attributes selected by the IG. Finally, the selected optimal feature subset is used for training and testing in the RF classifier. To better prove the performance of the proposed method, the model is compared with the support vector machine (SVM), decision tree (DT), and logistic regression (LG) methods. Experimental results show that the detection accuracy of the proposed model using the NSL-KDD dataset is 99.84% and that using the CICIDS2017 dataset is 99.80%. At the same time, compared with the existing methods, the proposed model has good classification detection performance.

Keywords: Feature Selection; Information Gain; Intrusion Detection; Principal Component Analysis; Random Forest Classifier

1 Introduction

With the continuous integration of industrialization and informatization, more and more computers and information technologies are applied to industrial control networks, especially the proposal of "Industry 4.0" in German has accelerated the process of opening up industrial control networks to the outside world, and industrial control networks are facing more serious threats. Throughout the various virus attacks in recent years, it can be seen that industrial control networks intrusion is powerful and destructive, and can widely launch attacks [13]. Therefore, research on the security of industrial control networks has more important research significance [29].

Aiming at the security problems in industrial control networks, 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 [10]. At present, in the research of industrial control network intrusion detection, researchers use data mining technology, deep learning and machine learning technology to deal with intrusion detection problems [22]. Traditional intrusion detection methods based on machine learning include naive Bayes, K-means, C4.5, *etc.* [21]. These methods can accurately detect network attacks, and have better detection performance than traditional methods. However, the existing network intrusion detection original dataset has large amount of data, high feature dimension, and contains redundant and irrelevant features, which seriously affects the performance of the intrusion detection model, making the traditional machine learning-based intrusion detection technology rely on feature method [9, 15].

In the industrial control network intrusion detection problem, there are problems such as high data dimension, attribute redundancy, and high calculation cost, which seriously affect the performance of the intrusion detection model and may lead to low classification accuracy and high false alarm rate. Feature selection has been proved to be an important mechanism of IDS, it can extract relevant features, also eliminate irrelevant features that cause false positives, reduce data dimension, and improve detection accuracy [20]. Therefore, In order to reduce the influence of high-dimensional network traffic data in the industrial control networks on the intrusion detection model, and improve the detection accuracy of the intrusion detection method, we presents an intrusion detection model based on feature selection and random forest (IG-PCA-RF). The proposed model first uses a mixed method of IG and PCA for feature selection. By calculating the information entropy of each attribute, each attribute feature is arranged in descending order to extract relevant attributes. Then, PCA is used to perform secondary feature reduction for relevant attributes selected by IG to select the optimal feature subset. Finally, the optimal feature subset selected by IG-PCA is used to train and detect the model, and constantly adjust the parameters of the RF classifier to improve the classification precision of the model. The main contributions of this paper are as follows:

1) The combination of feature selection and RF classifier is used to solve the problem of low performance of intrusion detection models caused by massive high-dimensional data.

2) Combining IG with PCA, a hybrid feature selection algorithm is proposed, which solves the problems of high dimensionality of data features, large redundancy and uncertainty of data features, and can select the optimal features subset.

3) The classification performance of the combination of RF and IG-PCA is compared with that of other methods (DT, SVM, LG) combined with IG-PCA, which proves that the combination of RF and IG-PCA has better detection performance.

The remaining part of this paper is organized as follows. Section 2 introduces related work. Section 3 introduces related theories in detail. Section 4 gives a description of the proposed intrusion detection model and related algorithms. Section 5 gives the experimental results and performance analysis as compared with existing methods. Finally, we conclude our paper in Section 6.

2 Related Work

As an important means of industrial control networks security protection, intrusion detection technology has been constantly developed, and has become a hotspot in industrial control networks security research. Research scholars have applied mature network intrusion detection technologies, such as data mining, machine learning, and deep learning to industrial control networks security, and constantly innovating. With the emergence of network traffic data with high dimensionality, serious noise redundancy, and unbalanced data class, feature selection has been widely used in the field of intrusion detection as a means of data dimensionality reduction. Feature selection selects the feature subset that has the greatest effect on the classification from the original feature set, so as to effectively reduce the computation and improve the classification accuracy. Dong *et al.* [10] proposed an industrial internet intrusion detection model based on mutual information, the mutual information feature selection method is used to reduce the features of NSL-KDD. In [32], random forests are used to eliminate recursive features, and the deep multilayer perceptron (DMLP) structure is used for intrusion detection with an accuracy rate of 91%. In [33], an intrusion detection method based on feature selection and ensemble learning technology is proposed, according to the correlation between features, heuristic dimensionality reduction algorithm CFS-BA is used to select the optimal subset. Amrita [2] proposed 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. Atkison *et al.* [3] studied the optimization of feature extraction algorithms for the development of industrial control network intrusion detection systems (IDS) based on network telemetry.

Combining data mining with industrial control networks intrusion detection technology has many benefits. Sapozhnikova *et al.* [25] proposed an industrial network IDS based on data mining, which utilizes PCA feature select and SVM classification detection, with a high detection rate. In [5], an improved industrial control systems intrusion detection model is proposed. This method combines SMLC hierarchical clustering with sequence coverage similarity to improve the model, and improves the accuracy of anomaly detection in industrial control systems. In [18], an improved multi-innovation Kalman particle swarm algorithm is proposed, and the improved algorithm is used to optimize the parameters of the support vector machine industrial control intrusion detection model. Optimized model has significantly improved detection rate. Chen *et al.* [6] proposed a new intrusion detection system (IDS) based on information gain criteria to select relevant features from network traffic records, and proposed a new support vector domain description to classify extracted features and detect new intrusions.

Traditional machine learning algorithms have achieved certain results in industrial control networks intrusion detection. Liang *et al.* [19] proposed an industrial network intrusion detection algorithm, which is based on a multi-feature data clustering optimization model. This method is mainly aimed at the abnormal behavior of the intrusion attack data in the industrial network, and the detection rate of NSL-KDD can reach 97.8%, but the training process of the model is complex. In [7], for the security problem of industrial control networks, a multi-level adaptive coupling method combining white list technology and machine learning is proposed. In the machine learning process, PCA preprocessing is performed on the dataset, and the adaptive coupling algorithm is used for training and detection. The detection rate fluctuates to a certain extent. In [26], in order to solve the problem of the application layer network protocol of industrial control systems, an intrusion detection method based on fuzzy C-means clustering and SVM is proposed. This method effectively reduces training time and improves classification accuracy.

Deep learning is different from traditional machine learning technology, which can learn data features through deep hidden layers. In recent years, researchers have begun to apply deep learning to the field of intrusion detection. In [8], an industrial control intrusion detection method based on deep learning model is proposed. It uses the long-short-term memory model to detect network packets carrying information, which has good detection accuracy, but cannot accurately extract features from high-dimensional and complex changing data. Hu *et al.* [14] used convolutional neural network (CNN) to extract features of network data by analyzing the network data of the industrial control systems (ICS), and classified the abnormal behavior in KDD99, but the detection accuracy rate was low and the false alarm rate was high. In [16], a hybrid structure with deep feature selection process is proposed, and a hybrid node structure is classified using a back propagation neural network (BPNN). In [1], the deep neural network (DNN) and decision tree (DT) classifiers are used to detect network attacks on industrial control systems. The method has good classification performance and needs to be improved in data processing.

Data mining and traditional machine learning methods can effectively improve the detection rate, but they have great limitations for mass high-dimensional data processing. Although deep learning can fully learn the data features, its network structure and training process are relatively complicated, which increases the difficulty of model training. Therefore, in view of the problems of high dimensionality of massive data features in industrial control networks, including redundant and irrelevant data, unbalanced data class, and insufficient extraction of data features by traditional intrusion detection methods. This paper proposes intrusion detection model based on feature selection and random forest, feature selection based on IG and PCA is performed on the dataset, and then uses random forest for training detection. Moreover, this paper also compares with other classifiers and some existing research methods.

3 Related Theories

3.1 Information Gain

Information gain (IG) [22] is a filtering method that reduces the dimensionality of the dataset. The main idea of this method is to sort the attribute subsets in descending order by calculating the information entropy of each attribute. Information gain can be used to measure the amount of information after certain attribute eliminates uncertainty. Let X and Y be the random feature attributes in the dataset, and the IG calculation is as Equation (1):

$$IG(X,Y) = H(X) - H(X \mid Y) \tag{1}$$

where $H(X)$ represents the information entropy of the random feature attributes X, which is defined as Equation (2):

$$H(X) = -\sum_i p(x_i)\log_2(p(x_i)) \tag{2}$$

where $p(x_i)$ is the proportion of class i samples in X, and $H(X|Y)$ represents the information entropy of X under the condition of Y, and its definition is shown in Equation (3):

$$H(X|Y) = -\sum_j p(y_j)\sum_i p(x_i|y_j)\log_2(p(x_i,y_j)) \tag{3}$$

where $p(y_j)$ is the proportion of j class samples in Y, and $p(x_i|y_j)$ represents the probability of the occurrence of variable X under the condition of Y.

In feature selection, the IG is for a feature item. The greater the IG of a feature, the contribution of the feature to the classification the bigger, and those features with large IG should be retained.

3.2 Principal Component Analysis

Principal component analysis (PCA) [28] is a dimensionality reduction method, which can reduce the dimensionality of the dataset and remove redundant information while retaining the important information of the data. The basic idea is to transform multiple features into a few unrelated new features ordered by importance through linear transformation. These new features are linear combinations of the original features, and these new features can be used to describe the original sample without changing its main components.

Specifically, Let $s(t)$ be a random data set selected by the IG feature selection method, where $t = 1, 2, ..., n$. The covariance matrix of $s(t)$ is shown in Equation (4):

$$R = 1\diagup(n-1)\sum_{t=1}^{n}[s(t)s(t)^T] \tag{4}$$

The linear transformation from $s(t)$ to $v(t)$ in PCA is shown in Equation (5):

$$v(t) = N^T S(t) \tag{5}$$

where N is an $n\times n$ matrix, and the ith eigenvector is the ith column of the covariance matrix R. The characteristic value is calculated as Equation (6):

$$\lambda_i q_i = Rq_i \tag{6}$$

where λ_i represents the eigenvalue of R ($\lambda_1 > \lambda_2 > ... > \lambda_n$), and q_i represents the corresponding eigenvector.

On the basis of Equation (6), the principal component is calculated as Equation (7):

$$v_i(t) = q_i^T s(t), i = 1, \ldots, n. \tag{7}$$

where $v_i(t)$ represents the ith principal component.

3.3 Random Forest

Random forest (RF) [23] is a collection of multiple decision trees. The training process of RF can be decomposed into multiple decision trees for individual training. Multiple traffic data subsets are extracted from the dataset, and different subsets of multiple decision trees are trained, the integrated RF classifier is obtained.

The formula of the decision tree in the RF is expressed as Equation (8):

$$F(a, b_i), (i = 1, 2, \ldots, n) \tag{8}$$

where a represents input data, b_i represents independent and identically distributed feature vectors, and n represents the number of decision trees in the RF.

The final result of the RF is obtained by a simple majority vote, and the calculation formula is as Equation (9):

$$F^*(a) = \arg\max_c(\sum \psi(F_i(a, b_i)) = c) \tag{9}$$

where c represents the prediction result, and $F^*(a)$ is used to count the types with the most predicted votes in multiple decision trees $F(a, b_i)$, and it is used as the final output of the RF.

4 The Proposed Model

4.1 Intrusion Detection Model Based on IG-PCA

Aiming at the problems of low detection rate caused by large amount of intrusion data, redundant attributes and uncorrelated attributes in industrial control networks. This paper uses a combination of IG, PCA and RF algorithm to research the intrusion detection model. The proposed model is mainly composed of dataset, data preprocessing module, classification training module, attack response module and results. Figure 1 shows the flow chart of intrusion detection model based on IG-PCA.

As can be seen from Figure 1, the proposed model first uses IG feature selection method to select relevant attributes. The attributes selected from the IG method can be used directly for classification, but with attribute bias towards a wide range of possible values, which will affect feature attributes real sorting. Then the PCA method is applied to select the optimal feature subsets from the relevant attributes extracted by IG, realize feature selection, attribute reduction and redundancy removal for high-dimensional datasets, reducing the complexity and dimension of the data. In the end, the RF ensemble classifier is used to detect the data to obtain the classification results, and also improve the detection effect.

The detailed processing steps of the random forest intrusion detection model based on IG-PCA are as follows:

Step 1: Data preprocessing. The preprocess the original dataset, convert the character type data in the dataset to integer type, and normalize the attribute feature value.

Step 2: Feature selection. The preprocessed data uses IG feature selection method to reduce the dimension and extract the relevant attributes.

Step 3: Feature extraction. The PCA method is applied to the relevant attributes extracted by IG to select the optimal attribute subset, so as to realize feature extraction, attribute reduction and redundancy removal on high-dimensional datasets, reducing the complexity and dimension of data.

Step 4: Classification training. The model of RF classifier is constructed, and the model is trained with the training dataset. The model parameters are adjusted constantly to achieve a certain accuracy and reduce the prediction error.

Step 5: Attack response. The pre-processed testing dataset is input into the trained detection model, and decision-voting classification is performed on the testing dataset, and the results of decision voting are classified and integrated to judge whether it is abnormal or normal behavior, and divide into the correct classification.

4.2 Feature Selection Algorithm

The feature selection method combining IG with PCA is used to reduce the dimension of the input dataset, eliminate redundant and unrelated features, and meanwhile improve the classification efficiency. Algorithm 1 is the detailed processing of feature selection.

Algorithm 1 IG-PCA feature selection algorithm

```
Input: Training dataset and Testing dataset
Output: Select the optimal feature subset Xbest.
Procedure Compute $IG(X, Y)$
Compute the information entropy of feature attribute $X$ (cf. Equation (2))
Compute the information entropy of $X$ under the condition of $Y$ (cf. Equation (3))
Calculate the IG of feature attribute $X$ (cf. Equation (1))
The $k$ attribute with the highest score–$s$
Return $s$
Procedure Compute PCA($s$)
Compute the covariance matrix $R$ of $s$ (cf. Equation (4))
Compute the eigenvector $(q_1, \cdots, q_i)$ and eigenvalue$(\lambda_1, \cdots, \lambda_i)$ (cf. Equation (6))
Calculate the principal component to get the eigenvector with the largest eigenvalue–$X_{best}$ (cf. Equation (6))
Return $X_{best}$
End
```

4.3 Construction of RF Classifier

First, the corresponding sub-sample selection should be completed, and then the modeling should be completed with the help of decision tree construction. In the multi-classifier combination process, CART algorithm can be

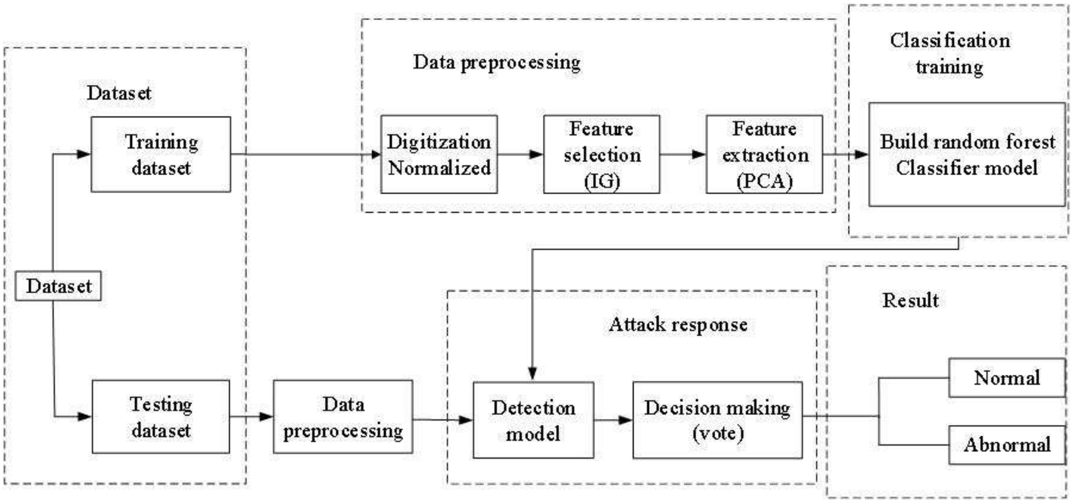


Figure 1: Flow chart of intrusion detection model based on IG-PCA

used for pruning. In the output strategy classification, the overall processing should be completed by voting method. Figure 2 shows the process of RF construction.

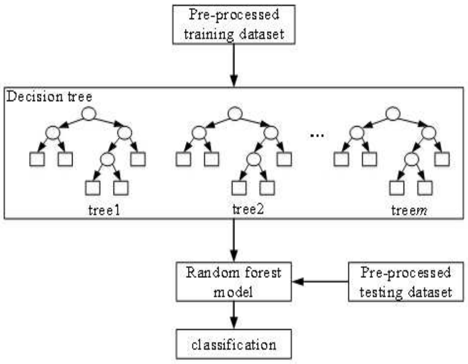


Figure 2: The process of RF construction

The specific steps for building the RF classifier model are as follows:

Step 1: Randomly select n samples from the pre-processed training dataset samples.

Step 2: Select k features from all features. When nodes find features for splitting, randomly extract feature subsets from the features, find the optimal solution in the feature subsets, and apply to the nodes for splitting. Select the best segmentation attribute as a node to build a decision tree.

Step 3: Repeat Steps 1 and 2 to build m decision trees.

Step 4: The m decision trees form the RF, and the classification result is obtained by voting to determine whether an abnormal attack has occurred in the network.

5 Experimental Results and Analysis

The experimental environment is: Windows 10 64-bit operating system, CPU Intel Core i5-4210U 2.49GHZ, memory 8G. The experimental programming language is python3, and weka 3.8.3 software is used to process the data. Two datasets of NSL-KDD [24] and CICIDS2017 [27] were selected for the experimental dataset.

In the feature selection experiment, PCA is used for secondary attribute reduction in order to prevent IG from producing more false positives due to its preference for attributes. In the classification experiment, the RF classifier is trained and tested, and the parameter set by the RF classifier is 100. The training dataset is used to train the model, the model parameters are adjusted constantly to achieve a certain accuracy, and reduce the prediction error. The trained detection model is used to classify and integrate the testing set data, and judge whether it is abnormal or normal behavior, and classify it into correct classification. Furthermore, in order to prove the superiority of the intrusion detection performance of this method, the existing decision tree (DT), support vector machine (SVM) and logistic regression (LG) three classifiers are compared.

5.1 Dataset Description

5.1.1 NSL-KDD Dataset

The NSL-KDD [24] dataset is an improvement to the KDD99 dataset, which solves the inherent problems in the KDD99 dataset, such as deleting redundant records, deleting duplicate records. The number of records in the training and testing sets are reasonable, which makes it affordable to run the experiments on the complete set without the need to randomly select a small portion. The training set of the NSL-KDD dataset does not contain redundant records, so the classifier does not bias towards more frequent records. There is no duplicate records in the NSL-KDD testing set, so the detection performance of the classifier model is not affected by duplicate records. Each data in the NSL-KDD dataset consists of 41 feature attributes and 1 class label. The 41-dimensional features in the dataset are divided into the following three types: the basic features of the network connection, traffic features, and network connection content features. According to the features of the dataset attack, it is divided into the following four types of attacks: denial of service attacks (DoS), probe attacks (Probe), user to root attacks (U2R), root to local attacks (R2L). There are 17 new attack types in the testing set of this dataset, but these attacks will not appear in the training set.

5.1.2 CICIDS2017 Dataset

The CICIDS2017 [27] dataset is an intrusion detection and intrusion prevention dataset. It was opened by the Canadian Cybersecurity Institute in 2017. It designed a real attack scenario, and collected traffic data by designing attack networks and victim networks. The CICIDS2017 dataset contains benign and most recent common attacks, similar to real-world data. It also includes the results of the network traffic analysis using CICFlowMeter with labeled flows based on the time stamp, source, and destination IPs, source and destination ports, protocols and attacks. Monday is the normal day and only includes the benign traffic. The implemented attacks include Brute Force FTP, Brute Force SSH, DoS, Heartbleed, Web Attack, Infiltration, Botnet and DDoS. They have been executed both morning and afternoon on Tuesday, Wednesday, Thursday and Friday. For each data, the author extracted 78 attributes and provided additional metadata about IP addresses and attacks. The CICIDS2017 dataset contains 2,830,743 records. These records are designed in 8 files, and each data contains 78 features and 1 label.

Table 1 shows the samples used in the experimental dataset.

Table 1: Selected samples of NSLKDD and CICIDS2017

Dataset	Class	Instances	Total instances
NSL-KDD	Normal Attack	75,247 71,463	146710
CICIDS2017	BENING Attack	55,992 55,647	108639

5.2 Data Preprocessing

5.2.1 Data Filtering

The data captured in the CICIDS2017 dataset in five days spanned a total of 8 files, which the 'Fwd Header Length' feature appeared twice, and some of the attack sample feature values were empty or infinite, and the dataset had a high Class imbalance problem. The dataset contains a total of 14 types of attacks and normal benign traffic. The largest number of category samples can reach hundreds of thousands, and the smallest number of category samples is only 11. In this case, if it is used for training classifier or detector, the model tends to be biased towards the majority class, while ignoring the minority class, resulting in poor model performance, lower accuracy, and higher false alarm rate.

Therefore, this paper solves the problem of data dispersion by merging files, merges the data of five days, selecting 10% of attack data from Tuesday to Friday and 10% of benign data on Monday, and deleting samples containing missing values and infinite values. At the same time, the dataset is divided into two parts, 80% for training and 20% for testing the model.

5.2.2 Data Transforming

The NSL-KDD dataset has three non-numerical features of protocol type, service and flag, and the remaining 38 characteristics are all numerical features. The classification modules of the intrusion detection model all need to calculate the numerical flow features, so non-numerical features must be converted into numerical features. For instance, the protocol-type feature in the NSL-KDD dataset contains three types of protocols, namely TCP, UDP and ICMP, which are replaced by 1, 2, and 3 respectively. As well, the 70 service attributes and 11 flag attributes in the dataset are also numeric in the same way.

5.2.3 Data Normalization

Data normalization is a process of scaling the value of each attribute to a relatively good range, in order to eliminate the preference for features with larger values from the dataset. Since the NSL-KDD dataset and the CICIDS2017 dataset have data with no fixed upper and lower bounds and continuous values, it is necessary to use min-max standardization to map the feature data to the standard range of [0, 1], and each feature is standardized using Equaption (10):

$$f(x) = (x - min)\diagup\ (max - min), x\epsilon[min, max] \qquad (10)$$

where x is the feature in the dataset, min and max are the minimum and maximum of feature x. The normalized data can be directly input into the intrusion detection model.

5.3 Evaluation Indexes

To evaluate the performance of the intrusion detection model, Accuracy (Acc), Precision (Precision), False acceptance rate (FAR), Recall rate and F1-score are used to measure the performance of the model. When evaluating the effectiveness of the model, most commonly used indicators can be calculated from the confusion matrix in Table 2.

In Table 2, TP indicates that the true value is a normal sample and is predicted to be the number of normal samples. FN indicates that the true value is a normal sample

Table 2: Confusion matrix

True value	Predictive value	
	Normal	Abnormal
Normal	TP	FN
Abnormal	FP	TN

and is predicted to be the number of abnormal samples. FP indicates that the true value is an abnormal sample and is predicted to be the number of normal samples. TN means that the true value is an abnormal sample and is predicted to be the number of normal samples.

Accuracy (Acc): Accuracy represents the percentage of the dataset that the model correctly classifies the true value of the sample. When the various samples in the dataset are relatively average, this is a good measure. However, when the various types of samples is unbalanced, it cannot reflect the true classification effect of the model. The calculation formula is as follows:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{11}$$

Precision: It represents the probability of actually being a positive sample among all the samples predicted to be positive. The calculation formula is as follows:

$$Precision = \frac{TP}{TP+FP} \tag{12}$$

FAR: Indicates an acceptable error rate. The calculation formula is as follows:

$$FAR = \frac{FPR+FNR}{2} \tag{13}$$

Recall rate: The proportion of positive samples among all samples predicted to be positive. The calculation formula is as follows:

$$Recall = \frac{TP}{TP+FN} \tag{14}$$

F1-score: Harmonized average of precision and recall. The calculation formula is as follows:

$$F1-score = \frac{2 \cdot (P \cdot R)}{P+R} \tag{15}$$

MBT (s): Model construction time consumed by a single training of intrusion detection model.

5.4 Performance Analysis

According to the ability of intrusion detection model to classify network traffic data, the performance of intrusion detection model is evaluated. To avoid the impact of data sampling when evaluating the intrusion detection model, the experiment was carried out using the ten-fold cross-validation method. More specifically, the proposed method was first evaluated using the NSL-KDD dataset; secondly, the performance of the proposed method was further verified using the CICIDS2017 dataset; and finally the performance of the method and some of the latest detection methods are compared from the case of Acc, Precision, FAR, F1-score and no feature selection.

In the IG feature selection module, Weka software [30] is used to feature selection on the NSL-KDD and CICIDS2017 datasets. IG performs ten-fold cross-validation on the dataset itself by calling the Ranker algorithm. By calculating the information entropy of each attribute, the features of each attribute are arranged from high to low, and the threshold is continuously adjusted to extract the relevant attributes. The features selected by the IG feature selection algorithm for the two datasets are shown in Table 3.

Table 3: IG feature selection attributes

Dataset	**Selected characteristics**
NSL-KDD	3, 4, 6, 12, 23, 25, 26, 29, 30, 33,34, 35, 37, 38, 39
CICIDS2017	1, 2, 5, 7, 9, 10, 15, 16, 17,18,19,22, 23, 30, 31, 37, 54, 64

Due to the preference of attributes selected by IG for attributes with a wide range of possible values, the ranking of these features may not be the correlation ranking with the training set, which will cause more false positives. In order to reduce this limitation, PCA is used to perform the second stage reduction step on the features listed in Table 3, to select the optimal subset of attributes, so as to further reduce the attributes in the feature selection stage. To avoid the deviation between the training set and the testing set, the PCA feature reduction step is only performed on the training set, mainly to ensure that the information of the testing set will not leak to the training set. If the whole dataset is used to perform PCA feature reduction steps, the performance of the model will be poor when new attacks are input into the model. As well, if the PCA feature reduction is performed on the training set and the testing set respectively, it will cause two datasets to be incompatible, so the classifier can not be trained by the training set, and then applied to the testing set. Therefore, using the same statistics of the training set, the same feature attributes are selected from the testing set.

IG feature selection can find attributes that have strong classification capabilities, thereby deleting attributes that are redundant, have an attribute value of 0, and have no effect on classification. Taking IG entropy 0.2 as the boundary, 15 dimensions are selected from the 41-dimensional features of the NSL-KDD data set, and 18 dimensions are selected from the 78-dimensional features of the CICIDS2017 data set. Since the number of features in the two data sets is far greater than the number of categories, the information gain will become very large and the generalization ability will decrease. It is necessary to perform PCA secondary feature reduction to reduce the 15-dimensional and 18-dimensional features selected by IG to 12 and 13 dimensions, respectively. These selected features are all features whose attribute value is not 0 and have a greater impact on classification. Compared with the two data sets, the feature attribute values in CICIDS2017 are generally higher than those in NSL-KDD, and the false alarm rate of classification is lower.

For evaluate the performance of the proposed model, the proposed feature selection method is compared with the non-feature selection method to distinguish between attacks and normal instances. Because the proposed IG-PCA algorithm selects the optimal feature subset, the performance of the detection model under the indicators of Acc, Precision, FAR, and F1-score has been significantly improved. Table 4 summarizes the performance based on the NSL-KDD dataset, including the results of the original feature, IG feature selection, and IG-PCA feature selection. The results show that if the feature selection is not realized, the RF classifier is not ideal in some indicators. When the proposed IG-PCA algorithm selects 12 optimal feature attributes of NSL-KDD, the proposed IG-PCA RF method performs best under these four indicators, which is obviously better than other classifiers. Among them, based on the NSL-KDD dataset, the accuracy of the model in this paper reaches 0.9984, Precision is 0.9980, FAR is 0.0210, and F1-score is 0.9984.

In order to further verify the performance of the model, this paper uses the CICIDS2017 dataset to conduct the verification experiment. The IG feature selection and PCA feature reduction of the CICIDS2017 dataset have been completed previously. This paper selects 13 optimal feature subsets of the CICIDS2017 dataset for experiments. The comparison results based on the CICIDS2017 dataset are shown in Table 5. It can be seen from the table that the proposed method still has good classification performance, which is better than other classifiers.

Moreover, because the feature dimension of the dataset is reduced, the method proposed in this paper reduces the time cost when constructing the model. Table 4 and Table 5 also compare the model construction time (MBT) based on different feature numbers. As can be seen from the table, compared with the random forest model using all features or IG alone, the random forest model using IG and PCA greatly reduce the model construction time. For the NSL-KDD dataset, although it does not take too much time to build the random forest classifier model, the model construction time (MBT) is reduced when IG-PCA feature selection method is applied, which is nearly 80% less than the original MBT. For the CICIDS2017 dataset, due to the large amount of data and high dimensionality, the RF classification method that maintains the original features requires about 240s. Because of the feature selection method, compared with the classification model using all the original features, the model in this paper greatly reduces the impact of MBT, and the construction time of intrusion detection model on these two datasets is reduced to within 50s. Especially for the CICIDS2017 dataset, when the IG-PCA-based hybrid feature selection method is used, the MBT of the random forest classifier is significantly reduced from 242.5s to 38.75s.

5.5 Performance Comparison with Other Feature Selection Methods

The intrusion detection datasets NSL-KDD and CICIDS2017 reflect that the network security environment has been more and more threat.The increasing number of various types of attacks, the large number of redundant and irrelevant features in their data, as well as the highly imbalance data in the dataset, which bring great challenges to machine learning methods. For further evaluate the feature selection method proposed in this paper, it is compared with some feature selection methods based on the two datasets NSL-KDD and CICIDS2017, namely IG (information gain) [22], IGR (information gain ratio) [11], PSO (Particle Swarm Optimization) [31], Wrapper (packaging method) [12]. In comparison with other feature selection methods, accuracy, the number of selected features, Precision and FAR are used as metrics. Figure 3 summarizes the performance of the proposed method and other feature selection methods based on the same classifier.

As shown in Figure 3(a), the detection accuracy of the proposed method on each dataset is better than other feature selection algorithms. The proposed IG-PCA-RF method achieves 99.84% and 99.8% accuracy rates on NSL-KDD and CICIDS2017 datasets, respectively. As well, Figure 3(b) shows the number of features selected using different features selection algorithms. It can be seen from the combination of Figure 3(a) and Figure 3(b), compared with Wrapper, Wrapper selects fewer features, but the detection accuracy is not high. Compared with IG, GA, and IGR, the feature selection of proposed method selects fewer features than IG, GA, and IGR, and it can be seen from Figure 3(a) that the accuracy of the proposed model is higher than the three methods. Figure 3(c) and Figure 3(d) show the model detection accuracy and false acceptance rate of the same classifier using different feature selection methods, which can indicate the efficiency of IDS. According to Figure 3(c), it can be observed that the detection accuracy of our proposed model on both data sets reached 99.8%, which is significantly better than any other feature selection methods based on the same classifier. In addition, as shown in Figure 3(d), the minimum FAR of the IG-PCA based intrusion detection model proposed in this paper on the NSL-KDD and CICIDS2017 datasets is 0.16% and 0.10%, respectively. Compared with other feature selection methods, the model proposed by us greatly reduces the false acceptance rate of each data set and ensures the effectiveness of IDS. Based on the above analysis, the IG-PCA feature selection method proposed in this paper is superior to other feature selection methods in performance and efficiency.

Table 4: Best performance classification based on each stage of NSL-KDD dataset

Classifier	**Acc**	**Precision**	**FAR**	**F1-Score**	**MBT (s)**
(a).Performance results based on original features(41)					
SVM	0.881	0.847	0.128	0.885	26.34
LG	0.917	0.892	0.181	0.919	23.91
DT	0.929	0.907	0.108	0.931	43.71
RF	0.920	0.894	0.098	0.922	181.55
(b).Performance results based on IG feature selection(15)					
IG-SVM	0.933	0.908	0.063	0.934	5.28
IG-LG	0.941	0.934	0.124	0.940	7.87
IG-DT	0.967	0.960	0.056	0.968	15.65
IG-RF	0.976	0.971	0.041	0.977	81.56
(c).Performance results based on IG-PCA feature selection(12)					
IG-PCA-SVM	0.953	0.936	0.040	0.954	3.65
IG-PCA-LG	0.971	0.960	0.062	0.971	4.86
IG-PCA-DT	0.982	0.976	0.032	0.982	8.57
IG-PCA-RF	0.9984	0.9984	0.0016	0.998	36.28

Table 5: Best performance classification based on each stage of the CICIDS2017 dataset

Classifier	**Acc**	**Precision**	**FAR**	**F1-Score**	**MBT (s)**
(a).Performance results based on original features(78)					
SVM	0.889	0.872	0.098	0.883	131.6
LG	0.894	0.889	0.101	0.908	123.9
DT	0.915	0.910	0.078	0.919	195.8
RF	0.932	0.925	0.046	0.934	242.5
(b).Performance results based on IG feature selection(18)					
IG-SVM	0.947	0.938	0.074	0.950	52.84
IG-LG	0.954	0.949	0.058	0.957	78.76
IG-DT	0.960	0.957	0.039	0.973	98.42
IG-RF	0.968	0.960	0.026	0.983	128.6
(c).Performance results based on IG-PCA feature selection(13)					
IG-PCA-SVM	0.974	0.976	0.024	0.974	26.53
IG-PCA-LG	0.981	0.982	0.032	0.981	34.86
IG-PCA-DT	0.996	0.992	0.012	0.996	32.23
IG-PCA-RF	0.998	0.998	0.001	0.998	38.75

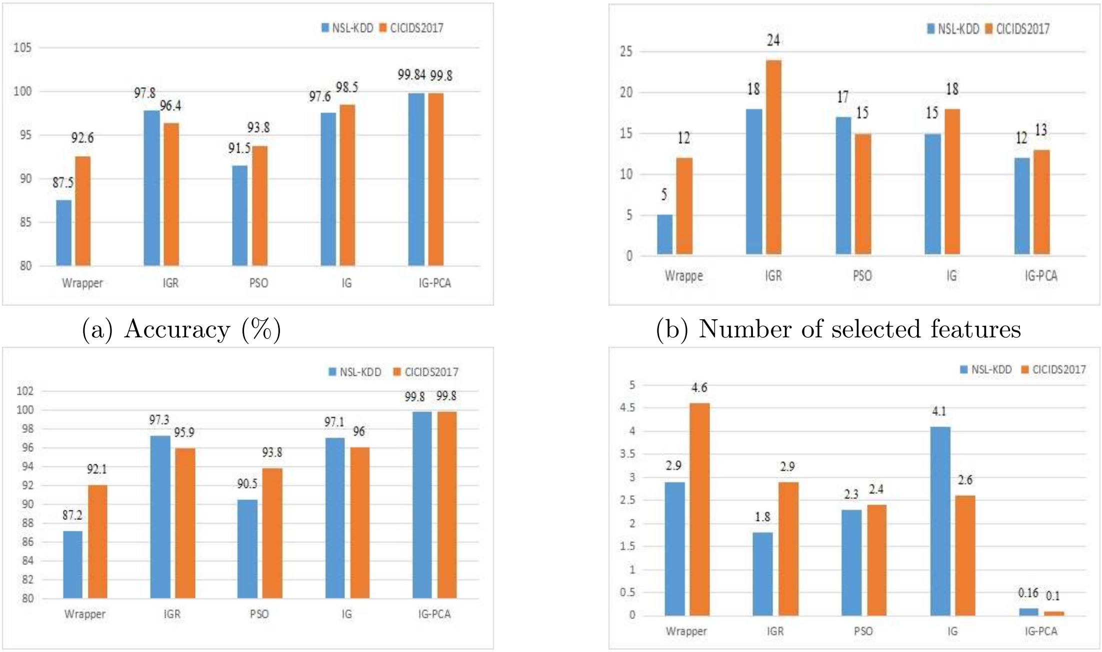
Table 6: Results compared with other existing research methods

Method	Dataset	Feature selection	Features	Acc(%) (s)	FAR(%) (s)
DE-ELM [12]	NSL-KDD	Wrapper	5	88.40	5
MDCOM [19]	NSL-KDD	-	41	97.80	-
IDM-MI [10]	NSL-KDD	MI	14	99.40	0.27
NSGA2-DT [17]	NSL-KDD	NSGA2-BLR	9-19	99.65	0.18
CFS-BA-ensemble [33]	NSL-KDD	CFS-BA	10	99.81	0.08
NSGA2-DT [17]	CICIDS2017	NSGA2-BLR	7-25	87.95	0.31
RF-DMLP [32]	CICIDS2017	RF	10	91.00	-
XGBoost-IDS [4]	CICIDS2017	-	78	91.36	12
ALDD [16]	CICIDS2017	BPNN	32	99.23	-
CFS-BA-ensemble [33]	CICIDS2017	CFS-BA	13	99.89	0.12
Proposed method	NSL-KDD	IG-PCA	12	99.84	0.16
Proposed method	CICIDS2017	IG-PCA	13	99.80	0.10

5.6 Performance Comparison with Existing Methods

For the sake of further highlight the performance of the method in this paper, the performance of the proposed method is compared with the [4, 10, 12, 16, 17, 19, 32, 33] from the aspects of feature selection method, number of selected features, accuracy and FAR, the comparison results are shown in Table 6.

As can be seen from Table 6, the Acc value of the



(a) Accuracy (%)

(b) Number of selected features

(c) Precision (%)

(d) FAR (%)

Figure 3: Compares with other feature selection methods based on the two datasets

model based on the NSL-KDD dataset is higher than that of DE-ELM [12], MDCOM [19], NSGA2-DT [17], IDM-MI [10] and CFS-BA-ensemble [33]. The accuracy of the method in this paper is higher than these methods, but the error acceptance rate of CFS-BA-ensemble [33] is lower. In the model detection task based on the CICIDS2017 dataset, the detection Acc of the model is higher than NSGA2-DT [17], RF-DMLP [32], XGBoost-IDS [4], ALDD [16], but lower than CFS-BA-ensemble [33]. Compared with CFS-BA-ensemble [33], the detection accuracy of this method is 0.09% lower, but the FAR of CFS-BA-ensemble [33] is 0.02% higher than that of this method. The main reason is that the method in this paper adopts the hybrid feature selection algorithm, which selects fewer features and improves the classification accuracy of the model. The classification performance is lower than individual methods, because the proposed method will be affected by the selected classifier, as well as the specific selected features and network environment.

6 Conclusions and Future Work

In this paper,we proposes an intrusion detection model based on feature selection and RF classifier, which solves the problem of low intrusion detection performance caused by massive high-dimensional data in industrial control networks. At the same time, IG and PCA are combined to give a hybrid feature selection algorithm, which solves the problems of high dimension of data features, large redundancy and uncertainty of data features, and can select the optimal feature subset. Through experimental simulation on NSL-KDD and CICIDS2017 datasets, the classification accuracy of the method in this paper is 99.84% for 12 feature subsets of NSL-KDD dataset, and the classification accuracy rate of 13 feature subsets based on CICIDS2017 dataset is 99.80%. Compared with the performance of DT, SVM and LG classifiers, it is proved that the IG-PCA-RF model in this paper has the best classification performance. In addition, compared with the existing latest methods such as DE-ELM, Cosine-PIO, NSGA2-DT, CFS-BA-ensemble and other two classification methods, the classification accuracy of this method is more higher, which can effectively improve the performance of intrusion detection.

The shortcomings is that this paper does not consider the influence of dynamic network environment changes, and further research will consider the adaptability of the intrusion detection model to the dynamic network environment.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (No. 61862041, 61363078). The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

References

[1] A. Al. Abassi, H. Karimipour, A. Dehghantanha, and M. P. Reza, "An ensemble deep learning-based cyber-attack detection in industrial control system," *IEEE Access*, vol. 8, pp. 83965–83973, 2020.

[2] K. R. Amrita, "A hybrid intrusion detection system: Integrating hybrid feature selection approach with heterogeneous ensemble of intelligent classifiers," *International Journal of Network Security*, vol. 21, no. 3, pp. 438–450, 2019.

[3] T. Atkison, S. Ponomarev, R. Smith, and B. Chen, "Feature extraction optimization for network intrusion detection in control system networks," *International Journal of Network Security*, vol. 20, no. 5, pp. 853–861, 2018.

[4] A. Bansal and S. Kaur, "Extreme gradient boosting based tuning for classification in intrusion detection systems," *International Conference on Advances in Computing and Data Sciences, Springer*, vol. 905, pp. 372–380, 2018.

[5] M. E. Boujnouni and M. Jedra, "New intrusion detection system based on support vector domain description with information gain metric," *International Journal of Network Security*, vol. 20, no. 1, pp. 25–34, 2018.

[6] D. Q. Chen, P. H. Zhang, and H. Z. Wang, "Intrusion detection for industrial control systems based on an improved svm method," *Journal of Tsinghua University*, vol. 58, no. 4, pp. 380–386, 2018.

[7] W. Z. Chen, T. J. Liu, Y. Tang, and D. S. Xu, "Multi-level adaptive coupled method for industrial control networks safety based on machine learning," *Safety Science*, vol. 120, pp. 268–275, 2019.

[8] A. K. Chu, Y. X. Lai, and J. Liu, "Industrial control intrusion detection approach based on multiclassification googlenet-lstm model," *Security and Communication Networks*, vol. 2019, pp. 1–11, 2019.

[9] R. H. Dong, X. Y. Li, Q. Y. Zhang, and H. Yuan, "Network intrusion detection model based on multivariate correlation analysis – long short time memory network," *IET Information Security*, vol. 14, no. 2, pp. 166–174, 2020.

[10] R. H. Dong, D. F. Wu, and Q. Y. Zhang, "Mutual information-based intrusion detection model for industrial internet," *International Journal of Network Security*, vol. 20, no. 1, pp. 131–140, 2018.

[11] R. H. Dong, H. H. Yan, and Q. Y. Zhang, "An intrusion detection model for wireless sensor network based on information gain ratio and bagging algorithm," *International Journal of Network Security*, vol. 22, no. 2, pp. 218–230, 2020.

[12] H. A. Faezah, L. A. Wathiq, and K. I. Ali, "Differential evolution wrapper feature selection for intrusion detection system," *Procedia Computer Science*, vol. 167, pp. 1230–1239, 2020.

[13] Y. Hu, A. Yang, H. Li, Y. Y. Sun, and L. M. Sun, "A survey of intrusion detection on industrial control systems," *International Journal of Distributed Sensor Networks*, vol. 14, no. 8, pp. 1–14, 2018.

[14] Y. B. Hu, D. H. Zhang, G. Y. Cao, and Q. Pan, "Network data analysis and anomaly detection using cnn technique for industrial control systems security," in *IEEE International Conference on Systems, Man and Cybernetics (SMC'19)*, pp. 593–597, Bari, Italy, Oct. 2019.

[15] P. Illavarason and B. K. Sundaram, "A study of intrusion detection system using machine learning classification algorithm based on different feature selection approach," in *The Third International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, pp. 295–299, Palladam, India, Dec 2019.

[16] J. G. Jiang, Q. Yu, M. Yu, and G. Li, "ALDD: A hybrid traffic-user behavior detection method for application layer DDoS," in *The 17th IEEE International Conference On Trust, Security And Privacy in Computing and Communications/ 12th IEEE International Conference on Big Data Science and Engineering (TrustCom/BigDataSE)*, pp. 1565–1569, Aug. 2018.

[17] C. Khammassi and S Krichen, "A NSGA2-LR wrapper approach for feature selection in network intrusion detection," *Computer Networks*, vol. 172, 2020.

[18] H. Li, B. Wang, and X. Xie, "An improved content-based outlier detection method for ICS intrusion detection," *EURASIP Journal on Wireless Communications and Networking*, vol. 103, no. 2020, pp. 1–15, 2020.

[19] W. Liang, K. C. Li, J. Long, X. Y. Kui, and A. Y. Zomaya, "An industrial network intrusion detection algorithm based on multifeature data clustering optimization model," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 3, pp. 2063–2071, 2020.

[20] S. Maza and M. Touahria, "Feature selection algorithms in intrusion detection system: A survey," *KSII Transactions on Internet and Information Systems*, vol. 12, no. 10, pp. 5079–5099, 2018.

[21] P. Mishra, V. Varadharajan, U. Tupakula, and E. S. Pilli, "A detailed investigation and analysis of using machine learning techniques for intrusion detection," *IEEE Communications Surveys and Tutorials*, vol. 21, no. 2, pp. 686–728, 2019.

[22] M. R. Mohamed, A. A. Nasr, I. F. Tarrad, and M. Z. Abdulmageed, "Exploiting incremental classifiers for the training of an adaptive intrusion detection model," *International Journal of Network Security*, vol. 21, no. 2, pp. 275–289, 2019.

[23] P. A. Resende and A. C. Drummond, "A survey of random forest based methods for intrusion detection systems," *ACM Computing Surveys*, vol. 51, no. 3, pp. 1–48, 2018.

[24] S. Roshan, Y. Miche, A. Akusok, and A. Lendasse, "Adaptive and online network intrusion detection system using clustering and extreme learning machines," *Journal of the Franklin Institut*, vol. 354, no. 4, pp. 1751–1779, 2018.

[25] M. U. Sapozhnikova, A. V. Nikonov, and A. M. Vulfin, "Intrusion detection system based on data mining technics for industrial networks," in *International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM'18)*, pp. 1–5, May 2018.

[26] W. L. Shang, J. R. Cui, C. H. Song, J. M. Zhao, and P. Zeng, "Research on industrial control anomaly detection based on fcm and svm," in *The 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/ 12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE)*, pp. 218–222, Aug. 2018.

[27] I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "Toward generating a new intrusion detection dataset and intrusion traffic characterization," in *International Conference on Information Systems Security and Privacy (ICISSP'18)*, pp. 108–116, 2018.

[28] Z. D. Shen, Y. H. Zhang, and W. Y. Chen, "A bayesian classification intrusion detection method based on the fusion of pca and lda," *Security and Communication Networks*, vol. 2019, pp. 1–11, 2019.

[29] L. Y. Shi, H. Q. Zhu, W. H. Liu, and J. Liu, "Industrial control system intrusion detection based on related information entropy and cnn-bilstm," *Journal of Computer Research and Development*, vol. 56, no. 11, pp. 2330–2338, 2019.

[30] K. Sumathi, S. Kannan, and K. Nagarajan, "Data mining: Analysis of student database using classification techniques," *International Journal of Computer Applications*, vol. 141, no. 8, pp. 22–27, 2016.

[31] B. A. Tama, M. Comuzzi, and K. H. Rhee, "A two-stage classifier ensemble for intelligent anomaly-based intrusion detection system," *IEEE Access*, vol. 7, pp. 94497–94507, 2019.

[32] S. Ustebay, Z. Turgut, and M. A. Aydin, "Intrusion detection system with recursive feature elimination by using random forest and deep learning classifier," in *International Congress on Big Data, Deep Learning and Fighting Cyber Terrorism (IBIGDELFT'18)*, pp. 71–76, Dec. 2018.

[33] Y. Y. Zhou, G. Cheng, S. Q. Jiang, and D. Mian, "Building an efficient intrusion detection system based on feature selection and ensemble classifier," *Computer Networks*, vol. 174, pp. 1–21, 2020.

Biography

Dong Rui-hong. 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.

Shui Yong-li. received a bachelor's degree in management from Jilin University of Finance and Economics in 2018. Currently, she is studying for a master's degree in Lanzhou University of Technology. The main research directions are network and information security, industrial control network security, intrusion detection, *etc.*

Zhang Qiu-yu. Researcher/Ph.D. 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.

An Electronic Voting Scheme Using Secure Multi-Party Computation Based on Secret Sharing

Hongquan Pu[1,2,3], Zhe Cui[1,2], and Ting Liu[1,2]
(Corresponding author: Hongquan Pu)

Chengdu Institute of Computer Applications, Chinese Academy of Sciences[1]
No.9, South Renmin Road, Sec.4, Chengdu 610041, China
Email: 774149765@qq.com
School of Computer and Control Engineering, University of Chinese Academy of Sciences[2]
No.19(A), Yuquan Road, Shijingshan District, Beijing 100049, China
Guangxi Key Laboratory of Hybrid Computation and IC Design Analysis[3]
No.188, East University Road, Nanning, 530006, China

(Received Jan. 2, 2020; Revised and Accepted Oct. 5, 2020; First Online Oct. 19, 2021)

Abstract

The security requirement is a challenging and critical issue in electronic voting systems. This paper proposes a method that uses a blind signature to authenticate voters and then achieves the voting process through secure multi-party computation (SMPC) based on Shamir'(t, n) secret sharing. Voters use a secret distributor to submit each secret share of votes to the vote counters. The secret shares received are calculated by vote counters, and then the total number of votes is computed by using an interpolation method. In this manner, there is no need to restore all the votes, and the security requirements, such as anonymity, receipt-free, verifiability, and accuracy, are well met. The experimental studies have demonstrated the effectiveness and the efficiency of the proposed method in different electronic voting systems.

Keywords: Blind Signature; Electronic Voting; Homomorphism; Secret Sharing; SMPC

1 Introduction

With the development of electronic information technology, voting is converted from paper voting to electronic voting [4, 5, 13]. Electronic voting has the advantages of fast and accurate counting, saving human resources and expenditure, and easy use of voting, which cannot be achieved by traditional voting ways. In the electronic voting scheme, the security and reliability of the voting process are guaranteed by cryptography theory. The current electronic voting system can be roughly divided into three categories [12,21,30]. The first category is that voters fill in the vote information on the voting paper after arriving at the polling station. Then a piece of electronic equipment is used to count the paper voting. The second one is that the voting system requires voters to go to the polling station to use the voting machine to complete the voting. The third one is that the voting system allows voters to use arbitrary terminals, such as computers, mobile phones, and pads, to complete voting online through the Internet. This paper focuses on addressing the issues in the third category. No matter which form of electronic voting, ensuring the security of the voting process is the most critical requirement. A secure and reliable electronic voting protocol is urgently demanded, so electronic voting based on cryptography has been a hot topic in the field of electronic voting.

Since Chaum proposed the first electronic voting scheme in 1980s [8], more cryptographic electronic voting schemes have been proposed, which includes the mix-net of electronic voting, the blind signature of electronic voting, the homomorphic encryption of electronic voting, and the secret sharing of electronic voting. In the scheme based on mix-net [18, 22, 26], the voting manager is the server of the voting site. Every server needs zero-knowledge proof, and thus it costs a lot. The scheme based on a blind signature cannot solve such problems due to voting abstention, collision avoidance and no receipt [15, 16, 20, 23, 29]. The scheme based on homomorphic encryption cannot satisfy large-scale voting activities [2, 9, 10, 19, 25]. Electronic voting based on secret sharing needs to split votes into multiple secret shares and distribute them, which has high communication complexity [3, 6, 14, 24, 27, 28, 31, 34].

In 1992, Fujioka et al. proposed a famous electronic voting scheme (FOO) using blind signature technology [15], which focuses on some security problems in large-scale electronic voting activities. Many famous electronic voting systems, such as Sensus [11,23,30] and E-Vox [17,23,30], have been developed based on FOO electronic voting scheme. However, it is still unable to solve the problems of vote collision and abstention.

In 2017, Quanyu Zhao et al. proposed an electronic voting scheme based on secret sharing and k-anonymity (SSK) [34]. The scheme satisfies the security requirements of electronic voting to a certain extent, but the scheme shares secret shares among different voters. With the increase of voting scale, the efficiency decreases gradually.

In this paper, the blind signature is used to authenticate voters, and secret sharing based on secure multi-party computation is used to realize the voting process, which guarantees the efficiency and security of the voting system.

2 Composition and Security Requirements for Electronic Voting Systems

A complete electronic voting system should consist of four parts [13,30]: registry institution, vote issuing institution, voting institution and counting institution. As shown in Figure 1, modules can be added or reduced according to the actual situation in practice.

1) Registry institution. The voting qualification of voters shall be examined in accordance with the relevant provisions in the electoral process, and relevant certificates shall be issued to those who meet the requirements.

2) Vote issuing institution. The legitimate vote is distributed to legitimate voters who are eligible to vote.

3) Voting institution. The client of voters register accounts, vote, poll results query and other functions.

4) Counting institution. In the counting phase, the validity of the vote papers is checked and the final voting results are counted.

Usually, a complete electronic voting process is as follows:

Step 1. Voters who have the right to vote to apply for an account with the registry.

Step 2. After the registration institution receives the application for registration account, it examines the applicant's voting qualifications and issues a digital certificate if it is satisfied. Otherwise, it rejects the application.

Step 3. Voting issuing agencies send vote to eligible voter.

Step 4. After receiving the blank vote, the voter votes and sends digital certificate and vote to the counting institution.

Step 5. After the voting is over, the counting institution examines the vote, counts the legitimate votes and publishes the results of the voting.

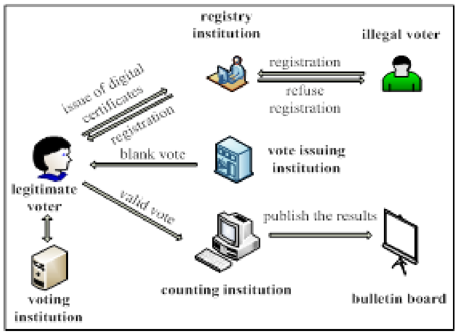


Figure 1: Composition of electronic voting

The specific electronic voting scheme includes entities that must adjust to specific requirements, and the voting process may be expanded or reduced. The security and reliability of every link in the process of electronic voting have always been the most important attribute and characteristic of the electronic voting system. Nowadays, some mature electronic voting systems are mostly designed based on FOO electronic voting protocol [15]. FOO protocol defines seven security requirements of electronic voting, which are regarded as the basic security requirements of electronic voting, as follows:

1) Completeness. All legitimate votes should be counted correctly and no invalid votes should be counted.

2) Soundness. It can resist the illegal actions of malicious voters.

3) Privacy. No third party or dishonest voter can obtain the identity information of the voters and disrupt the voting process.

4) Unreusability. Only one legitimate vote is counted correctly for each voter.

5) Eligibility. Voters participating in the voting shall be authorized to obtain voting qualifications in advance.

6) Fairness. In the process of electronic voting, no third party can be informed of the intermediate

7) Verifiability. Voters can check that their votes are correctly counted.

In addition to the above seven basic require-ments, there are higher security requirements, such as: no receipt, overall verifiability, and multi-party privacy.

3 Preliminaries

In this section, we will briefly introduce blind signature, secure multi-party computation, Shamir'(t, n) secret sharing and homomorphism proof.

3.1 Blind Signature

Blind signature technology is a special signature technology, which was first proposed by Professor Chaum in 1982 [7]. This technology has two characteristics:

1) The message provider performs blind transformation ahead of time by randomly generating blind factor. The signer signs the message after blind transformation and cannot obtain the correct message content.

2) The message provider can get the message signed by the signer using the blind inverse transformation of the signed message. This protects the message that needs to be signed from being leaked. Figure 2 shows the whole process of blind signature, where Alice is a message provider and Bob is a signer.

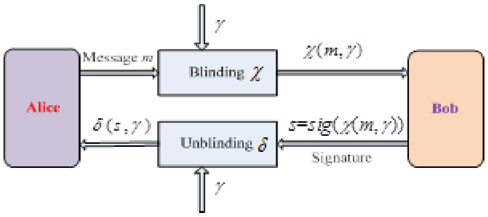


Figure 2: Processing of blind signature

3.2 Secure Multi-party Computation

Secure Multi-Party Computation (SMPC) is first proposed by Yao.A.C who is a Turing Award Winner in 1986 [32, 33]. The mathematical definition of Secure Multi-Party Computation is assumed that n participants want to compare their secrets without revealing their secrets. We denote a set of participants, $p = \{p_1, p_2, \cdots, p_n\}$, and each participant $p_i(1 \leq i \leq n)$ has a secret x_i, then we can obtain the result $f(x_1, x_2, \cdots, x_n)$ via the secure multi-party computation f. The purpose of secure multi-party computing is to obtain some output of secret-related results without revealing the secret, as shown in Figure 3.

Secret sharing is a typical method to realize secure multi-party computing. This paper mainly focuses on how to apply the secret sharing in electronic voting.

3.3 Shamir'(t, n) Secret Sharing

Shamir'(t, n) secret sharing scheme [3, 27] is based on Lagrange interpolation formula, which consists of three phases.

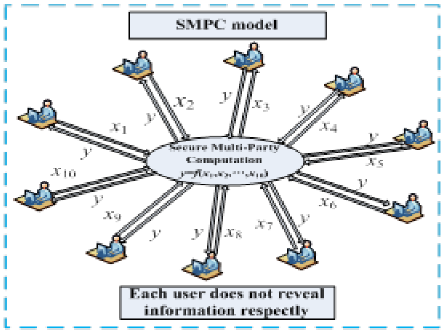

SMPC model
Secure Multi-Party Computation
y=$f(x_1,x_2,\cdots,x_{10})$
x_1
x_2
x_3
x_4
x_5
x_6
x_7
x_8
x_9
x_{10}
y
Each user does not reveal information respectly

Figure 3: The model of SMPC

1) Initialization phase:
Secret distributor D randomly select n different non-zero elements $x_1, x_2, \cdots, x_n$ from $GF(q)$, where $GF(\cdot)$ denotes q is prime and $q > n$. D assigns x_i to $P_i(i = 1, 2, \cdots, n)$, P_i is a participant. The value of x_i is public.

2) Secret distribution phase:
If D intends to have n participants $P_1, P_2, \cdots, P_n$ shares secret $s \in GF(q)$. D randomly chooses $t-1$ elements $a_1, a_2, \cdots, a_{t-1}$ from $GF(q)$. Then constructing $t - 1$ polynomials $f(x) = s + a_1x + a_2x^2 + \cdots + a_{t-1}x^{t-1}$, D calculates $y_i = f(x_i)$ $(1 \leq i \leq n)$ and assigns y_n to participant P_i as sub-secret safely.

3) Secret recovery phase:
Any t participants in the n participants may be set as $P_1, P_2, \cdots, P_t$, show their sub-secrets, and get t pairs of points (x_1, y_1), (x_2, y_2), $\cdots$, (x_t, y_t), so the polynomial $f(x)$ and the shared secret value s can be reconstructed, as shown in Equations (1) and (2).

$$f(x) = \sum_{i=1}^{t} y_i \prod_{j=1, j \neq i}^{t} \frac{x - x_j}{x_i - x_j} \quad (1)$$

$$s = \sum_{i=1}^{t} y_i \prod_{j=1, j \neq i}^{t} \frac{x_j}{x_j - x_i} \quad (2)$$

4 Homomorphism of Shamir'(t, n) Secret Sharing

The homomorphism of secret sharing is pro-posed by Beneloh [1] in 1987. If two participants have their own secret s_1 and s_2, they want to obtain the result of the operation $s_1 \oplus s_2$. D obtains the secret share of two participants through mathematical operation, and then distributes secret shares to all participants. All or part of the participants can get the result of $s_1 \oplus s_2$ directly after showing their secret shares, without restoring the secret s_1 and s_2 of the two participants. The secret sharing scheme satisfying the above conditions is $\oplus$ homomorphism.

Theorem 1. *The shamir'(t, n) secret sharing is additive homomorphism in $x_i (i = 1, 2, \cdots, n)$*

Proof. If participants P_a and P_b need to share their secrets a and b. D randomly selects n different non-zero elements $x_1, x_2, \cdots, x_n$ from $GF(q)$ (q is prime and $q > n$) and D assigns x_i to P_a and $P_b (i = 1, 2, \cdots, n)$, then D randomly selects $a_1, a_2, \cdots, a_{t-1}$ from $GF(q)$ and randomly selects $b_1, b_2, \cdots, b_{t-1}$ from $GF(q)$.

We have

$$\begin{aligned} f_a(x_i) &= a + a_1 x_i + a_2 x_i^2 + \cdots + a_{t-1} x_i^{t-1} \\ f_b(x_i) &= b + b_1 x_i + b_2 x_i^2 + \cdots + b_{t-1} x_i^{t-1} \end{aligned}$$

D shares $f_a(x_i)$ and $f_b(x_i)$ $(i = 1, 2, \cdots, n)$ to n participants respectively, Any t of n participants can recover a and b respectively, and then calculate $(a + b)$. Each participant can also calculate the sum of their secret share, then recover the result $(a + b)$.

$$\begin{aligned} f_a(x_i) + f_b(x_i) &= a + a_1 x_i + a_2 x_i^2 + \cdots + a_{t-1} x_i^{t-1} \\ &\quad + b + b_1 x_i + b_2 x_i^2 + \cdots + b_{t-1} x_i^{t-1} \\ &= (a + b) + (a_1 + b_1) x_i + (a_2 + b_2) x_i^2 \\ &\quad + \cdots + (a_{t-1} + b_{t-1}) x_i^{t-1} \end{aligned}$$

It is clear that

$$\begin{aligned} a + b &= \sum_{i=1}^{t} f_a(x_i) \prod_{j=1, j \neq i}^{t} \frac{x_j}{x_j - x_i} \\ &\quad + \sum_{i=1}^{t} f_b(x_i) \prod_{j=1, j \neq i}^{t} \frac{x_j}{x_j - x_i} \\ &= \sum_{i=1}^{t} (f_a(x_i) + f_b(x_i)) \prod_{j=1, j \neq i}^{t} \frac{x_j}{x_j - x_i} \end{aligned}$$

Thus, the shamir'(t, n) method is additive homomorphism in x_i $(i = 1, 2, \cdots, n)$. □

5 An Electronic Voting Scheme

The scheme is divided into four phases: preparatory phase, authorization verification phase, voting phase and counting phase.

5.1 Preparatory Phase

Suppose there are q voters, denoted as $V_1, V_2, \cdots, V_q$. The votes are completed by q voters as $s_1, s_2, \cdots, s_q$. The voter V_i $(i = 1, 2, \cdots, q)$ selects and fills a vote s_i. Then, it generates a random number k_i as the key of the bit commitment and encrypts the vote with bit commitment function. The result is as $s_i = \epsilon(s_i, k_i)$. The voter V_i randomly selects a blind factor r_i and executes the blinding processing for s_i using blind signature function χ as $s_i = \chi(s_i, r_i)$. At last, V_i sends s_i to the management center (MC).

5.2 Authorization Verification Phase

When MC receiving the information s_i from the voter V_i, it checks whether the voter V_i has the right to vote, and refuses to verify the request if V_i does not have the right to vote. If the voter V_i has a legitimate right to vote, MC checks whether V_i has applied for the verification request. If the voter has applied, MC refuses to verify the application. Then MC signs s_i, as $s_i = \rho(s_i)$, and sends s_i to the voter V_i as an authorization certificate and generates an ID_i which is unique to send to the voter V_i at the same time. Finally, MC publishes the total number of successful authorized verification signatures, and sends ID_i to each vote counter.

5.3 Voting Phase

Step 1. Voter V_i retrieves the information s_i by $s_i = \delta(s_i, r_i)$, where δ is a de-blinding function, and then it checks the correctness of the signature of s_i. If is correct, go to the next step, or go back to the preparatory phase.

Step 2. Suppose there are n vote counters, de-noted as $C_1, C_2, \cdots, C_n$. The votes completed by q voters as $s_1, s_2, \cdots, s_q$. Each voter V_i $(i = 1, 2, \cdots, q)$ divides the vote into n secret shares by using Shamir (t, n) through secret distributor (SD), where x_i, $(i = 1, 2, \cdots, n)$ is public.

$$\begin{cases} f_{V_1}(x_i) = s_1 + s_{11} x_i + s_{12} x_i^2 + \cdots + s_{1(t-1)} x_i^{t-1} \\ f_{V_2}(x_i) = s_2 + s_{21} x_i + s_{22} x_i^2 + \cdots + s_{2(t-1)} x_i^{t-1} \\ \vdots \\ f_{V_q}(x_i) = s_q + s_{q1} x_i + s_{q2} x_i^2 + \cdots + s_{q(t-1)} x_i^{t-1} \end{cases}$$

Step 3. SD sends each voter V_i $(i = 1, 2, \cdots, q)$ of secret share $f_{V_i}(x_j)$ $(j = 1, 2, \cdots, n; i = 1, 2, \cdots, q)$ and ID_i to the vote counter C_j $(j = 1, 2, \cdots, n)$ through a secure channel as shown in Figure 4.

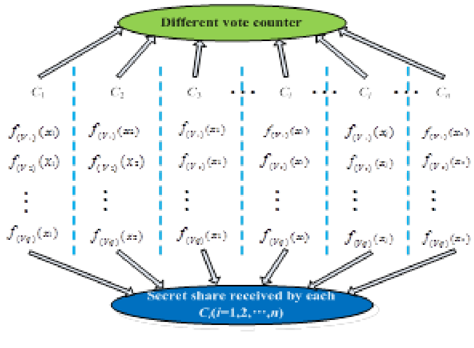


Figure 4: Secret distribution device sends secret share to voter counter

5.4 Counting Phase

Step 1. Each vote counter checks whether the ID number exists for the secret share received, discards it if it does not exist, and performs the following operations if it exists. After vote counter C_i $(i = 1, 2, \cdots, n)$ receiving each voter of secret share $f_{V_j}(x_i)$ $(i = 1, 2, \cdots, n; j = 1, 2, \cdots, q)$, it calculates the sum of each secret share via Equation (3).

$$\begin{aligned} f(C_i) &= f_{V_1}(x_i) + f_{V_2}(x_i) + \cdots, f_{V_q}(x_i) \\ &= s_1 + s_{11}x_i + s_{12}x_i^2 + \cdots + s_{1(t-1)}x_i^{t-1} \\ &\quad +s_2 + s_{21}x_i + s_{22}x_i^2 + \cdots + s_{2(t-1)}x_i^{t-1} \\ &\quad + \cdots + s_q + s_{q1}x_i + s_{q2}x_i^2 + \cdots \\ &\qquad +s_{q(t-1)}x_i^{t-1} \\ &= (s_1 + s_2 + \cdots + s_q) + (s_{11} + s_{21} + \cdots \\ &\quad +s_{q1})x_i + (s_{12} + s_{22} + \cdots + s_{q2})x_i^2 \\ &\quad + \cdots + (s_{1(t-1)} + s_{2(t-1)} + \cdots \\ &\qquad +s_{q(t-1)}x_i^{t-1} \end{aligned} \quad (3)$$

By transforming Equation (3), the final results can be restored through t of n vote counters.

Step 2. The final voting result $s_1 + s_2 + \cdots + s_q$ can be restored by t of the n vote counters and announced on the bulletin board via the following equation:

$$s_1 + s_2 + \cdots + s_q = \sum_{i=1}^{t} f(C_i) \prod_{j=1, j\neq i}^{t} \frac{x_j}{x_j - x_i}$$

Figure 5 is an electronic vote process based on secret sharing. The secret distributor is on the voter terminal, and the voter himself can decide how to share the secret.

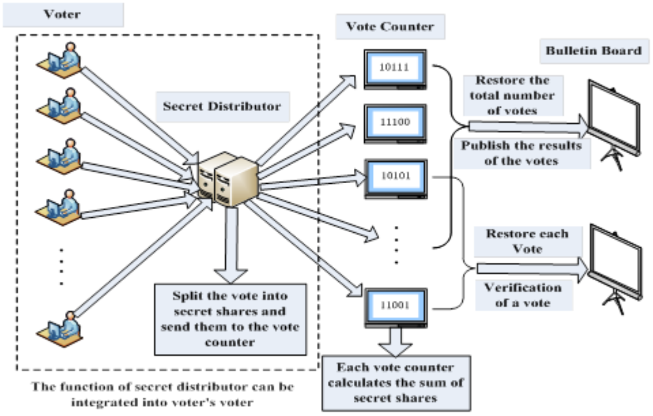

Voter
Vote Counter
Bulletin Board
Secret Distributor
10111
11100
10101
11001
Restore the total number of votes
Publish the results of the votes
Restore each Vote
Verification of a vote
Split the vote into secret shares and send them to the vote counter
Each vote counter calculates the sum of secret shares
The function of secret distributor can be integrated into voter's voter

Figure 5: Complete secret sharing process

6 Security Analysis

Completeness. Vote of voter is successfully verified by MC will be secretly shared to the vote counter, and the vote that has not been verified will not be shared.

Soundness. In the preparatory phase, each voter's identity needs to be verified and authorized to sign. Unqualified voters will be rejected. In the voting phase, the vote counter only accepts the secret share of authorized signature, and the secret share without verification will not be accepted.

Privacy. Third-party organizations or malicious voters will not be able to obtain the information of the vote. The vote will be shared in the form of secret shares to the vote counter. In the process of voting, they will not be able to obtain a single vote and the results of votes.

Unreusability. Because each voter has a validated signature and a unique ID number, it can ensure that only one valid vote can be counted for each voter.

Eligibility. Each voter can not vote unless sign blindly to obtain authentication before voting.

Fairness. In this voting process, no one else can get votes or interfere with the voting process.

Verifiability. Before the end of the vote, if a voter doubts whether his or her vote is recorded in the final number of votes, the vote counters can restore the votes of the voter according to ID number using Equation (4).

$$s_i = \sum_{j=1}^{t} f_{V_i}(x_j) \prod_{j=1, j\neq i}^{t} \frac{x_j}{x_j - x_i} \quad (4)$$

7 Experimental Study

This section compares the efficiency of the proposed scheme(BSS) with FOO [15] and SSK [34] schemes for different number of voters. The efficiency of 5000 and 20000 voters for a different number of vote counters is compared. At the same time, the number of secret shares between BSS and SSK scheme is compared. The experiments are based on the assumption of voting at the same time. The experimental process simulates choosing one of the three candidates and fills in the votes using a random strategy. The experiments were implemented on a machine with dual-core 2.3 GHz CPU and 8 GB memory running 64 bit Ubuntu 18.04.

1) Figure 6 shows the efficiency result for a different number of voters. The number of vote counters in this algorithm is set to 10.

   The experimental results show that the efficiency of the proposed scheme is higher than FOO and SSK, with the increasing number of voters, the efficiency advantage is more obvious.

2) Figure 7 shows the results of time efficiency of different number of vote counters for 5000 and 20000 voters.

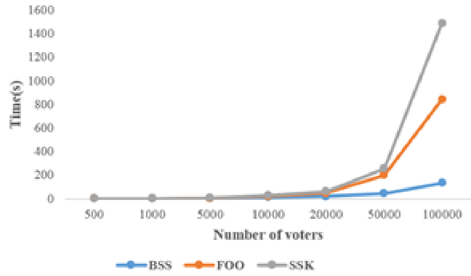


Figure 6: Comparison of three schemes for different number of voters

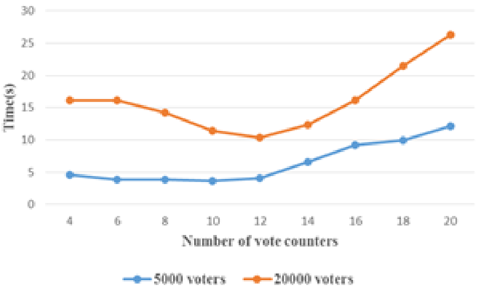


Figure 7: Time efficiency of different number of vote counters for 5000 and 20000 voters

It shows that the number of vote counters will affect the efficiency of this scheme. With the gradual increase of vote counter, the time efficiency first decreases and then increases, and it exists an optimum value.

3) Compared with the number of secret shares of BSS and SSK, setting n voters, k vote counters, the number of secret shares of BSS is kn, and the number of secret shares of SSK is n^2. With the increase of voters, the growth rate of secret shares of SSK algorithm is significantly higher than BSS, resulting in a large number of redundancy of secret shares and affecting efficiency.

8 Conclusions

This paper introduced in detail the necessary composition and security requirements of electronic voting, blind signature, Shamir'(t, n) secret share, and proposed an electronic voting scheme based on blind signature and Shamir'(t, n) secret share. This scheme uses blind signature technology to verify the identity of voters, and secretly shares votes to vote counters in the voting process. The vote counters then perform homomorphic addition operation on the secret share received and finally recovers the final result of voting. By the theoretical analysis, the method meets the relevant security requirements. At the same time, the experimental study shows that the BBS is more efficient than FOO and SSK in the three-in-one electronic voting system. Furthermore, the efficiency of BBS is affected by the number of vote counters, and there is an optimum value for the number of vote counters. At last, the number of secret shares of BBS is significantly lower than that of SSK as the number of voters increases.

The method in this paper only validates the single candidate. Whether the multi-candidate electronic voting is applicable or not remains to be investigated. Also, the number of vote counters should be considered in the security and scale of the electronic voting process. The electronic voting process is not an ideal model in practice. Robustness of the method to different requirements and characteristics of electronic voting is what we need to be considered and studied in the future.

Acknowledgments

This study was supported by the Guangxi Key Laboratory of Hybrid Computation and IC Design Analysis (HCIC201701).

References

[1] J. C. Benaloh, "Secret sharing homomorphisms: Keeping shares of a secret secret," in *Advances in Cryptology (CRYPTO'86)*, vol. 263, no. 1, pp. 251-260, 1986.

[2] J. Benaloh, M. J. Fischer, "A robust and verifiable cryptographically secure election scheme," in *Proceedings of the 26th IEEE Symposium on the Foundations of Computer Science*, pp. 372-382, 1985.

[3] G. Blakley, "Safeguarding cryptographic," in *AFIPS National Computer Conference*, pp. 313-317, 1979.

[4] J. W. Bryans, B. Littlewood, P. Y. A. Ryan and L. Strig-ini, "E-voting: Dependability requirements and design for dependability," in *International Conference on Availability*, pp. 1-8, 2006.

[5] Caltech/MIT Voting Technology Project. (`http://www.votingtechnologyproject.org/`)

[6] H. Chahar, B. N. Keshavamurthy, M. Chirag, "Privacy-preserving distributed mining of association rules using Elliptic-curve cryptosystem and Shamir's secret sharing scheme," *Academy Proceedings in Engineering Sciences*, vol. 42, no. 4, pp. 1997-2007, 2017.

[7] D. Chaum, "Blind signatures for untraceable payments," in *Advances in Cryptology Proceedings of Crypto*, pp. 199-203, 1982.

[8] D. Chaum, "Untraceable electronic mail, return address, and digital pseudonyms," *Communications of the ACM*, vol. 24, no. 2, pp. 84-90, 1981.

[9] I. Chillotti, N. Gama, M. Georgieva, M. lzabachene, "A homomorphic LWE based e-voting scheme," in *Proceedings of 7th International Workshop*, pp. 245-265, 2016.

[10] R. Cramer, R. Gennaro, B. Schoenmakers, "A secure and optimally efficient multi-authority election

scheme," in *International Conference on the Theory and Applications of Cryptographic Techniques*, pp. 103-118, 1997.

[11] L. F. Cranor, R. K. Cytron, "Sensus: Asecurity - conscious electronic polling system for the Internet," in *Proceedings of 30th Annual Hawaii Conference on System Sciences*, pp. 561-570, 1997.

[12] G. Dini, "A secure and available electronic voting service for a large-scale distributed system," *Future Generation Computer Systems*, vol. 19, no. 1, pp. 69-85, 2003.

[13] J. Epstein, "Electronic voting," *Computer*, vol. 40, pp. 8, pp. 92-95, 2007.

[14] N. Fazio, R. Gennaro, T. Jafarikhah, W. E. S. III, "Homomorphic secret sharing from paillier encryption," in *International Conference on Provable Security*, pp. 381-399, 2017.

[15] A. Fujioka, T. Okamoto, K. Ohta, "A practical secret voting scheme for large scale elections," in *Proceedings of Advances in Cryptology-AUSCRYPT*, pp. 244-251, 1992.

[16] P. Grontas, A. Pagourtzis, A. Zacharakis, "Coercion resistance in a practical secret voting scheme for large scale elections," in *Proceedings of the 14th International Symposium on Pervasive Systems, Algorithms and Networks & The 11th International Conference on Frontier of Computer Science and Technology & The Third International Symposium of Creative Computing*, 2017. DOI: 10.1109/ISPAN-FCST-ISCC.2017.79.

[17] M. A. Herschberg, "Secure electronic voting over the world wide web," *Master Thesis in Electrical Engineering and Computer Science*, 1997. (http://hdl.handle.net/1721.1/43497)

[18] N. Islam, K. M. R. Alam, S. Tamura, Y. Morimoto, "A new e-voting scheme based on revised simplified verifiable re-encryption mixnet," in *International Conference on Networking, Systems and Security (NSysS'17)*, 2017. DOI: 10.1109/NSysS.2017.7885795.

[19] B. Lee, K. Kim, "Receipt-free electronic voting through collaboration of voter and honest verifier," in *Proceedings of JWISC*, pp. 101-108, 2000.

[20] L. López-García, L. J. D. Perez, F. Rodríguez-Henríquez, "A pairing-based blind signature e-voting scheme," *The Computer Journal*, vol. 57, no. 10, pp. 1460-1471, 2014.

[21] M. Mursi, G. M. R. Assassa, A. Abdelhafez, K. M. A. Samra, "On the development of electronic voting: A survey," *International Journal of Computer Applications*, vol. 61, no. 16, pp. 1-11, 2013.

[22] C. A. Neff, "A verifiable secret shuffle and its application to e-voting," in *Proceedings of the 8th ACM conference on Computer and Communications Security*, pp. 116-125, 2011.

[23] M. Ohkubo, F. Miura, M. Abe, A. Fujioka, T. Okamoto, "An improvement on a practical secret voting scheme," in *International Workshop on Information Security*, pp. 225-234, 1999.

[24] L. J. Pang, Q. Q. Pei, H. X. Li, Q. J. Xu, *Secret Sharing Technology and Its Applications*, Posts and Telecommunications Press, Beijing, 2017

[25] K. Peng, R. Aditya, C. Boyd, E. Dawson, B. Lee, "Multiplicative homomorphic e-voting," in *Progress in Cryptology*, pp. 61-72, 2004.

[26] K. Sako, J. Kilian, "Receipt-free mix-type voting scheme - a practical solution to the implementation of a voting booth," in *Proceedings of the International Conference on the Theory and Application of Cryptographic Techniques (EUROCRYPT'95)*, pp. 393-403, 1995.

[27] A. Shamir, "How to share a secret," *Communications of ACM*, pp. 22, no. 11, pp. 612-613, 1979.

[28] R. Shi, L. Huang, Y. Luo, H. Zhong, "A threshold multi-secret sharing scheme," in *IEEE International Conference on Networking*, 2008. DOI: 10.1109/ICNSC.2008.4525497.

[29] C. Y. Song, C. R. Zhang, S. Cao, "Blind signature scheme and its application in electronic voting protocol," *Computer Engineering*, vol. 38, no. 6, pp. 139-141, 2012.

[30] M. H. Sun, *Modern cryptography research on secure multi-party Computation protocols*, Publishing House of Electronics Industry, Beijing, 2016

[31] C. C. Yang, T. Y. Chang, M. S. Hwang, "A (t, n) mul-ti-secret sharing scheme," *Applied Mathematics and Computation*, vol. 151, no. 2, pp. 483-490, 2004.

[32] A. C. Yao, "Protocols for secure computations," in *Proceedings of 23th Annual IEEE Symposium on Foun-dations of Computer Science*, pp. 160-164, 1982.

[33] A. C. Yao, "How to generate and exchange se-crets," in *Proceedings of the 27th Annual Symposium on Foundations of Computer Science*, pp. 162-167, 1986.

[34] Q. Zhao, Y. Liu, "E-voting scheme using secret sharing and K-anonymity," *World Wide Web*, vol. 22, pp. 1657-1667, 2018.

Biography

Hongquan Pu received the M.S. degree in computer application technology from University of Chinese Academy of Sciences in 2014. He is currently a Ph.D. candiadate in University of Chinese Academy of Sciences. His current research interests include Electronic voting, Secret Sharing, LUC Secret System, Secure Multi-Party Computation(SMPC).

Zhe Cui received the degree of Bachelor in Electronic Precision Machinery from University of Electronic Science and Technology of China in 1992. He received the M.S. degree in Computer Application Technology from Chengdu Institute of Computer Applications, Chinese Academy of Sciences in 1995. He received the Ph.D. degree in Computer Software and Theory from Chengdu Institute of Computer Applications, Chinese Academy of Sciences in 2011. He is currently a Ph.D. supervisor at the

University of Chinese Academy of Sciences. The main research fields include pattern recognition and information security.

Ting Liu received the M.S. degree in Computer Software and Theory from Xi'an Technological University in 2011. He is currently a Ph.D. candiadate in University of Chinese Academy of Sciences. His research interests include Electronic voting, Blockchain and Secret Sharing.

Elliptic Curve Scalar Multiplication Algorithm Based on Side Channel Atomic Block over GF(2^m)

Shuang-Gen Liu, Yan-Yan Hu, and Lan Wei
(Corresponding author: Shuang-Gen Liu)

School of Cyberspace Security, Xi'an University of Posts and Telecommunications
Xi'an 710121, China
Email: liusgxupt@163.com

(Received Feb. 28, 2020; Revised and Accepted Dec. 10, 2020; First Online May 30, 2021)

Abstract

Elliptic Curve Cryptography (ECC) has become one of the research hotspots in cryptography in recent years. Scalar multiplication is the most crucial operation in ECC, and it largely determines the efficiency of ECC. To improve ECC's speed, we propose a new secure and efficient scalar multiplication algorithm on elliptic curves over GF(2^m). In addition, we present the new composite operation formulas $3P_1$ and $2P_1 + P_2$ using only x-coordinate, where P_1 and P_2 are points on an elliptic curve. To ensure the safety and efficiency of the proposed algorithm, we constitute an atomic block by adding dummy operations and using the Montgomery trick.

Keywords: Elliptic Curve Cryptography; Scalar Multiplication; Side Channel Atomic Block; Simple Power Analysis

1 Introduction

As the hacking techniques become more and more powerful, safe and efficient encryption technology is needed. Since Miller [17] and Koblitz [9] independently proposed Elliptic Curve Cryptography (ECC) in 1985, it has become one of the research hotspots in the field of cryptography due to its short key and high security. ECC can provide the same functions as the RSA cryptosystem and it requires a shorter key length than RSA under the same security. It is generally used for digital signature, authentication, encryption, decryption [8, 19, 25]. Because of its advantages in security, encryption and decryption performance, and space consumption, ECC has a wide range of applications, such as transport layer security (TLS), cryptocurrency, SM2 public key cryptography and government agencies, etc. Besides, it is especially suitable for environments with limited storage resources, such as smart cards and secure storage chips.

In ECC, it is easy to obtain the point Q when $Q = kP$, and the number k and point P are given. But it is difficult to find k when point P and point Q are given. This is the classical discrete logarithm problem (DLP). ECC uses this feature to encrypt where point Q is the public key, big number k is the private key and point P is the base point on an elliptic curve. The most crucial operation in ECC is scalar multiplication kP that largely determines the speed of ECC. There are two main methods to improve the efficiency of scalar multiplication. The first method is to reduce computation by optimizing the bottom arithmetic formulas, such as reducing the number of field inversion operations by transforming coordinates. The second method is to decrease the number of point addition and doubling in the scalar multiplication algorithm by studying the expanded form of k, such as double-base chain [27] and symmetric ternary form (STF) [13].

Side channel analysis (SCA) is a method to attack the cryptographic devices by analyzing the leaked side channel information such as time consumption, power consumption or electromagnetic radiation during the operation of cryptographic devices [24]. Power analysis is a form of SCA. It is an attack by collecting power consumption information generated by cryptographic devices or cryptographic chips during encryption, decryption or signature operations, and analyzing the key by using statistics, cryptography and other relevant knowledge. Power analysis can be divided into simple power analysis (SPA) and differential power analysis (DPA). SPA has a direct threat to cryptographic devices. It can directly analyze the power information collected during the execution of cryptographic algorithm. When the device performs encryption or decryption, the key can be derived from the difference in power consumption trajectories. The key in this paper refers to the private key k.

In 1987, the Montgomery algorithm was proposed by Montgomery [18]. The basic idea is that each loop has a point addition and doubling so that the energy consumed by each loop is basically the same. In 1999, Lopez and Dahab [16] optimized the Montgomery ladder algorithm on elliptic curves over GF(2^m). The new point addition and doubling formulas eliminated the calculation of y-coordinate, which improved the calculation speed of the algorithm. In 2008, the new point addition and doubling formulas proposed by Yu *et al.* [**?**] not only omitted the y-coordinate but also dislodged the field inversion operation. In 2013, Sung *et al.* [5] posed the new composite formulas $4P_1$, $3P_1 + P_2$ and $2P_1 + 2P_2$ with only x-coordinate, and presented the extended quaternary Montgomery ladder algorithm over GF(2^m). In 2016, Lai and Zhang [10] proposed Co_Z point addition algorithm, conjugate point addition algorithm and point doubling-point addition algorithm with omitting Z-coordinate on Hessian elliptic curves and applied them to the traditional Montgomery ladder algorithm. In 2017, Yu *et al.* [26] optimized the Montgomery algorithm using the Co_Z technique in projective coordinates over GF(3^m). In 2019, Liu *et al.* [14] proposed the ternary Montgomery ladder algorithm, which combines the original Montgomery ladder algorithm with the ternary representation of the scalar k.

To obtain a safe and efficient scalar multiplication algorithm, we first propose the new composite operation formulas $3P_1$ and $2P_1 + P_2$ using only x-coordinate in affine coordinate system to reduce the bottom field operations and we apply them to the ternary Montgomery ladder algorithm. Then we constitute an atomic block by adding dummy operations to the proposed composite operation formulas to prevent SPA. Last, we use Montgomery trick in the atomic block to optimize the computational cost, which can decrease the number of field inversion operations.

The rest of this paper is presented as follows. In section II, we briefly introduce Elliptic Curve Cryptography and the Montgomery ladder algorithm. In section III, we give a detailed presentation on new composite operation formulas and the anti-SPA scalar multiplication algorithm based on side channel atomic block. In section IV, we compare the performance of the proposed algorithm with existing algorithms.

2 Elliptic Curve Cryptography and Montgomery Ladder Algorithm

2.1 Elliptic Curve Cryptography

Definition 1. *The equation of a non-super singular elliptic curve E over GF(2^m) is given as follows:*

$$E/GF(2^m) : y^2 + xy = x^3 + ax^2 + b. \tag{1}$$

with $a, b \in GF(2^m), b \neq 0$. All points on E and the infinity point $\mathcal{O}$ form an abelian group. Assume $P_1 = (x_1, y_1) \in E(GF(2^m))$, $P_2 = (x_2, y_2) \in E(GF(2^m))$, $-P_1 = (x_1, x_1 + y_1)$ and $P_2 \neq -P_1$.

If $P_1 \neq P_2$, $P_3 = P_1 + P_2 = (x_3, y_3)$, then point addition operation:

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

If $P_1 = P_2$, $P_3 = 2P_1 = (x_3, y_3)$, then point doubling operation:

$$\begin{cases} x_3 = (x_1 + \frac{y_1}{x_1})^2 + \frac{y_1}{x_1} + x_2 + a \\ y_3 = (x_1 + \frac{y_1}{x_1})(x_1 + x_3) + x_3 + y_1 \end{cases} \tag{3}$$

It can be seen that the computational costs of point addition and doubling are both $1I + 2M + 1S$, where I, M, S are the representations of field inversion, field multiplication and field squaring, respectively.

2.2 Montgomery Ladder Algorithm

The Montgomery algorithm was initially proposed to improve the speed of scalar multiplication. The left-to-right Montgomery ladder algorithm [20] is described by Algorithm 1, which is a classical way to compute the scalar multiplication.

Algorithm 1 Left-To-Right Montgomery Ladder Algorithm

```
Input: $P = (x, y) \in E(GF(2^m))$, and $k = (k_{n-1}k_{n-2} \cdots k_1k_0)_2$
Output: $Q = kP \in E(GF(2^m))$
$R_0 = P, R_1 = 2P$
for $i \leq n-2, \cdots, 0$ do
    if $k_i = 1$ then
        $R_0 = R_0 + R_1, R_1 = 2R_1$
    else if $k_i = 0$ then
        $R_1 = R_0 + R_1, R_0 = 2R_0$
    end if
end for
Return $Q = R_0$
End
```

Based on the original Montgomery ladder algorithm, Liu *et al.* [14] proposed the ternary Montgomery ladder algorithm, which is described by Algorithm 2.

3 New Algorithm Based on the Ternary Montgomery Ladder Algorithm

3.1 Composite Operation Formulas

Improving the performance of the Montgomery ladder algorithm by using only x-coordinate method was first

Algorithm 2 The Ternary Montgomery Ladder Algorithm

```
Input: $P = (x, y) \in E(GF(2^m))$, and $k = (k_{n-1}k_{n-2}\cdots k_1k_0)_3$, where $k_{n-1} = 1$ or 2
Output: $Q = kP \in E(GF(2^m))$
$R_0 = k_{n-1}P, R_1 = (k_{n-1} + 1)P$
for $i \leq n-2, \cdots, 0$ do
    if $k_i = 0$ then
        $R_2 = 3R_0, R_1 = 2R_0 + R_1$
    else if $k_i = 1$ then
        $R_2 = 2R_0 + R_1, R_1 = 2R_1 + R_0$
    else if $k_i = 2$ then
        $R_2 = 2R_1 + R_0, R_1 = 3R_1$
    end if
    $R_0 = R_2$
end for
Return $Q = R_0$
End
```

introduced by Lopez & Dahab [16]. Then several x-coordinate-only methods were presented [5, 22, 28]. Assume P_i is a point on an elliptic curve E. Let $P_1 = (x_1, y_1)$, $P_2 = (x_2, y_2)$, $-P_1 = (x_1, x_1 + y_1)$, $P_2 - P_1 = P = (x, y)$, and $P_2 \neq -P_1$, then we can obtain

$$x_3 = \begin{cases} x + \frac{x_1}{x_1+x_2} + \left(\frac{x_1}{x_1+x_2}\right)^2 & P_1 \neq P_2 \\ x_1^2 + \frac{b}{x_1^2} & P_1 = P_2 \end{cases} \tag{4}$$

The formula for restoring the y coordinate at the last step is

$$y_1 = (x_1 + x)\{(x_1 + x)(x_2 + x) + x^2 + y\}/(x + y) \tag{5}$$

It can be seen from Equation (4) that the costs of both two operations are $1I + 1M + 1S$. Based on the idea of Lopez & Dahab, this paper proposes two composite operation formulas $3P_1$ and $2P_1 + P_2$.

Theorem 1. *Let $P_1 = (x_1, y_1)$ be a point on an elliptic curve E over $GF(2^m)$. Then, x_{3P_1} can be gained:*

$$x_{3P_1} = x_1 + \frac{{x_1}^3}{{x_1}^4 + {x_1}^3 + b} + \left(\frac{{x_1}^3}{{x_1}^4 + {x_1}^3 + b}\right)^2 \tag{6}$$

with cost $1I + 1M + 3S + 1C$, where C is the representation of field cubing.

Proof. Let $3P_1$ be computed as $2P_1 + P_1$. Equation (4) gives

$$x_{3P_1} = x_1 + \frac{x_{P_1}}{x_{P_1} + x_{2P_1}} + \left(\frac{x_{P_1}}{x_{P_1} + x_{2P_1}}\right)^2 \tag{7}$$

Then, we obtain Equation (6). □

Theorem 2. *Let $P_1 = (x_1, y_1), P_2 = (x_2, y_2)$ be points on an elliptic curve E over $GF(2^m)$. Then, $x_{2P_1+P_2}$ can be gained:*

$$\begin{aligned} x_{2P_1+P_2} = x_2 &+ \frac{x_1{(x_1+x_2)}^2}{(x+x_1){(x_1+x_2)}^2 + x_1x_2} \\ &+ \left(\frac{x_1{(x_1+x_2)}^2}{(x+x_1){(x_1+x_2)}^2 + x_1x_2}\right)^2 \end{aligned} \tag{8}$$

with cost $1I + 2S + 4M$.

Proof. Let $2P_1 + P_2$ be computed as $(P_1 + P_2) + P_1$ and $P_2 - P_1 = P(x, y)$ which is an input. Equation (4) gives

$$x_{2P_1+P_2} = x_2 + \frac{x_{P_1}}{x_{P_1} + x_{P_1+P_2}} + \left(\frac{x_{P_1}}{x_{P_1} + x_{P_1+P_2}}\right)^2 \tag{9}$$

Then, we obtain Equation (8). □

The probability that k_i is equal to 0, 1, and 2 is 1/3 [11]. When Algorithm 2 is computed by Equation (6) and Equation (8), the average calculation costs are $2I + 6M + 14/3S + 2/3C$ per loop.

3.2 New Algorithm Based on Side Channel Atomic Block

In view of SCA attack, in 2004, Mames, ciet and joye proposed a method that almost does not increase the amount of computation: Side channel atomic block method [4]. Its main idea is to decompose the operations on elliptic curves into a series of indistinguishable atomic blocks with multiple side channels. The general method is to add dummy operations so that there is no difference in the side channel analysis of different execution processes.

In this paper, as can be seen from Algorithm 2 that the discrimination of each loop is $3P_1$ and $2P_1 + P_2$. To make it anti-SPA, we can add some dummy operations to $3P_1$ and $2P_1 + P_2$ to make the costs of $3P_1$ and $2P_1 + P_2$ indistinguishable so that the amount of calculations in each loop is exactly the same. In this way, the scalar multiplication can be represented as a series of indistinguishable atomic blocks of code, so the attacker cannot obtain the side channel information by SPA attack.

The Montgomery trick is an efficient way to improve performance by computing field inversions simultaneously. For instance, a^{-1} and b^{-1} can be computed as $a^{-1} = (ab)^{-1} \cdot b$, $b^{-1} = (ab)^{-1} \cdot a$. It converts two field inversion operations into one field inversion operation and three field multiplication operations, which can save $1I - 3M$ calculation costs per loop. Therefore, we apply the Montgomery trick to Equation (6) and Equation (8) in the atomic block to reduce the amount of field inversion operations to optimize the algorithm.

As stated above, we constitute the atomic block by adding dummy operations in each loop to make it anti-SPA and using Montgomery trick to reduce the amount of field inversion operations. Table 1, called the atomic block elliptic curve triple and double-and-add, *i.e.* AETD, describes the atomic block in detail.

In the upper section, Algorithm 2 can be computed efficiently by using the proposed composite operation formulas Equation (6) and Equation (8), with cost $2I+14/3S+6M+2/3C$ per loop. However, the computation costs of AETD just require $1I+4S+10M$. Applying AETD to Algorithm 2, Algorithm 3 is obtained. Algorithm 3 saves $I+2/3S-4M+2/3C$ compared with Algorithm 2 computed by Equation (6) and Equation (8), and saves $3I-2M$ compared with Algorithm 2 computed by Equation (2) and Equation (3) in each loop. From Algorithm 3, we can conclude that only one atomic block is used in each loop, so each loop requires the same amount of calculations regardless of the value of k_i.

Algorithm 3 Anti-SPA Scalar Multiplication Algorithm Based On Side Channel Atomic Block

```
Input: $P = (x,y) \in E(GF(2^m))$, and $k = (k_{n-1}k_{n-2}\cdots k_1k_0)_3$, where $k_{n-1} = 1$ or 2
Output: $Q = kP \in E(GF(2^m))$
$R_0 = k_{n-1}P, R_1 = (k_{n-1}+1)P$
for $i \le n-2, \cdots, 0$ do
    $(R_0, R_1) = AETD[k_i](R_0, R_1)$
end for
Return $Q = R_0$
End
```

4 Performance Analysis

4.1 Security Analysis

In ECC, if a scalar multiplication algorithm has different power consumption according to k_i, it is vulnerable to SPA. In other words, if the algorithm has the same power consumption regardless of the value of k_i, it is resistant to SPA. Therefore, all countermeasures against SPA have to modify the algorithm to obtain a uniform power consumption trace. In general, there are three main ways to anti-SPA. The first way is uniform algorithm behavior, such as Montgomery ladder algorithm. The second way is uniform point addition and doubling formulas, such as Edwards curve [2]. The third way is to add dummy field operations [6].

To improve the efficiency of the ternary Montgomery ladder algorithm, the composite operation formulas $3P_1$ and $2P_1 + P_2$ are proposed. However, the power consumption of $3P_1$ and $2P_1 + P_2$ is different. Algorithm 2 computes $3P_1$ and $2P_1 + P_2$ when k_i is equal to 0 or 2, while it computes $2P_1 + P_2$ twice when k_i is equal to 1. SPA gains the key according to the peak shape of the energy graph [12], so it is easy to obtain the value of k_i by observing the power consumption curve leaked during execution of the algorithm. Therefore, we adopt the third way to add dummy field operations to constitute an atomic block. It can be seen from Table 1 and Algorithm 3 that the field operation of each step of every atomic block is the same and only one atomic block is used in each loop, so the power consumed by each loop is the same whatever $k_i = 0$, 1, or 2, which is secure to resist SPA. In addition, Algorithm 3 can also resist DPA so long as randomize the scalar k.

4.2 Efficiency Analysis

Because the extra calculations of algorithms are negligible, in this paper, we mainly compare the calculations of main iteration of algorithms. In this section, the efficiency of the proposed composite operation formulas and Algorithm 3 is analyzed.

Table 2 shows the computation costs of Algorithm 2 under different calculation formulas. From it, we can draw the conclusion that Algorithm 2 can be computed efficiently by using Equation (6) and Equation (8) proposed in this paper. It requires $2I + 14/3S + 6M + 2/3C$ on average, with saving $2I - 2M - 2/3S - 2/3C$ than Equation (4) and saving $2I + 2M - 2/3S - 2/3C$ than Equation (2) and Equation (3) in each loop.

Given an integer k, assume that $m = \lceil \log_3 k \rceil$ is the length of the ternary representation and $n = \lceil \log_2 k \rceil$ is the length of the binary representation, $m = n\log_3 2$, *i.e.* 160-binary is equivalent to 101-ternary and 192, 256, 600-binary [21] is equivalent to 122, 162, 379-ternary, respectively. We suppose $n = 160$ bits, $m = 101$ bits. According to the experiment of Bernstein [3], we assume $I/M = 8$, $S/M = 0.8$.

Table 3 shows the comparison of Algorithm 3 and existing algorithms over GF(2^m). It can be seen that Algorithm 3 has a good improved efficiency compared with the algorithms of [12, 23] and [15]. In comparison, the improved efficiency of Algorithm 3 is 13.6%, 33.9%, 8.7%, 13.4%, 1.6%, and 15.4%, respectively.

In order to analyze the dynamic changes of the improved efficiency of Algorithm 3 than existing algorithms, we suppose

$$I/M = \beta \tag{10}$$

$S/M = 0.8$. The improved efficiency of Algorithm 3, *i.e.* ε can be given as follows:

$$\varepsilon{=}1 - \frac{(m-1)(\#I_1 + \#M_1)}{\ell(\#I_2 + \#M_2)} \tag{11}$$

$(\#I_1 + \#M_1)$ represents the amount of calculations of Algorithm 3 per loop, and $(\#I_2 + \#M_2)$ represents the amount of calculations of existing algorithms in each loop. $(m-1)$ and ℓ indicate the number of iterations of Algorithm 3 and existing algorithms.

Field inversion operations can be efficiently computed by the Extended Euclidean Algorithm (EEA) over GF(2^m), which uses $gcd(a,b) = gcd(b{+}ca, a)$ for all binary polynomials. According to [1], when the field size is 163 bits, performance of a field inversion operation using the EEA is equal to about 6.67-10.33 field multiplication operations in binary field, which means β is about 6.67-10.33.

Figure 1 shows the comparison of Algorithm 3 and existing algorithms. I/M is the x-axis and the improved

Table 1: The atomic block elliptic curve triple and double-and-add (AETD)

Input:$T_1=P_1=x_1,T_2=P_2=x_2,T_3=P=x$ Output:$(3P_1,2P_1+P_2)$ or $(2P_1+P_2,2P_2+P_1)$ or $(2P_2+P_1,3P_2)$		
$k_i=0$	$k_i=1$	$k_i=2$
$(T_1,T_2)=(3P_1,2P_1+P_2)$	(T_1,T_2)=$(2P_1+P_2,2P_2+P_1)$	$(T_1,T_2)=(2P_2+P_1,3P_2)$
$T_4 \leftarrow T_1+T_2(x_1+x_2)$	$T_4 \leftarrow T_1+T_2(x_1+x_2)$	$T_4 \leftarrow T_1+T_2(x_1+x_2)$
$T_5 \leftarrow {T_1}^2({x_1}^2)$	$T_5 \leftarrow {T_4}^2((x_1+x_2)^2)$	$T_5 \leftarrow {T_4}^2((x_1+x_2)^2)$
$T_6 \leftarrow T_3+T_2(dummy)$	$T_6 \leftarrow T_3+T_2(x+x_2)$	$T_6 \leftarrow T_3+T_2(x+x_2)$
$T_6 \leftarrow T_5 \cdot T_5({x_1}^4)$	$T_6 \leftarrow T_6 \cdot T_5((x+x_2)(x_1+x_2)^2)$	$T_6 \leftarrow T_6 \cdot T_5((x+x_2)(x_1+x_2)^2)$
$T_7 \leftarrow T_1 \cdot T_5({x_1}^3)$	$T_7 \leftarrow T_2 \cdot T_5(x_2(x_1+x_2)^2)$	$T_7 \leftarrow T_2 \cdot T_5(x_2(x_1+x_2)^2)$
$T_5 \leftarrow b$	$T_4 \leftarrow b$	$T_5 \leftarrow b$
$T_8 \leftarrow T_1 \cdot T_2(x_1x_2)$	$T_8 \leftarrow T_1 \cdot T_2(x_1x_2)$	$T_8 \leftarrow T_1 \cdot T_2(x_1x_2)$
$T_5 \leftarrow T_5+T_7(b+{x_1}^3)$	$T_4 \leftarrow T_5+T_7(dummy)$	$T_4 \leftarrow T_5+T_7(dummy)$
$T_5 \leftarrow T_6+T_5(A)$	$T_4 \leftarrow T_6+T_8(A)$	$T_4 \leftarrow T_6+T_8(A)$
$T_6 \leftarrow T_3+T_1(x+x_1)$	$T_6 \leftarrow T_3+T_1(x+x_1)$	$T_6 \leftarrow T_3+T_1(dummy)$
$T_4 \leftarrow {T_4}^2((x_1+x_2)^2)$	$T_3 \leftarrow {T_3}^2(dummy)$	$T_6 \leftarrow {T_2}^2({x_2}^2)$
$T_9 \leftarrow T_1 \cdot T_4(x_1(x_1+x_2)^2)$	$T_9 \leftarrow T_1 \cdot T_5(x_1(x_1+x_2)^2)$	$T_9 \leftarrow T_2 \cdot T_6({x_2}^3)$
$T_4 \leftarrow T_6 \cdot T_4((x+x_1)(x_1+x_2)^2)$	$T_5 \leftarrow T_6 \cdot T_5((x+x_1)(x_1+x_2)^2)$	$T_6 \leftarrow T_6 \cdot T_6({x_2}^4)$
$T_6 \leftarrow T_6+T_9(dummy)$	$T_6 \leftarrow T_6+T_9(dummy)$	$T_6 \leftarrow T_6+T_9$
$T_4 \leftarrow T_4+T_8(B)$	$T_5 \leftarrow T_5+T_8(B)$	$T_5 \leftarrow T_6+T_5(B)$
$T_6 \leftarrow T_5 \cdot T_4(AB)$	$T_6 \leftarrow T_4 \cdot T_5(AB)$	$T_6 \leftarrow T_5 \cdot T_4(AB)$
$T_6 \leftarrow {T_6}^{-1}((AB)^{-1})$	$T_6 \leftarrow {T_6}^{-1}((AB)^{-1})$	$T_6 \leftarrow {T_6}^{-1}((AB)^{-1})$
$T_5 \leftarrow T_6 \cdot T_5(B^{-1})$	$T_4 \leftarrow T_6 \cdot T_4(B^{-1})$	$T_5 \leftarrow T_6 \cdot T_5(A^{-1})$
$T_4 \leftarrow T_6 \cdot T_4(A^{-1})$	$T_5 \leftarrow T_6 \cdot T_5(A^{-1})$	$T_4 \leftarrow T_6 \cdot T_4(B^{-1})$
$T_4 \leftarrow T_4 \cdot T_7(A^{-1}{x_1}^3)$	$T_5 \leftarrow T_5 \cdot T_7(A^{-1}x_2(x_1+x_2)^2)$	$T_4 \leftarrow T_4 \cdot T_7(B^{-1}x_2(x_1+x_2)^2)$
$T_6 \leftarrow {T_4}^2$	$T_6 \leftarrow {T_5}^2$	$T_6 \leftarrow {T_4}^2$
$T_4 \leftarrow T_4+T_6$	$T_5 \leftarrow T_5+T_6$	$T_4 \leftarrow T_4+T_6$
$T_4 \leftarrow T_1+T_4(3P_1)$	$T_5 \leftarrow T_1+T_5(2P_2+P_1)$	$T_4 \leftarrow T_1+T_4(2P_2+P_1)$
$T_5 \leftarrow T_5 \cdot T_9(B^{-1}x_1(x_1+x_2)^2)$	$T_4 \leftarrow T_4 \cdot T_9(B^{-1}x_1(x_1+x_2)^2)$	$T_5 \leftarrow T_5 \cdot T_9(A^{-1}{x_2}^3)$
$T_9 \leftarrow {T_5}^2$	$T_9 \leftarrow {T_4}^2$	$T_9 \leftarrow {T_5}^2$
$T_5 \leftarrow T_5+T_9$	$T_4 \leftarrow T_4+T_9$	$T_5 \leftarrow T_5+T_9$
$T_2 \leftarrow T_2+T_5(2P_1+P_2)$	$T_1 \leftarrow T_2+T_4(2P_1+P_2)$	$T_2 \leftarrow T_2+T_5(3P_2)$
$T_1 \leftarrow T_4(3P_1)$	$T_2 \leftarrow T_5(2P_2+P_1)$	$T_1 \leftarrow T_4(2P_2+P_1)$
$(A={x_1}^4+{x_1}^3+b;$	$(A=(x+x_2)(x_1+x_2)^2+x_1x_2;$	$(A=(x+x_2)(x_1+x_2)^2+x_1x_2;$
$B=(x+x_1)(x_1+x_2)^2+x_1x_2)$	$B=(x+x_1)(x_1+x_2)^2+x_1x_2)$	$B={x_2}^4+{x_2}^3+b)$

Table 2: The computation costs of Algorithm 2 under different calculation formulas

Formulas	$3P_1$	$2P_1+P_2$	Average costs of main iteration	Anti-SPA
(2)(3)	$2I+4M+2S$	$2I+4M+2S$	$4I+8M+4S$	yes
(4)	$2I+2M+2S$	$2I+2M+2S$	$4I+4M+4S$	yes
(6)(8)	$1I+1M+3S+1C$	$1I+4M+2S$	$2I+6M+14/3S+2/3C$	no

Table 3: The computation costs of different scalar multiplication algorithms

Algorithm	Total costs of main iteration	$n=160$bits,$m=101$bits	Anti-SPA
Unprotected NAF [23]	$(10M+20/3S)n$	$2453.3M$	No
Co_Z protected NAF [23]	$(85/6M+265/36S)n$	$3208.9M$	Yes
Mont2 [12]	$3(1I+1M+1S)(n/2-1)$	$2322.6M$	No
Mont3 [12]	$37/24(1I+1M+1S)(n+2)$	$2447.6M$	No
Co_Z STF [15]	$(52/3M+5S)m$	$2154.7M$	No
Co_Z Anti-SPA STF [15]	$(20M+6S)m$	$2504.8M$	Yes
Algorithm 3	$(1I+10M+4S)(m-1)$	$2120M$	Yes

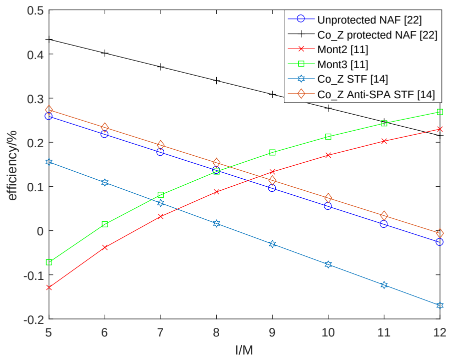


Figure 1: The comparison of Algorithm 3 and existing algorithms

efficiency of Algorithm 3 than existing algorithms is the y-axis. When $I/M = 8$, Table 3 can be obtained. It can be seen from Figure 1, for algorithms of [23] and [15], the improved efficiency of Algorithm 3, *i.e.* ε, decreases linearly as β increases. For algorithms of [12], ε increases as β increases and the larger β, the slower ε increases. When β is 6.67-10.33, Algorithm 3 is more efficient than other algorithms except for Co_Z STF algorithm [15]. However, Algorithm 3 performs better than Co_Z STF algorithm [15] when β is less than 8.3. In summary, Algorithm 3 has a good improvement in efficiency compared with existing algorithms.

5 Conclusions

In this paper, we proposed an anti-SPA scalar multiplication algorithm based on side channel atomic block over GF(2^m). Besides, we have optimized the bottom field operations by presenting new composite operation formulas $3P_1$ and $2P_1 + P_2$. Figure 1 intuitively shows the comparison of the proposed algorithm and existing algorithms. When $I/M = 8$, it can be seen from Table 3 that the proposed algorithm is more efficient than existing algorithms, ranging from 1.6% to 33.9%. Then we can apply it to the specific environments, such as wireless sensor networks with resource-limited. Next, we will try to transform the coordinate to optimize the proposed composite operation formulas and then propose a more efficient scalar multiplication algorithm.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (61872058), the Key Research and Development Program of Shaanxi (Program No.2021NY-211). The authors gratefully acknowledge the anonymous reviewers for their valuable comments.

References

[1] R. Avanzi and N. Thériault, "Effects of optimizations for software implementations of small binary field arithmetic," in *International Workshop on the Arithmetic of Finite Fields*, pp. 69–84, 2007.

[2] D. J. Bernstein and T. Lange, "Faster addition and doubling on elliptic curves," in *International Conference on the Theory and Application of Cryptology and Information Security*, pp. 1–20, 2007.

[3] D. J. Bernstein and T. Lange, "Analysis and optimization of elliptic-curve single-scalar multiplication," *Contemporary Mathematics*, vol. 461, no. 461, pp. 1, 2008.

[4] B. Chevallier-Mames, M. Ciet, and M. Joye, "Low-cost solutions for preventing simple side-channel analysis: Side-channel atomicity," *IEEE Transactions on computers*, vol. 53, no. 6, pp. 760–768, 2004.

[5] S. M. Cho, S. C. Seo, T. H. Kim, Y. H. Park, and S. Hong, "Extended elliptic curve Montgomery ladder algorithm over binary fields with resistance to simple power analysis," *Information Sciences*, vol. 245, pp. 304–312, 2013.

[6] L. Elmegaard-Fessel, "Efficient scalar multiplication and security against power analysis in cryptosystems based on the NIST elliptic curves over prime fields," *IACR Cryptology ePrint Archive*, vol. 2006, pp. 313, 2006.

[7] R. R. Farashahi, H. F. Wu, and C. A. Zhao, "Efficient arithmetic on elliptic curves over fields of characteristic three," in *International Conference on Selected Areas in Cryptography*, pp. 135–148, 2012.

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

[9] N. KOBLITZ, "Elliptic curve cryptosystems," *Mathematics of computation*, vol. 48, no. 177, pp. 203–209, 1987.

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

[11] L. Li, "Research on the ternary algorithm in the elliptic curve operations," *Journal of Network Safety Technology and Application (in Chinese)*, no. 11, pp. 94–96, 2015.

[12] Y. Li, J. L. Wang, X. W. Zeng, and X. Z. Ye, "A segmented Montgomery scalar multiplication algorithm with resistance to simple power analysis," *Computer Engineering and Science (in Chinese)*, vol. 39, no. 1, pp. 92–102, 2017.

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

[14] S. G. Liu, R. R. Wang, Y. Q. Li, and C. L. Zhai, "An improved ternary Montgomery ladder algorithm on

elliptic curves over GF(3ˆm)," *International Journal Network Security*, vol. 21, no. 3, pp. 384–391, 2019.

[15] S. G. Liu, Y. Y. Ding, R. Shi, and S. M. Lu, "Co_Z addition on elliptic curves over finite fields GF(2ˆm)," *Journal of Wuhan University (in Chinese)*, vol. 65, no. 2, pp. 207–212, 2019.

[16] J. López and R. Dahab, "Fast multiplication on elliptic curves over GF(2ˆ m) without precomputation," in *International Workshop on Cryptographic Hardware and Embedded Systems*, pp. 316–327, 1999.

[17] V. S. Miller, "Use of elliptic curves in cryptography," in *Conference on the Theory and Application of Cryptographic Techniques*, pp. 417–426, 1985.

[18] P. L. Montgomery, "Speeding the Pollard and elliptic curve methods of factorization," *Mathematics of Computation*, vol. 48, no. 177, pp. 243–264, 1987.

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

[20] T. Oliveira, J. López, and F. Rodríguez-Henríquez, "The Montgomery ladder on binary elliptic curves," *Journal of Cryptographic Engineering*, vol. 8, no. 3, pp. 241–258, 2018.

[21] S. F. Tzeng and 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, 2004.

[22] H. Wang, B. Li, and W. Yu, "Montgomery algorithm on elliptic curves over finite fields of character three," *Journal on Communications (in Chinese)*, vol. 29, no. 10, pp. 25–29, 2008.

[23] J. Wei, X. Liu, H. Liu, and W. Guo, "A low-time-complexity and secure dual-field scalar multiplication based on co-z protected NAF," *IEICE Electronics Express*, vol. 11, no. 11, pp. 20140361–20140361, 2014.

[24] X. S. Yan, X. G. Zhang, H. F. Zhang, and L. Liu, "Countermeasures in CPU for timing and power side channel attack," *DEStech Transactions on Engineering and Technology Research*, 2018. DOI: 10.12783/dtetr/pmsms2018/24880.

[25] J. You, Q. Zhang, C. D'Alves, B. O'Farrell, and C. K. Anand, "Using z14 fused-multiply-add instructions to accelerate elliptic curve cryptography," in *Proceedings of the 29th Annual International Conference on Computer Science and Software Engineering*, pp. 284–291, 2019.

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

[27] W. Yu, S. A. Musa, and B. Li, "Double-base chains for scalar multiplications on elliptic curves," in *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pp. 538–565, 2020.

[28] N. Zhang and Q. Q. Peiand G. Z. Xiao, "Elliptic curve scalar multiplication with x-coordinate," *Wuhan University Journal of Natural Sciences*, vol. 12, no. 1, pp. 163–166, 2007.

Biography

Shuang-Gen Liu, born in 1979. He received the PH.D. in cryptography form Xidian University in 2008. He is currently an associate professor with the school of cyber security, Xian University of Posts and Telecommunications, Xi?an, China. He is a member of the China Computer Federation, and a member of the Chinese Association for Cryptologic Research. His recent research interests include crptography and information security.

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

Lan Wei, born in 1998. An undergraduate student of Xi'an University of posts and telecommunications.She is mainly engaged in the research of elliptic curve cryptosystem.

A Blockchain-based Revocable Certificateless Signature Scheme for IoT Device

Yushuang Chen[1,2,3], Dong Zheng[1,3,4], Rui Guo[1,2,3,4], Yinghui Zhang[1,3,4], and Xiaoling Tao[4]
(Corresponding author: Yushuang Chen)

School of Cyberspace Security, Xi'an University of Posts and Telecommunications[1]
710121, Xi'an, China
State Key Laboratory of Integrated Services Networks, Xidian University[2]
710071, Xi'an, China
National Engineering Laboratory for Wireless Security, Xi'an University of Posts and Telecommunications[3]
Guangxi Cooperative Innovation Center of Cloud Computing and Big Data[4]
Guilin University of Electronic Technology, 541004, Guilin, China
Email: chenyushuang16@163.com

(Received Aug. 2, 2020; Revised and Accepted June 2, 2021; First Online Oct. 19, 2021)

Abstract

With the rapid development of Internet of Things (IoT) technology, completing a transaction with intelligent devices has been an essential communication mode in our daily life. Generally, lightweight smart devices conduct transactions relying on the third-party payment platform, which takes several inherent challenges of mutual distrust, data authentication and low efficiency in the IoT environment. To mitigate these issues, in this paper, we design a paring-free revocable certificateless signature (CLS) scheme with high efficiency based on the blockchain technique. Our scheme is a viable approach to enhance the performance without bilinear parings operation in the IoT lightweight equipment. And there are two different address lists stored in the blockchain to record the secret and public transaction data, respectively. According to this method, not only can the key management of the user be implemented in an efficient solution, but also enables the public verification to be performed transparently between the two parties. Furthermore, with the help of cloud server, it generates a time key for the user to accomplish revocation with low computation cost. We also make an analysis on the security to demonstrate that our proposal is existential unforgeability against adaptive chosen message attacks (EUF-CMA) under the elliptic curve discrete logarithm (ECDL) problem. Finally, the performance comparison indicates that our protocol is better than other related schemes.

Keywords: Blockchain; Certificateless Signature; Cloud Server; Paring-Free; Revocation

1 Introduction

The emergence of Internet of Things (IoT) allows lightweight computing devices to connect to each other via the wireless network, which brings an intelligent and convenient life for humans. However, opportunities often come with challenges. Indeed, one of the considerable challenges is the lack of necessary security technologies for data security and user privacy [27]. In IoT communication, malicious users may intercept wireless communication signals to eavesdrop, delete, edit, and replay messages. Therefore, it is important to ensure the data authenticity of the messages. Also, the user privacy preserving issue in IoT is crucial. The personal information should be retained confidential and not disclosed. Recently, with the exponential growth of IoT devices, providing data security and privacy protection as an important research direction has attracted widespread attention from academia and industry. However, lightweight equipment deploys distributed topology for information interaction in the IoT environment, which involves the interest game between the user and the device, and it affects the user privacy and data security in the absence of a trusted mechanism [17]. Meanwhile, it leads to high cost of establishing credits and low efficiency of performing communication between IoT devices.

There are many measures taken to satisfy these requirements in the IoT communication. It is necessary to design cryptographic protocols for strengthening the mutual trust between IoT equipment [26]. The relevant protocols include utilizing the digital signature to guarantee the data authenticity and designing the pseudonym mechanism to achieve user privacy preservation. Also, the public key infrastructure (PKI) plays a crucial role in establishing mutual trust, and it is regarded as a secure platform for IoT devices collaboration. Nevertheless, in the PKI systems, the trusted certificate authority (CA) is responsible for issuing certificates to users, suffering from the inherent certificate management weakness. The tasks of certificate management are costly and burdensome, involving revocation, storage, distribution and verification, *etc.* To minimize the cost of certificate management, Shamir [19] first introduced the concept of identity-based signature (IBS) scheme in 1984. The IBS scheme is built on a private key generator (PKG), which calculates a corresponding private key based on the user's identity. Unfortunately, in this signature scheme, it considers the user's relevant identity information as the public key in communication and sustains the key escrow drawback. To solve the problem, Al-Riyami and Paterson [1] proposed the idea of certificateless public key cryptosystem (CL-PKC). As a kind of CL-PKC, certificateless signature (CLS) scheme offers a solution for the IoT equipment to check the authenticity of the communication data. The CLS scheme is different from the IBS scheme and the user sends the identity information to a key generation center (KGC), then KGC produces a corresponding partial private key. Afterwards, the user fetches a full private key for communication with the partial private key. CLS scheme does not only require the user's public key for authentication, but also the KGC cannot obtain the user's complete private key. Therefore, it is free from the certificate management and key escrow in CLS schemes.

Although IoT communications benefit from the advantage of lightweight key management by using the CLS algorithm, KGC undertakes the work of issuing keys, which will lead to a heavy computational burden for the user revocation in a revocable CLS system. Obviously, this does not apply to the distributed IoT environment. The blockchain technique and cloud computing platforms can ensure the secure and efficient services for IoT devices communication. The cloud computing technology is utilized to provides outsourcing services for revocation, which reduces calculating costs and improves the revocation efficiency. However, during the outsourcing in the cloud server, the key issues on the security are the integrity of the outsourced data and the trust of the users, which should adopt some methods to avoid the malicious adversary to distorted the stored data in the cloud. The problems can be solved with the help of the blockchain technique. In recent years, the blockchain technology is developed as a decentralized ledger to support services in many areas [15]. Especially, depending on the properties of blockchain technique, such as decentralization, transparency, immutability and so on, it can be utilized in the efficient and secure distributed IoT environment. It links all the blocks together to track and coordinate transactions, and has the ability to record the information for billions of IoT equipment in the network. In the distributed and de-trusted form, blockchain technology has become an important cornerstone for solving decentralization and building trust [17]. A paring-free revocable CLS scheme built on the blockchain-based IoT communication not only ensures the data security, but also lowers the operation costs. And it provides high-performance blockchain-cloud services employing the cloud computing for the communication between the IoT devices.

1.1 Our Contributions

Based on the blockchain technique, this paper designs a novel data transaction system using a revocable paring-free certificateless signature scheme applied to the IoT devices communication. Specifically, our contributions are as follows:

1) We present a revocable certificateless signature scheme without bilinear parings, which enhances the efficiency of signing and verification and reduces the computational cost on lightweight devices.

2) Our scheme is based on cloud computing to generate a time key achieving the revocation of the users, which improves the revocation efficiency. Meanwhile, the time cost of the transaction on the blockchain is also advanced by outsourcing to the cloud server.

3) There are two lists designed in the proposed scheme. One is the private address list that is utilized to record the private information, including the real identity and private key of the user, is kept confidentially by the user nodes on the blockchain. The other list denotes the public information, including verification key and time key, is broadcast on the blockchain. Thus, our scheme can not only protect the privacy of users, but also realize the transaction verification in public with the properties of distributed storage and transparency on blockchain.

4) The proposed scheme resists Type I, Type II, Type III and Type IV adversaries under the hardness of elliptic curve discrete logarithm (ECDL) problem. Besides, our protocol enjoys the anonymity of user in the revocation, and the cloud server knows nothing about the user's real identity when issuing a time key.

1.2 Realted Work

In 2003, Al-Riyami and Paterson [1] first presented a CLS scheme without suffering from the key escrow flaw. However, Huang *et al.* [11] soon indicated that the scheme [1] was weak in the key replacement attacks initiated by any malicious third party, then they proved that the improved solution was secure in the random oracle model. In 2004, Yum and Lee [28] designed a CLS construction from an identity-based signature scheme. Unfortunately, Hu *et al.* [10] pointed out that the proposed structure was insecure against the public key replacement attacks, then they developed a modification of the construction. Afterwards, the first concrete CLS scheme in the standard model was constructed in literature [16]. However, a malicious but passive adversary was not considered in the security analysis. Very recently, several proposals in the standard model were designed. In the literatures [2, 21], although the authors demonstrated the CLS schemes were secure in the standard model, the computational efficiency was low.

The majority of the proposals mentioned above remained the bilinear pairing, which greatly impaired the efficiency in the IoT environments. He *et al.* [9] proposed the first effective CLS scheme without bilinear pairings and provided the security proof in the random oracle model. It was efficient to lower the computation cost both in the sign and verify algorithms. Unfortunately, Tsai *et al.* [24] pointed out the scheme [9] contained some intrinsic flaws and was fail to resist the attack of Type II adversary, then they developed a corresponding improved CLS scheme. Recently, many other CLS schemes without bilinear pairings were put forward to enhance the proposed system performance [8, 20, 25]. However, Karati *et al.* [12] showed that most existing CLS schemes were based on map-to-point hash functions, which led to great challenges of the computing capacity and system expansion performance in the practical applications. The proposal in the literature [12] avoided the map-to-point hash function and the formal security proof was defined in the random oracle model. However, Zhang *et al.* [29] demonstrated the scheme [12] still suffered from the security defects.

Besides the system performance, the revocation of a user is also a considerable issue in the CLS scheme. The original revocation technique was that KGC periodically updated some private keys for users. However, the new key must be transmitted to the user through a secret channel. In the literature [1], the authors introduced an initial revocation solution. The KGC updated private keys for users who was not revoked, which inevitably needed more calculation. The scheme relied on the secure channel and could not ensure the security and efficiency requirements for IoT lightweight devices. In 2016, Sun *et al.* [22] provided an effective revocable CLS scheme without bilinear pairings. Unfortunately, the attack modes in the literature [3] indicated that the scheme [22] was not secure in resisting Type I and Type II adversaries. The design in the literature [5] first suggested to apply cloud computing to the CLS scheme, the authors promoted a revocable certificateless signature scheme. It concluded that cloud services played a critical role in lightweight applications.

In the past few years, with the improvement of CLS schemes, many research efforts had paid more attention to apply CLS algorithms to various IoT application environments thanks to the characteristics of certificate-free and escrow-free. In 2015, Tsai *et al.* [23] designed a certificateless short signature scheme. The proposal was secure against the attacks of Type I and Type II adversaries, and could be applied to the lightweight devices which had limited storage capacity. A paring-free CLS scheme [7] was introduced for healthcare wireless medical senor networks, and the authors also presented an aggregation technique to reduce the computational complexity. Other schemes in the literatures [6, 13], the secure protocols were deployed under the various IoT environments.

As we know, IoT equipment communication involves the cooperation between multiple peer-to-peer entities. Besides, there are still intermediate operators, which lead to the limitations of high operating cost, low execution efficiency, and poor expansion performance for the system. Blockchain, as a cryptocurrency mechanism, is available for the operating costs, security and privacy issues with the features of distributed and tamper-proof. Moreover, it supports massive devices expansion in IoT. In 2016, Christidis *et al.* [4] discussed the role of the blockchain and showed several applications in the IoT environment. Nevertheless, the authors did not provide a detailed solution. The most recent work [14], a distributed data storage scheme based on the blockchain and certificateless cryptography was put forward. It gave an effective protocol for the authentication of IoT equipment. The design indicated that it was feasible for applying certificateless cryptosystem to blockchain technology.

Above all, our proposed protocol benefits from the paring-free revocable CLS algorithm with blockchain. And it does not rely on a secure channel for the communication data. In addition, the scheme overcomes the heavy computational burden with cloud computing for IoT devices.

1.3 Organization

The remainder of this paper is organized as follows: Section 2 introduces the background with the computational complexity assumptions and the basic technology in our scheme. Section 3 describes the system model and the security model of our proposed system. Section 4 details the presented revocable certificateless signature scheme. Then we analyse the security and performance in Sections 5 and 6, respectively. Finally, Section 7 draws the conclusion of this paper.

2 Preliminaries

2.1 Computational Assumption and Mathematical Definition

Definition 1. *Elliptic Curve Discrete Logarithm (ECDL) problem*
Given an elliptic curve additive group G with prime order q, and P is a generator of G. The ECDL problem defines that there is a point $X \in G$, where $X = aP$, given $a \in Z_q^*$, then it is computationally infeasible for every probabilistic polynomial time (PPT) algorithm B to compute a with non-negligible probability. The advantage ε of B in solving the solution of the ECDL problem is considered as

$$|\Pr[P, X \in G; a \leftarrow B\ (P, X) : X = aP]| \geq \varepsilon$$

Definition 2. *Cryptographic hash function [12] denotes that it is computationally infeasible for every PPT algorithm B to solve x from a given hashed value $H(x)$. The* *advantage ε of B in working out another solution x' is formulated as*

$$\left|\Pr\left[\begin{array}{c} x \in_R \{0,1\} \\ z \leftarrow H(x) \end{array}\middle| x' \leftarrow B(z), H(x') = z\right]\right| \geq \varepsilon$$

2.2 Blockchain

Blockchain was first described in a paper on bitcoin written by Satoshi Nakamoto in 2008. Essentially, the blockchain technique is a distributed ledger with decentralization, tamper resistance and traceability. And the transaction data between two nodes is recorded in the blockchain. In detail, a blockchain structure is shown in Figure 1. It shows the elements of a block, and the relation between blocks. One block is composed of two parts: The block header and block body. The block header involves a nonce, a timestamp, the hash value of the previous block, and the root hash. Moreover, the block body mainly records the detailed transaction information and transaction number of nodes. Specifically, two nodes on the blockchain conduct transactions with the consensus mechanism (e.g. PoW) and smart contracts. The transaction process between two nodes is shown in Figure 2. Assume that a user A needs to transfer *wallet* for a merchant server B by the transaction T_X, we define the current transaction $T_X(i) = \{body_i = (H_i, data_i), \sigma_i, t_i\}$, where H_i denotes the hash value of the previous transaction, σ_i is a signature of payer A, $data_i$ represents some information that stores in the block body $body_i$, t_i is a valid timestamp in transaction. After inputting the signature σ_i, the merchant server B verifies the valid of σ_i and outputs the verification results.

The implementation of blockchain can be applied to the domains such as IoT, cloud computing, and big data, *etc.* and blockchain technique has some good characteristics that satisfies the requirements of security, storage of data when the entities communicate with each other. The following describes some of the necessary features:

Decentralization: Record data as a distributed ledger and support interoperability;

Immutable: The data recorded on the blockchain is tamper-resistant;

Authentication: The nodes that are deployed and approved on the blockchain can access the blockchain network.

3 System Model and Security Model

3.1 The System Model

As shown in Figure 3, the system model of the presented certificateless signature scheme consists of four entities: The user U_i, Key Generation Center (KGC), Cloud Server (CS) and Merchant Server (MS). These entities work as the nodes to implement consensus and mutual supervision by consensus mechanism on the blockchain through the Internet. The following are the details.

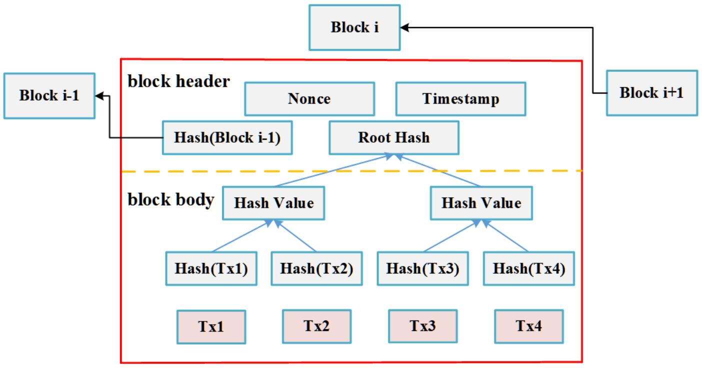


Figure 1: The blockchain structure

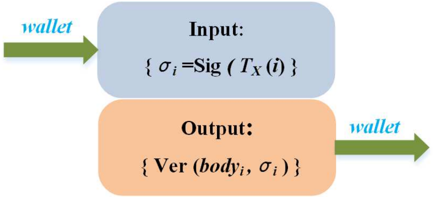

wallet
Input:
{ σ_i =Sig (T_X (i) }
Output:
{ Ver ($body_i$, σ_i) }
wallet

Figure 2: The transaction process based on blockchain

- **Key generation center:** KGC is responsible for calculating the public key and the system master key. Besides, it acts as the intermediate node and generates partial-private key for the user U_i.
- **User:** As a payer, U_i communicates with MS on the blockchain, he first calculates the full private key and public key after receiving the partial-private key from KGC. Then he signs the transaction message using the private key and sends the wallet information with signature to MS.
- **Merchant server:** MS establishes a transaction connection with U_i and performs a verification to confirm the validity of the received transaction information. Once the deal is done, MS broadcasts the transaction status into the blockchain.
- **Cloud server:** CS provides cloud computing services. It has a massive amount of computing power and generates a time key for U_i when receiving the U_i's revocation request.

3.2 Security Model

In this subsection, we introduce the security model. There are four different types adversaries in the proposed scheme, namely Type I adversary $\mathcal{ADV}_1$, Type II adversary $\mathcal{ADV}_2$, Type III adversary $\mathcal{ADV}_3$, and Type IV adversary $\mathcal{ADV}_4$. The details as below.

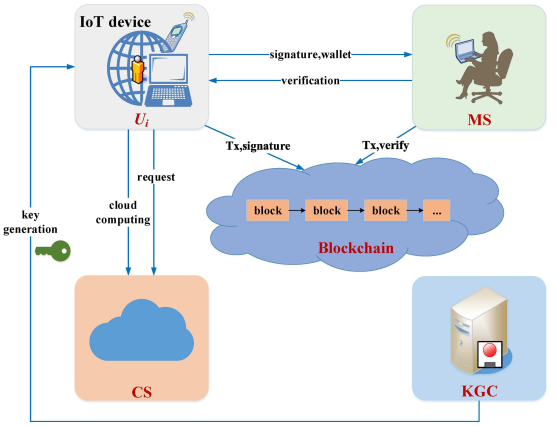


Figure 3: The system model

1) Adversary $\mathcal{ADV}_1$ is regarded as a dishonest user, and it does not know the system master key, but has the ability to replace the user's public key with choosing a random value.

2) Adversary $\mathcal{ADV}_2$ acts as a malicious-but-passive KGC. It can hold the system master key and the user's partial private key. $\mathcal{ADV}_2$ can neither replace the public key nor extract the secret value of user.

3) Adversary $\mathcal{ADV}_3$ simulates as a curious CS. We suppose that $\mathcal{ADV}_3$ knows the time master key, partial private keys and the secret values of other users besides the target user.

4) Adversary $\mathcal{ADV}_4$ models as a malicious revoked user who can extract the partial private key and the secret value of with a legal ID of the user, but cannot obtain the user's time key.

Next, the security is defined with the following games between a challenger and an adversary.

Game 1: Let $\mathcal{ADV}_1$ be a dishonest user.

Setup. A challenger $\mathcal{C}$ performs the Setup algorithm with inputting a secure parameter k, and obtains the system master key s and the public parameters $params$. Then $\mathcal{C}$ keeps s secretly and sends the $params$ to $\mathcal{ADV}_1$.

Queries. In this phase, the adversary $\mathcal{ADV}_1$ can adaptively ask the following oracle queries.

***Establish-Identity*:** If $\mathcal{ADV}_1$ does an Establish-Identity query with a valid identity ID_i, $\mathcal{C}$ first executes the Extract-Partial-Private-Key algorithm to fetch the user's partial private key, then performs the Set-Secret-Value algorithm to generate the secret value x_i. In addition, $\mathcal{C}$ outputs the user's public key pk_i and CS's private key sk_c and public key pk_c by running the Generate Key algorithm. After that, $\mathcal{C}$ creates a list L^{list} and stores $\langle ID_i, sk_i = (x_i, R_i, D_i), pk_i, sk_c, pk_c\rangle$ into it, and returns pk_i to $\mathcal{ADV}_1$.

***Extract-Partial-Private-Key*:** Upon receiving the query with an identity ID_i, $\mathcal{C}$ performs the Extract-Partial-Private-Key algorithm, then outputs the partial private key to $\mathcal{ADV}_1$.

***Set-Secret-Value*:** $\mathcal{ADV}_1$ can make such a query with the user's identity ID_i, then $\mathcal{C}$ returns the secret value x_i to $\mathcal{ADV}_1$.

***Replace-Public-Key*:** Suppose that $\mathcal{ADV}_1$ has the ability to replace the user's public key pk_i with a new random value pk_i'.

***Update-Time-Key*:** $\mathcal{ADV}_1$ can ask the Update-Time-Key query with taking an identity ID_i and a valid time stamp t_i as input, then $\mathcal{C}$ executes the Update-Time-Key phase and returns a time key tk_i to $\mathcal{ADV}_1$.

***Sign*:** In this query, $\mathcal{ADV}_1$ has the ability to request a signature σ with inputting $\langle ID_i, m, pk_i, t_i\rangle$, and then the challenger $\mathcal{C}$ outputs σ to $\mathcal{ADV}_1$ by running the Sign algorithm.

Forgery. Finally, $\mathcal{ADV}_1$ produces a tuple $\langle\sigma', ID', m', pk_i', t_i\rangle$ and wins Game 1 once the following conditions are met:

- Always outputs "valid" when verifies the given $\langle\sigma', ID', m', pk_i', t_i\rangle$.
- The tuple $\langle ID', m', t_i\rangle$ has never been queried to the Sign oracle.
- ID' has never been queried to the Extract-Partial-Private-Key oracle.

Definition 3. *The proposed scheme is secure against Type I adversary if there is not a PPT adversary $\mathcal{ADV}_1$ who wins the Game 1 with non-negligible advantage ε.*

Game 2: Let $\mathcal{ADV}_2$ be a malicious but massive KGC.

Setup. The challenger $\mathcal{C}$ executes the Setup algorithm in an identical way as that in Game 1. Besides, $\mathcal{C}$ returns the master key s to the adversary $\mathcal{ADV}_2$.

Queries. In this phase, $\mathcal{ADV}_2$ adaptively asks the following oracle queries with the similar way as that in Game 1: *Establish-Identity, Set-Secret-Value, Update-Time-Key* and *Sign Queries.*

Forgery. Finally, $\mathcal{ADV}_2$ generates a tuple $\langle\sigma', ID', m', pk_i', t_i'\rangle$ and wins Game 2 once the following conditions are met:

- Always outputs "valid" when verifies the given tuple $\langle\sigma', ID', m', pk_i', t_i'\rangle$.
- The tuple $\langle ID', m', t_i'\rangle$ has never been queried to the Sign oracle.
- ID' has never been queried to the Set-Secret-Value oracle.

Definition 4. *The proposed scheme is secure against the Type II adversary if there is not a PPT adversary $\mathcal{ADV}_2$ who can win Game 2 with non-negligible advantage ε.*

Game 3: Let $\mathcal{ADV}_3$ be a curious CS.

Setup. $\mathcal{C}$ performs the setup phase in the similar way as that in Game 1. After that, $\mathcal{C}$ outputs the system master key s to $\mathcal{ADV}_3$.

Queries. In this phase, assume that the adversary $\mathcal{ADV}_3$ can adaptively ask the following oracle queries in the same way as that in Game 1: *Establish-Identity, Extract-Partial-Private-Key, Set-Secret-Value* and *Sign Queries.*

Forgery. The algorithm runs in an identical way as that in Game 1.

Definition 5. *The proposed scheme is secure against the Type III adversary if there is not a PPT adversary $\mathcal{ADV}_3$ who can win Game 3 with non-negligible advantage ε.*

Game 4: Let $\mathcal{ADV}_4$ be a malicious revoked user.

Setup. A challenger $\mathcal{C}$ performs the setup algorithm in the similar process as that in Game 1.

Queries. In this phase, suppose that the adversary $\mathcal{ADV}_4$ has the ability to adaptively ask the following oracle queries in an identical way as that in Game 1: *Establish-Identity, Extract-Partial-Private-Key, Set-Secret-Value, Update-Time-Key, Replace-Public-Key* and *Sign Queries.*

Forgery. Finally, $\mathcal{ADV}_4$ returns a tuple $\langle \sigma', ID', m', pk_i', t_i' \rangle$ and wins the Game 4 once the following conditions are met:

- Always outputs "valid" when verifies the given tuple $\langle \sigma', ID', m', pk_i', t_i' \rangle$.
- The tuple $\langle ID', m', t_i' \rangle$ has never been queried to the Sign oracle.
- The tuple $\langle ID', t_i' \rangle$ has never been queried to the Update-Time-Key oracle.

Definition 6. *The proposed scheme is secure against the Type IV adversary if there is not a PPT adversary $\mathcal{ADV}_4$ who can win Game 4 with non-negligible advantage ε.*

4 The Proposed Revocable Certificateless Signature Scheme

This section details the presented scheme, as shown in Figure 4. The notations used in the scheme are defined in Table 1.

Table 1: Notations in the scheme

Notation	Meaning
U_i	The user i.
MS	The merchant server.
ID_i	An identity of U_i.
psk_i	The partial private key of ID_i.
(pk_i, sk_i)	The key pair of ID_i.
tk_i	The time key of ID_i.
vk_i	The verification key of ID_i.
t_i	The current timestamp.
t_s	The starting timestamp.
t_e	The ending timestamp.
(pk_s, sk_s)	The key pair of MS.
(x_c, pk_c)	The key pair of cloud server.

4.1 Initialization

Step 1. The entities including KGC, CS, U_i, and MS connect with each other on the blockchain network. These entities are regarded as the nodes to finish a transaction. In the initialization phase, the system initials the block at first, then runs the written smart contract to generate two initialized empty lists for U_i, called A^{list} and R^{list}, the details of them are described as below.

1) The private address list A^{list}, which records the information about U_i's privacy address. In detail, the form of the list is defined as $ID_i||H(ID_i)||sk_i$, where ID_i is a real identity of U_i, $H(ID_i)$ is the hash value of ID_i, and sk_i is the private key of U_i. Here, the list A^{list} is kept by the U_i secretly.

2) The confirmation list R^{list}, which stores the public information of U_i. The system creates a new list R^{list} without doing anything to the previous address list A^{list}. Concretely, the form of the list is described as $H(ID_i)||pk_i||vk_i||tk_i||t_s||t_e||status$, where pk_i is the public key of U_i, vk_i is the U_i's verification key, tk_i is the revoked user U_i's time key, and t_i is the current transaction request timestamp of U_i. Moreover, t_s denotes the transaction starting time, t_e logs the end time of a transaction. There is a status recorded in the blockchain, which represents a result of the transaction between two nodes. In addition, the information broadcasts in the blockchain and any nodes can fetch the list R^{list}.

Step 2. This setup algorithm is run by KGC. The following steps are executed to generate the system parameters and master key. Given a security parameter k and a large prime number q. KGC chooses an elliptic curve additive group G with order q, where P is a

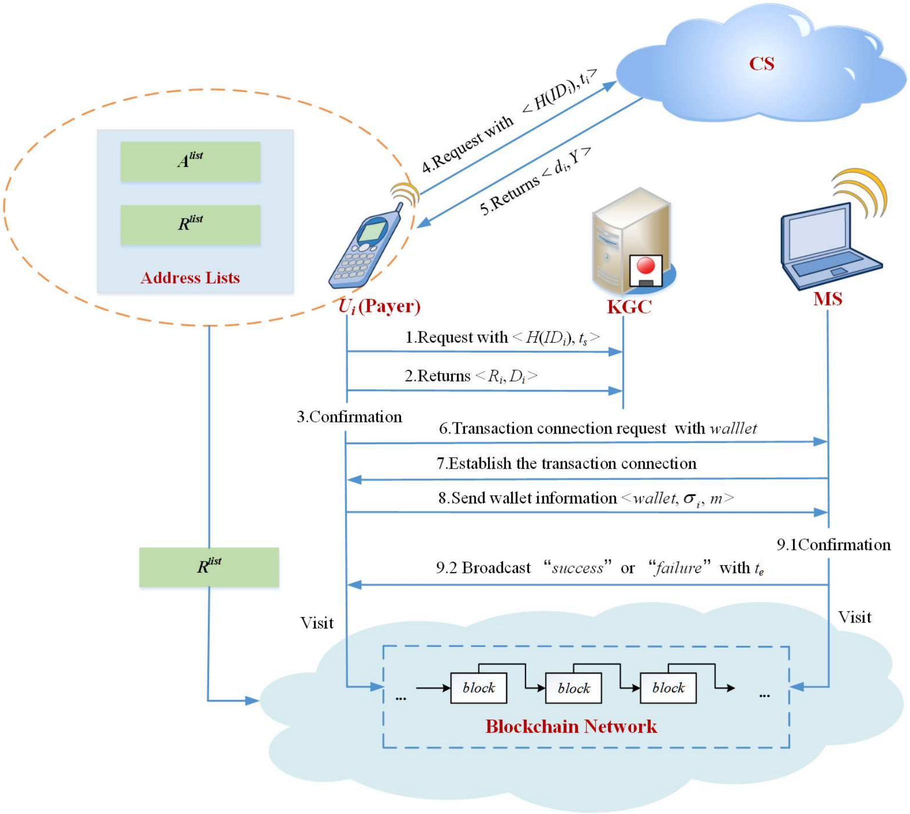

CS
4.Request with <H(ID_i), t_i>
5.Returns < d_i, Y >
A^list
R^list
Address Lists
U_i (Payer)
KGC
MS
1.Request with < H(ID_i), t_s >
2.Returns < R_i, D_i >
3.Confirmation
6.Transaction connection request with walllet
7.Establish the transaction connection
8.Send wallet information <wallet, σ_i, m>
9.1Confirmation
9.2 Broadcast "success" or "failure" with t_e
R^list
Visit
Visit
...
block
block
block
...
Blockchain Network

Figure 4: The transaction details of our proposed scheme

generator of G. Then KGC sets $s \in {}_R Z_q^*$ as the system master key and computes $P_{pub} = sP$. After that, KGC picks secure hash functions H_0, H_1, H_2, H_3 and H_4, where $H_1 : \{0,1\}^* \times G^2 \times \{0,1\}^* \to Z_q^*$, $H_2 : G \to Z_q^*$, $H_3 : \{0,1\}^* \times \{0,1\}^* \times G \times \{0,1\}^* \to Z_q^*$, and $H_4 : \{0,1\}^* \times G \times Z_q^* \to Z_q^*$. Finally, KGC keeps s secretly and publishes the system public parameters $params = \{G, k, q, P, P_{pub}, H_0, H_1, H_2, H_3, H_4\}$.

4.2 Transaction

In this subsection, U_i and MS establish a transaction connection and finish a full communication. The details proceed as below.

Step 1. Request. The User U_i starts a transaction with a request under an identity ID_i and a timestamp t_i.

Step 2. Extract partial private key. Upon receiving the request from U_i, KGC first picks a value $r_i \in {}_R Z_q^*$, and calculates $R_i = r_i P$, $h_0 = H_0 \ (ID_i, R_i, P_{pub}, t_i)$, $D_i = r_i + sh_0$, then sends the results $\langle R_i, D_i \rangle$ to U_i. When U_i fetches $\langle R_i, D_i \rangle$, he or she verifies the correctness by the equation $D_i P - R_i = h_0 P_{pub}$. If it holds, U_i determines $psk_i = \langle R_i, D_i \rangle$ as the partial private key, then U_i finishes the confirmation process.

Step 3. KeyGen. U_i selects two integers $x_i, \beta_i \in {}_R Z_q^*$, and sets x_i as the secret value, then computes $X_i = x_i P$ as the public key pk_i. To be specific, U_i issues the full private key $sk_i = \langle x_i, psk_i \rangle = \langle x_i, R_i, D_i \rangle$. Thus, U_i has a key pair $\langle sk_i, pk_i, vk_i \rangle$, where the verification key $vk_i = \beta_i$ is used for verifying. Afterwards, U_i records the transaction key pair $\langle sk_i, pk_i \rangle$ into the list A^{list} and stores the vk_i and pk_i into the list R^{list}.

Step 4. Update time key. In this phase, U_i obtains a time key tk_i with the help of cloud computing. The details proceed as follows.

1) When receiving a update request with the timestamp t_i and the hash value of an identity $H_1(ID_i)$ from U_i, CS first examines the validity of $H_1(ID_i)$ and the timestamp t_i by searching in the confirmation list R^{list}. If it matches, CS confirms that the identity of U_i is legal. Then CS chooses a value $y_t \in {}_R Z_q^*$ and computes $Y = y_t P$, $h_1 = H_1(ID_i, pk_i, Y, t_i)$, and sets $d_i = x_c + y_t h_1 - H_2 \ (x_c pk_i)$. After that, CS returns the tuple $\langle d_i, Y \rangle$ to U_i.

2) Once U_i acquires the tuple from CS, he or she first calculates $tk_i = d_i + H_2(x_i pk_c)$, then verifies the validity of tk_i by the formula $tk_i P = pk_c + h_1 Y$. If the verification holds, U_i outputs a time key tk_i.

3) U_i records the time key tk_i into the list R^{list} and broadcasts it on the blockchain network.

Step 5. Wallet request. U_i keeps the corresponding records into the two different lists, respectively. Afterwards, U_i sends a wallet request based on its verification key vk_i to MS. The details are given as below.

1) Upon receiving the wallet request from U_i, MS sets a value $x_s \in {}_R Z_q^*$ as the private key and determinesas $pk_s = x_s P$ the public key. After that, MS returns the pk_s to U_i, thus the transaction key pair of MS is $\langle x_s, pk_s \rangle$.

2) U_i computes a wallet information *wallet* after knowing the pk_s of MS, where the results is $wallet = (x_i + pk_s D_i)\beta_i^{-1}$.

3) Then MS verifies the wallet information with β_i by the formula $wallet \beta_i P - pk_i = pk_s \ (R_i + h_0 P_{pub})$. If it holds, MS accepts the *wallet* and agrees establish the transaction connection with U_i.

Step 6. Sign. After the transaction connection has established between U_i and MS, the signature generation algorithm is executed by U_i to output a signature σ_i on m, where $m \in \{0,1\}^*$. U_i does the following:

1) U_i picks an integer $\alpha \in {}_R Z_q^*$, and computes $Q = \alpha^{-1}P$, $h_3 = H_3(ID_i, m, tk_i, t_i)$, and $h_4 = H_4(ID_i, R_i, h_3)$.

2) Calculates $S = \alpha \ (h_3 x_i + h_4 D_i + tk_i)$.

3) Finally, U_i generates the signature $\sigma_i = \langle S, Q, Y, R_i \rangle$, then returns the full transaction information $\langle wallet, \sigma_i, m \rangle$ to MS.

Step 7. Verify. When MS receives the full transaction message $\langle wallet, \sigma_i, m \rangle$ from U_i, MS checks the validity of $\langle \sigma_i, m \rangle$ by the following equation:

$$QS \overset{?}{=} h_3 pk_i + h_4 \ (R_i + h_0 P_{pub}) + pk_c + h_1 Y$$

Once the verification holds, MS outputs the results "*success*" as the *status* and produces a corresponding timestamp t_e denoting the finish and accepts the transaction *wallet*, otherwise, outputs "*failure*".

4.3 Revocation

Finally, MS broadcasts the status of the transaction results ("*success*" or "*failure*") to the blockchain network, then U_i records the transaction *status* and t_e into the public list R^{list}. Thus, each nodes can fetch the recorded data. Afterwards, there are two lists called A^{list} and R^{list} to record the privacy information and public information on the blockchain, respectively. When U_i requests a revocation with a current timestamp t_i' and the hash value of an identity $H_1(ID_i)$ to CS, CS first searches the R^{list} and checks the validity of the time key by the received timestamp t_i', if it satisfies $t_i' \in \ (t_s, t_e)$, namely t_i' is beyond the scope of the recorded transaction start time t_s and the end time t_e, then the time key tk_i is legal. Otherwise, CS issues a new time key for U_i to perform the revocation algorithm. The details proceed as follows.

CS picks an integer $y_t' \in {}_R Z_q^*$ and calculates $Y' = y_t' P$, $h_1' = H_1 \ (ID_i, pk_i, Y', t_i')$, and sets $d_i' = x_c + y_t' h_1' - H_2 \ (x_c pk_i)$, then returns $\langle d_i', Y' \rangle$ to U_i. After that, U_i computes $tk_i' = d_i' + H_2 \ (x_i pk_c)$, where the tk_i' is a new time key for U_i.

5 Security Analysis

5.1 Consistency

1) Partial private key: The following Equation (1) holds when the user U_i's partial private key achieves consistency.

$$D_i P = \ (r_i + s h_0) P = r_i P + s h_0 P = R_i + h_0 P_{pub} \quad (1)$$

2) Time key: The following Equation (2) holds when the user U_i's time key achieves consistency.

$$\begin{aligned} tk_i P = &\ (d_i + H_2(x_c pk_i))P \\ = &\ (x_c + y_t h_1 - H_2 \ (x_c pk_i) + H_2(x_i pk_c))P \\ = &\ (x_c + y_t h_1 - H_2 \ (x_c x_i P) + H_2(x_i x_c P))P \quad (2) \\ = &\ (x_c + y_t h_1)P \\ = &\ pk_c + h_1 Y \end{aligned}$$

3) Signature: The following Equation (3) holds when the signature σ_i achieves consistency:

$$\begin{aligned} QS = &\ (\alpha^{-1}P) \ (\alpha \ (h_3 x_i + D_i h_4 + tk_i)) \\ = &\ P(h_3 x_i + D_i h_4 + tk_i) \\ = &\ h_3 pk_i + h_4 \ (R_i + h_0 P_{pub}) + tk_i P \quad (3) \\ = &\ h_3 pk_i + h_4 \ (R_i + h_0 P_{pub}) + pk_c + h_1 Y \end{aligned}$$

5.2 Security Proof

The security proof shows the proposed revocable certificateless signature scheme is secure under the security model, where satisfies secure under the existential unforgeability against adaptive chosen message attacks (EUF-CMA). We define four different theorems to prove the scheme satisfies unforgeability. The details are described as follows.

Theorem 1. *(Type I adversary security). If an adversary $\mathcal{ADV}_1$ can break our scheme in a polynomial time in the random oracle model with a non-negligible advantage ε after making at most q_{H_0} queries to the random oracle H_0, q_{EK} queries to the Extract-Partial-Private-Key-Extract oracle and q_S queries to the Sign oracle, then we can establish an algorithm $\mathcal{F}$ that uses $\mathcal{ADV}_1$ as a black box to solve the ECDL problem with a probability:*

$$\varepsilon' \geq \left(1 - \frac{q_{H_0} q_{EK} + q_S}{q}\right)\left(\frac{1}{q_{H_0}}\right)\varepsilon$$

Proof. Let the algorithm $\mathcal{F}$ be an effective algorithm for ECDL problem, given an ECDL random instance $(P, aP) \in G$, then the goal of $\mathcal{F}$ is to calculate a solution a. We define the adversary $\mathcal{ADV}_1$ can forge a signature σ, then $\mathcal{F}$ uses $\mathcal{ADV}_1$ as a subroutine to solve ECDL problem. Meanwhile, $\mathcal{F}$ performs as a challenger in Game 1. Now, the adversary $\mathcal{ADV}_1$ interacts with the challenger $\mathcal{F}$ by performing the following steps. □

Setup. $\mathcal{F}$ first selects a challenged identity ID_i. Moreover, $\mathcal{F}$ takes an element $s \in {}_R Z_q^*$ and computes $P_{pub} = sP$, then outputs the system public parameters $params = \{G, k, q, P, P_{pub}, H_0, H_1, H_2, H_3, H_4\}$ to $\mathcal{ADV}_1$.

Queries. In this phase, the challenger $\mathcal{F}$ creates the lists L^{list}, H_0^{list}, L_{sv}^{list}, L_{ppk}^{list}, L_{utk}^{list} which are initially empty. When adversary $\mathcal{ADV}_1$ does queries, $\mathcal{F}$ returns the corresponding response results.

Establish-Identity Query**:** When $\mathcal{ADV}_1$ does such a query with an identity ID_i, the challenger $\mathcal{F}$ first checks if the tuple $\langle ID_i, x_i, R_i, D_i, pk_i, t_i, h_0\rangle$ is in the list L^{list}. If so, $\mathcal{F}$ returns pk_i to $\mathcal{ADV}_1$ directly. Otherwise, $\mathcal{F}$ performs as below.

- If $ID_i \neq ID_i'$, then $\mathcal{F}$ chooses two values $r_i, h_0 \in {}_R Z_q^*$, computes $R_i = r_i P$ and sets $H_0\ (ID_i, R_i, P_{pub}, t_i) = h_0$, then determines $D_i = r_i + sh_0$.
- Otherwise, $\mathcal{F}$ picks a value $h_0 \in {}_R Z_q^*$, sets $R_i = aP$, $H_0\ (ID_i, R_i, P_{pub}, t_i) = h_0$, then determines $D_i = \perp$.

After that, $\mathcal{F}$ sends pk_i to $\mathcal{ADV}_1$ and stores the tuple $\langle ID_i, sk_i = (x_i, R_i, D_i), pk_i, t_i, h_0\rangle$ into the list L^{list}.

H_0 ***Query*** **:** Upon receiving $\langle ID_i, R_i, P_{pub}, t_i, h_0\rangle$ from $\mathcal{ADV}_1$, $\mathcal{F}$ first checks if the tuple $\langle ID_i, R_i, P_{pub}, t_i, h_0\rangle$ is in the list H_0^{list}. If so, $\mathcal{F}$ returns h_0 to $\mathcal{ADV}_1$ directly. Otherwise, $\mathcal{F}$ randomly chooses a value $h_0 \in {}_R Z_q^*$ and adds the tuple to the list H_0^{list}. Then it outputs h_0 to $\mathcal{ADV}_1$.

Set-Secret-Value Query **:** When the challenger $\mathcal{F}$ receives such a query with identity ID_i, $\mathcal{F}$ responds as follows.

- If $ID_i \neq ID_i'$, $\mathcal{F}$ picks $x_i \in {}_R Z_q^*$ and returns x_i to $\mathcal{ADV}_1$. Then $\mathcal{F}$ stores x_i into the list L_{sv}^{list}.
- Otherwise, $\mathcal{F}$ aborts.

Extract-Partial-Private-Key Query **:** $\mathcal{ADV}_1$ does this query with a created identity ID_i, $\mathcal{F}$ first checks if the tuple $\langle ID_i, R_i, D_i, t_i\rangle$ in the list L_{ppk}^{list}. If so, $\mathcal{F}$ returns the corresponding $\langle R_i, D_i\rangle$ to $\mathcal{ADV}_1$. Otherwise, $\mathcal{F}$ performs as following.

- If $ID_i \neq ID_i'$, then the challenger $\mathcal{F}$ recovers the tuple $\langle ID_i, sk_i = (x_i, R_i, D_i), pk_i, t_i, h_0\rangle$ from L_{ppk}^{list} and outputs $\langle R_i, D_i\rangle$ to $\mathcal{ADV}_1$.
- Otherwise, $\mathcal{F}$ aborts.

KeyGen Query **:** $\mathcal{F}$ maintains a list $L^{list} = \langle ID_i, sk_i = (x_i, R_i, D_i), pk_i, t_i, h_0\rangle$, upon receiving a KeyGen query, $\mathcal{F}$ returns the user's full private key $sk_i = (x_i, R_i, D_i)$ to $\mathcal{ADV}_1$. Moreover, $\mathcal{F}$ computes $X_i = x_i P$ and determines the user' public key $pk_i = X_i$.

Replace-Public-Key Query **:** Suppose that $\mathcal{ADV}_1$ asks a Replace-Public-Key query with the tuple (ID_i, pk'), $\mathcal{F}$ acts as below.

1) Fetches the tuple $\langle ID_i, sk_i = (x_i, R_i, D_i), pk_i, t_i, h_0\rangle$ from L^{list}.
2) Let $pk_i = pk'$, $x_i = \perp$.
3) Updates the corresponding tuple $\langle ID_i, sk_i = (x_i, R_i, D_i), pk_i, t_i, h_0\rangle$ based on the above record.

Update-Time-Key Query **:** $\mathcal{F}$ maintains a list L_{utk}^{list}, when $\mathcal{ADV}_1$ makes an Update-Time-Key query with the ID_i and a timestamp t_i, $\mathcal{F}$ executes the following algorithms.

1) Picks a value $y_t \in {}_R Z_q^*$ and calculates $Y = y_t P$.
2) Chooses $h_1 \in {}_R Z_q^*$ and sets $h_1 = H_1\ (ID_i, pk_i, Y, t_i)$.
3) Sets $d_i = x_c + y_t h_1 - H_1\ (x_c pk_i)$ and $tk_i = d_i + H_2\ (x_i pk_c)$.
4) Adds the tuple $(ID_i, Y, pk_c, t_i, tk_i, h_1)$ into the list L_{utk}^{list} and returns (tk_i, Y) to $\mathcal{ADV}_1$, where the user's time key is tk_i.

Sign Query**:** The adversary $\mathcal{ADV}_1$ does a Sign query with $\langle ID_i, m_i, t_i\rangle$, $\mathcal{F}$ proceeds the following.

- If $ID_i \neq ID_i'$, $\mathcal{F}$ runs the Sign algorithm to fetch a signature σ_i and sends it to $\mathcal{ADV}_1$.
- If $ID_i = ID_i'$, $\mathcal{F}$ picks the values $Q_i, h_3, h_4 \in {}_R Z_q^*$, and computes $Q_i S_i = h_3 pk_i + h_4\ (R_i + h_0 P_{pub}) + tk_i P$, then sets $h_3 = h_3\ (ID_i, m_i, tk_i, t_i)$, $h_4 = h_4\ (ID_i, R_i, h_3)$. Besides, $\mathcal{F}$ outputs the signature $\sigma_i = (S_i, Q_i, Y_i, R_i)$ to $\mathcal{ADV}_1$.

Forgery. Finally, $\mathcal{ADV}_1$ outputs a legal signature pairs $\langle ID, m, \sigma\rangle$ under an identity ID^*. By using the forking lemma, $\mathcal{F}$ replies $\mathcal{ADV}_1$ with the same random oracle. Here, $\mathcal{ADV}_1$ can generate the different forged signature pairs $\langle ID^*, m_i^*, \sigma_i^* = (S_i^*, Q_i^*, Y_i^*, R_i^*)\rangle$ and $\langle ID^*, m_i^*, \sigma_i' = (S_i', Q_i^*, Y_i^*, R_i^*)\rangle$. Therefore, we obtain Equation (4) and Equation (5) as follows.

$$Q_i^* S_i^* = h_3^* pk_i^* + h_4^*\ (R_i^* + h_0^* P_{pub}) + h_1^* Y_i^* + pk_c^* \quad (4)$$

$$Q_i^* S_i' = h_3^* pk_i^* + h_4'\ (R_i^* + h_0^* P_{pub}) + h_1^* Y_i^* + pk_c^* \quad (5)$$

Here, $Q_i^* = \alpha^{-1} P$, $R_i^* = aP$, $P_{pub} = sP$ and $h_4^* \neq h_4'$, the above equations are not linearly related to each other, then there is the Equation (6).

$$(S_i^* - S_i')\alpha^{-1} = (h_4^* - h_4')(a + sh_0^*) \quad (6)$$

Thus, $\mathcal{F}$ can compute the value of a by the following Equation (7).

$$a = \frac{S_i^* - S_i'}{\alpha\ (h_4^* - h_4')} - sh_0^* \tag{7}$$

where a is the solution of the challenging ECDL instance.

Advantage: Let's analyze whether $\mathcal{F}$ can solve the ECDL problem with the probability not less than ε'. $\mathcal{F}$ succeeds in the game when the following events E_1 and E_2 do not happen, but E_3 happens.

- E_1: The adversary $\mathcal{ADV}_1$ asks an Extract-Partial-Private-Key query under an identity ID^*.
- E_2: $\mathcal{ADV}_1$ returns a message-signature pair (m_i^*, σ_i^*) for an identity $ID^* \neq ID_i$ in the Forgery phase.
- E_3: σ_i^* is a valid forged signature on (ID^*, m_i^*).

Then, the successful probability that $\mathcal{F}$ breaches the ECDL instance in Game 1 is defined as below:

$$\begin{aligned}&\Pr[\neg E_1 \wedge \neg E_2 \wedge E_3]\\&= Pr[\neg E_1] \cdot Pr[\neg E_2 | \neg E_1] \cdot \Pr[E_3 | \neg E_1 \wedge \neg E_2]\end{aligned}$$

Since,

$$\begin{aligned}&\Pr[\neg E_1] \geq \left(1 - \frac{q_{H_0}}{q}\right)^{q_{EK}} \cdot \left(1 - \frac{1}{q}\right)^{q_S}\\&\Pr[\neg E_2 | \neg E_1] \geq \frac{1}{q_{H_0}}\\&\Pr[E_3 | \neg E_1 \wedge \neg E_2] = \varepsilon\end{aligned}$$

Thus, $\mathcal{F}$ has the overall advantage of breaking ECDL problem:

$$\begin{aligned}\varepsilon' =\ & \Pr[\neg E_1] \cdot Pr[\neg E_2 | \neg E_1] \cdot \Pr[E_3 | \neg E_1 \wedge \neg E_2]\\ \geq\ & \left(1 - \frac{q_{H_0}}{q}\right)^{q_{EK}} \cdot \left(1 - \frac{1}{q}\right)^{q_S} \varepsilon \left(\frac{1}{q_{H_0}}\right)\\ \geq\ & \left(1 - \frac{q_{H_0} q_{EK} + q_S}{q}\right)\left(\frac{1}{q_{H_0}}\right)\varepsilon\end{aligned}$$

Finally, $\mathcal{F}$ solves the ECDL problem with the probability

$$\varepsilon' \geq \left(1 - \frac{q_{H_0} q_{EK} + q_S}{q}\right)\left(\frac{1}{q_{H_0}}\right)\varepsilon$$

Theorem 2. *(Type II adversary security) If an adversary $\mathcal{ADV}_2$ can break our scheme in a polynomial time in the random oracle model with a non-negligible advantage ε after making at most q_{H_0} queries to the random oracle H_0, q_{SV} queries to the Set-Secret-Value oracle and q_S queries to the Sign oracle, then we can establish an algorithm $\mathcal{F}$ that uses $\mathcal{ADV}_2$ as a black box to solve the ECDL problem with a probability:*

$$\varepsilon' \geq \left(1 - \frac{q_{H_0} q_{SV} + q_S}{q}\right)\left(\frac{1}{q_{H_0}}\right)\varepsilon$$

Proof. Let the algorithm $\mathcal{F}$ be an effective algorithm for ECDL problem, given an ECDL random instance $(P, aP) \in G$, then the goal of $\mathcal{F}$ is to calculate a solution a. We define the adversary $\mathcal{ADV}_2$ can forge a signature σ, then $\mathcal{F}$ uses $\mathcal{ADV}_2$ as a subroutine to solve ECDL problem. Meanwhile, $\mathcal{F}$ performs as a challenger in Game 2. Now, the adversary $\mathcal{ADV}_2$ interacts with the challenger $\mathcal{F}$ by performing the following steps. □

Setup. $\mathcal{F}$ selects a challenged identity ID_i. Then, $\mathcal{F}$ takes a value $s \in {}_R Z_q^*$ as the system master key and calculates $P_{pub} = sP$, sets $params = \{G, k, q, P, P_{pub}, H_0, H_1, H_2, H_3, H_4\}$ to $\mathcal{ADV}_2$, then outputs the system parameters $params$ and s to $\mathcal{ADV}_2$.

Queries. In this phase, the challenger $\mathcal{F}$ creates the lists L^{list}, H_0^{list}, L_{sv}^{list}, L_{ppk}^{list}, L_{utk}^{list} which are initially empty. When adversary $\mathcal{ADV}_2$ does queries, $\mathcal{F}$ returns the corresponding response results.

***Establish-Identity Query*:** When $\mathcal{ADV}_2$ does such a query with an identity ID_i, the challenger $\mathcal{F}$ first checks if the tuple $\langle ID_i, x_i, R_i, D_i, pk_i, t_i, h_0\rangle$ is in the list L^{list}. If so, $\mathcal{F}$ returns pk_i to $\mathcal{ADV}_2$ directly. Otherwise, $\mathcal{F}$ performs as below.

- If $ID_i \neq ID_i'$, $\mathcal{F}$ chooses a random value $x_i \in {}_R Z_q^*$, and returns $pk_i = x_i P$.
- Otherwise, sets $pk_i = aP$.

Besides, $\mathcal{F}$ sends pk_i to $\mathcal{ADV}_2$ and adds the $\langle ID_i, sk_i = (x_i, R_i, D_i), pk_i, t_i, h_0\rangle$ into L^{list}.

H_0 ***Query, Set-Secret-Value Query, Update-Time-Key Query*** **and** ***Sign Query*:** $\mathcal{ADV}_2$ makes these queries in a similar way as them in the Theorem 1.

Forgery. Finally, $\mathcal{ADV}_2$ outputs a valid signature pairs $\langle ID^*, m^*, \sigma^*\rangle$ under an identity ID^*. By using the forking lemma, $\mathcal{F}$ replies $\mathcal{ADV}_2$ with the same random oracle. Here, $\mathcal{ADV}_2$ can generate the two different forged signature pairs $\langle ID^*, m_i^*, \sigma_i^* = (S_i^*, Q_i^*, Y_i^*, R_i^*)\rangle$ and $\langle ID^*, m_i^*, \sigma_i' = (S_i', Q_i^*, Y_i^*, R_i^*)\rangle$. Therefore, we can obtain the Equation (8) and Equation (9) as below.

$$Q_i^* S_i^* = h_3^* pk_i^* + h_4^*\ (R_i^* + h_0^* P_{pub}) + h_1^* Y_i^* + pk_c^* \tag{8}$$

$$Q_i^* S_i' = h_3' pk_i^* + h_4^*\ (R_i^* + h_0^* P_{pub}) + h_1^* Y_i^* + pk_c^* \tag{9}$$

Here, $h_3^* \neq h_3'$, $Q_i^* = \alpha^{-1} P$, $pk_i^* = aP$. Thus, $\mathcal{F}$ can obtain the solution a under the ECDL problem by the Equation (10).

$$a = \frac{S_i^* - S_i'}{\alpha\ (h_3^* - h_3')} - sh_0^* \tag{10}$$

Advantage: Let's analyze whether $\mathcal{F}$ can solve the ECDL problem with the probability not less than ε'. $\mathcal{F}$ succeeds in the game when the following events E_1 and E_2 do not happen, but E_3 happens.

- E_1: $\mathcal{ADV}_2$ asks the Set-Secret-Value query under an identity ID^*.
- E_2: $\mathcal{ADV}_2$ returns a message-signature pair (m_i^*, σ_i^*) for an identity $ID^* \neq ID_i$ in the Forgery phase.
- E_3: σ_i^* is a valid forged signature on (ID^*, m_i^*).

From the simulation, We have

$$\begin{aligned}
&\Pr[\neg E_1] \geq \left(1-\frac{q_{H_0}}{q}\right)^{q_{SV}} \cdot \left(1-\frac{1}{q}\right)^{q_S}\\
&\Pr[\neg E_2|\neg E_1] \geq \frac{1}{q_{H_0}}\\
&\Pr[E_3|\neg E_1 \wedge \neg E_2] = \varepsilon
\end{aligned}$$

Thus, $\mathcal{F}$ has the overall advantage of breaking ECDL problem:

$$\begin{aligned}
\varepsilon' =\ & \Pr[\neg E_1] \cdot Pr[\neg E_2|\neg E_1] \cdot \Pr[E_3|\neg E_1 \wedge \neg E_2]\\
\geq & \left(1-\frac{q_{H_0}q_{SV}+q_S}{q}\right)\left(\frac{1}{q_{H_0}}\right)\varepsilon
\end{aligned}$$

Finally, $\mathcal{F}$ solves the ECDL problem with the probability

$$\varepsilon' \geq \left(1-\frac{q_{H_0}q_{SV}+q_S}{q}\right)\left(\frac{1}{q_{H_0}}\right)\varepsilon$$

Theorem 3. *(Type III adversary security) If an adversary $\mathcal{ADV}_3$ can break our scheme in a polynomial time in the random oracle model with a non-negligible advantage ε after making at most q_{H_0} queries to the random oracle H_0 and q_S queries to the Sign oracle, then we can establish an algorithm $\mathcal{F}$ that uses $\mathcal{ADV}_3$ as a black box to solve the ECDL problem with a probability:*

$$\varepsilon' \geq \left(1-\frac{q_S}{q}\right)\left(\frac{1}{q_{H_0}}\right)\varepsilon$$

Proof. Let the algorithm $\mathcal{F}$ be an effective algorithm for ECDL problem, given an ECDL random instance $(P, aP) \in G$, then the goal of $\mathcal{F}$ is to calculate a solution a. We define the adversary $\mathcal{ADV}_3$ can forge a signature σ, then $\mathcal{F}$ uses $\mathcal{ADV}_3$ as a subroutine to solve ECDL problem. Meanwhile, $\mathcal{F}$ performs as a challenger in Game 3. Now, the adversary $\mathcal{ADV}_3$ interacts with the challenger $\mathcal{F}$ by performing the following steps. □

Setup. In this algorithm, $\mathcal{F}$ first randomly selects a challenged identity ID_i. Moreover, $\mathcal{F}$ takes an element $s \in {}_R Z_q^*$ and computes $P_{pub} = sP$, then outputs the system public parameters $params = \{G, k, q, P, P_{pub}, H_0, H_1, H_2, H_3, H_4\}$ to $\mathcal{ADV}_3$.

Queries.

***Establish-Identity Query*:** When $\mathcal{ADV}_3$ does such a query with an identity ID_i, the challenger $\mathcal{F}$ first checks if the tuple $\langle ID_i, x_i, R_i, D_i, pk_i, t_i, h_0\rangle$ is in the list L^{list}. If so, $\mathcal{F}$ returns pk_i to $\mathcal{ADV}_3$ directly. Otherwise, $\mathcal{F}$ performs as below.

- If $ID_i \neq ID_i'$, then $\mathcal{F}$ selects $r_i, h_0 \in {}_R Z_q^*$, computes $R_i = r_iP$, $H_0\ (ID_i, R_i, P_{pub}, t_i) = h_0$, then calculates $D_i = r_i + sh_0$.
- Otherwise, $\mathcal{F}$ picks a value $h_0 \in {}_R Z_q^*$, sets $R_i = aP$, $H_0\ (ID_i, R_i, P_{pub}, t_i) = h_0$, then determines $D_i = \perp$.

After that, $\mathcal{F}$ picks $x_i \in {}_R Z_q^*$ and determines $pk_i = x_iP$ as the public key, then sends pk_i to $\mathcal{ADV}_3$. Finally, $\mathcal{F}$ stores the tuple $\langle ID_i, sk_i = (x_i, R_i, D_i), pk_i, t_i, h_0\rangle$ into the list L^{list}.

H_0 ***Query, Set-Secret-Value Query, Extract-Partial-Private-Key Query*** and ***Sign Query:*** $\mathcal{ADV}_3$ asks the queries in a similar way as them in Theorem 1.

Forgery. Finally, $\mathcal{ADV}_3$ fetches a legal signature pairs $\langle ID^*, m^*, \sigma^*\rangle$ with an identity ID^*. By using the forking lemma, $\mathcal{F}$ replies $\mathcal{ADV}_3$ with the same random oracle. Here, $\mathcal{ADV}_3$ generates two different forged signature pairs $\langle ID^*, m_i^*, \sigma_i^* = (S_i^*, Q_i^*, Y_i^*, R_i^*)\rangle$ and $\langle ID^*, m_i^*, \sigma_i' = (S_i', Q_i^*, Y_i^*, R_i^*)\rangle$.

Thus, there are the following Equation (11) and Equation (12).

$$Q_i^* S_i^* = h_3^* pk_i^* + h_4^*\ (R_i^* + h_0^* P_{pub}) + h_1^* Y_i^* + pk_c^* \tag{11}$$

$$Q_i^* S_i' = h_3^* pk_i^* + h_4'\ (R_i^* + h_0^* P_{pub}) + h_1^* Y_i^* + pk_c^* \tag{12}$$

Here, $Q_i^* = \alpha^{-1}P$, $R_i^* = aP$, $P_{pub} = sP$ and $h_4^* \neq h_4'$, thus, from the above equations, we can solve the solution a of the challenging ECDL problem, where $a = (S_i^* - S_i')/\alpha\ (h_4^* - h_4') - sh_0^*$.

Advantage: Let's analyze whether $\mathcal{F}$ can solve the ECDL problem with the probability not less than ε'. Using the similar method as we used to analyze the successful probability in Theorem 1, we can obtain that the successful probability that $\mathcal{F}$ breaches the ECDL instance in Game 3 is defined as below:

$$\begin{aligned}
\varepsilon' =\ & \Pr[\neg E_1] \cdot Pr[\neg E_2|\neg E_1] \cdot \Pr[E_3|\neg E_1 \wedge \neg E_2]\\
\geq & \left(1-\frac{q_S}{q}\right)\left(\frac{1}{q_{H_0}}\right)\varepsilon
\end{aligned}$$

Finally, $\mathcal{F}$ solves the ECDL problem with the probability

$$\varepsilon' \geq \left(1-\frac{q_S}{q}\right)\left(\frac{1}{q_{H_0}}\right)\varepsilon$$

Theorem 4. *(Type IV adversary security) If an adversary $\mathcal{ADV}_4$ can break our scheme in a polynomial time in the random oracle model with a non-negligible advantage ε after making at most q_{H_0} queries to the random oracle H_0, then we can establish an algorithm $\mathcal{F}$ that uses $\mathcal{ADV}_4$ as a black box to solve the ECDL problem with a probability:*

$$\varepsilon' \geq \left(\frac{1}{q_{H_0}}\right)\varepsilon$$

Proof. Let the algorithm $\mathcal{F}$ be an effective algorithm for ECDL problem, given an ECDL random instance $(P, aP) \in G$, then the goal of $\mathcal{F}$ is to calculate a solution a. We define the adversary $\mathcal{ADV}_4$ can forge a signature σ, then $\mathcal{F}$ uses $\mathcal{ADV}_4$ as a subroutine to solve ECDL problem. Meanwhile, $\mathcal{F}$ performs as a challenger in Game 4. Now, the adversary $\mathcal{ADV}_4$ interacts with the challenger $\mathcal{F}$ by performing the following steps. □

Setup. $\mathcal{F}$ runs this algorithm in a same way as that in Theorem 1.

Queries.

Establish-Identity Query: When $\mathcal{ADV}_4$ does such a query with an identity ID_i, the challenger $\mathcal{F}$ first checks if the tuple $\langle ID_i, x_i, R_i, D_i, pk_i, t_i, h_0\rangle$ is in the list L^{list}. If so, $\mathcal{F}$ returns pk_i to $\mathcal{ADV}_4$ directly. Otherwise, $\mathcal{F}$ chooses two values $r_i, h_0 \in {}_R Z_q^*$, computes $R_i = r_i P$, sets $H_0\ (ID_i, R_i, P_{pub}, t_i) = h_0$ and determines $D_i = r_i + sh_0$. After that,m$\mathcal{F}$ sends pk_i to $\mathcal{ADV}_4$ and stores the tuple $\langle ID_i, sk_i = \ (x_i, R_i, D_i), pk_i, t_i, h_0\rangle$ into the list L^{list}.

H_0 ***Query, Set-Secret-Value Query, Extract-Partial-Private-Key Query*** and ***Replace-Public-Key Query:*** $\mathcal{ADV}_4$ asks these queries in an identical way as them in Theorem 1.

Update-Time-Key Query**:** $\mathcal{F}$ maintains a list L_{utk}^{list}, when $\mathcal{ADV}_4$ makes an Update-Time-Key query with the ID_i and a timestamp t_i, $\mathcal{F}$ executes the following algorithms.

- If $ID_i \neq ID_i'$, $\mathcal{F}$ picks the value $y_t, h_1 \in {}_R Z_q^*$, computes $Y = y_t P$, and sets $H_1(ID_i, pk_i, Y, t_i) = h_1$, then determines $d_i = x_c + y_t h_1 - H_1\ (x_c pk_i)$ and $tk_i = d_i + H_2\ (x_i pk_c)$.
- Otherwise, $\mathcal{F}$ selects a random value $h_1 \in {}_R Z_q^*$, sets $Y = aP$, $H_1(ID_i, pk_i, Y, t_i) = h_1$ and determines $tk_i = \perp$.

Furthermore, $\mathcal{F}$ inserts the tuple $(ID_i, Y, pk_c, t_i, tk_i, h_1)$ into the list L_{utk}^{list} and returns (tk_i, Y) to $\mathcal{ADV}_4$, where the user's time key is tk_i.

Sign Query**:** $\mathcal{ADV}_4$ asks the sign query in an identical way as that in Theorem 1.

Forgery. Finally, $\mathcal{ADV}_4$ outputs a legal signature pairs $\langle ID^*, m^*, \sigma^*\rangle$ with an identity ID^*. By using the forking lemma, $\mathcal{F}$ replies $\mathcal{ADV}_4$ with the same random oracle. Here, $\mathcal{ADV}_4$ can generate two different forged signature pairs $\langle ID^*, m_i^*, \sigma_i^* \ = \ (S_i^*, Q_i^*, Y_i^*, R_i^*)\rangle$ and $\langle ID^*, m_i^*, \sigma_i' = \ (S_i', Q_i^*, Y_i^*, R_i^*)\rangle$.
Thus, we can calculate the following Equation (13).

$$\begin{cases} Q_i^* S_i^* = h_3^* pk_i^* + h_4^*\ (R_i^* + h_0^* P_{pub}) + h_1^* Y_i^* + pk_c^* \\ Q_i^* S_i' = h_3^* pk_i^* + h_4^*\ (R_i^* + h_0^* P_{pub}) + h_1' Y_i^* + pk_c^* \end{cases} \quad (13)$$

Here, $Q_i^* = \alpha^{-1}P$, $Y_i^* = aP$ and $h_1^* \neq h_1'$, thus, from the above equations, we can solve the solution a of the challenging ECDL problem, where the result $a = \ (S_i^* - S_i')/\alpha\ (h_1^* - h_1')$.

Advantage: Let's analyze whether $\mathcal{F}$ can solve the ECDL problem with the probability not less than ε. Using the similar method as we used to analyze the successful probability in Theorem 1, we can obtain that the successful probability that $\mathcal{F}$ breaches the ECDL instance in Game 4 is defined as below:

$$\varepsilon' = \Pr[E_1] \cdot \Pr[E_2|E_1] \cdot \Pr[E_3|E_1 \wedge E_2] \geq \left(\frac{1}{q_{H_0}}\right) \varepsilon$$

Finally, $\mathcal{F}$ solves the ECDL problem with the probability $\varepsilon' \geq \left(\frac{1}{q_{H_0}}\right) \varepsilon$.

From the above theorems, we can conclude that the challenger $\mathcal{F}$ works out a solution a of ECDL problem, it is contradictory to the assumption of solving the ECDL problem. Therefore, the proposed scheme has the security properties of resisting forgery attacks.

Theorem 5. *Anonymity. Our proposed scheme has the property of anonymity.*

When the user U_i performs a request of updating time key, the communication between U_i and CS is based on researching the hash value of the identity ID_i, that is $H\ (ID_i)$. Obviously, CS can only query the hash value of an identity, based on the irreversibility and collision constraint of the hash function, the user U_i's real identity ID_i cannot be directly obtained by the attacker. Therefore, at the process that the cloud server calculates a time key for U_i, even in an insecure channel, the adversary cannot break the U_i's identity from d_i. In addition, the user only broadcasts the confirmation list R^{list} for public verification the transaction information and the validity of the time key in the transaction system based on the blockchain network, thus still cannot determine the identity ID_i on the condition that the R^{list} can only query the hash value.

6 Performance Analysis

We make a performance comparison among our proposed scheme and other relevant schemes in terms of the computational cost, experimental results, and secure features. The notations for cryptographic operations used in the scheme are shown in Table 2.

Here, the evaluations were run on a personal computer that the configuration is Intel (R) Core (TM) i5-5200U CPU @2.20GHZ, 8GB RAM and Windows 10 operating system. The codes are written in Ubuntu 16.04 LTS operating system. The experiment results are calculated by using C programming languages with the popular PBC library [18]. For the paring-based schemes, we consider a super-singular elliptic curve Type A curve $y^2 = x^3 + x$ with 512 bits group and the embedding degree is set as 2,

Table 2: The notations and execution time in our scheme

T_{bp}	The time of a bilinear paring
T_e	The time of an exponentiation
T_H	The time of a map-point hash function
T_{sm}	The time of a scalar multiplication of ECC
T_a	The time of a point addition of ECC
T_{psm}	The time of a scalar multiplication
T_{pa}	The time of a point addition

Table 3: The execution time of each operation (ms)

T_{bp}	T_e	T_H	T_{sm}	T_a	T_{psm}	T_{pa}
9.852	6.468	5.696	0.442	0.002	3.368	0.009

which can reach the equal security level as 1024 bits RSA. For the ECC-based schemes, we consider the ECC group on the Koblitz elliptic curve $y^2 = x^3 + ax + b$ defined on a finite field $F_{2^{163}}$, where the value of a is 1, b is a random prime number with 163 bits. As shown in Table 3, we calculate the execution time of each operation.

Table 4: Computation cost comparison

Schemes	Total cost (ms)	Improve (%)
[2]	$4T_{bp} + 8T_e \approx 91.15$	97.5%
[21]	$5T_{sm} + 3T_{bp} \approx 46.40$	95.2%
[25]	$7T_{sm} + 5T_{add} \approx 3.10$	28.5%
[12]	$1T_{bp} + 5T_e \approx 37.757$	94.7%
[29]	$4T_{sm} + 1T_{bp} + 3T_{add} \approx 23.40$	90.5%
[23]	$3T_{sm} + 1T_{bp} + 2T_{add} \approx 20.01$	88.9%
Ours	$5T_{sm} + 4T_{add} \approx 2.21$	—

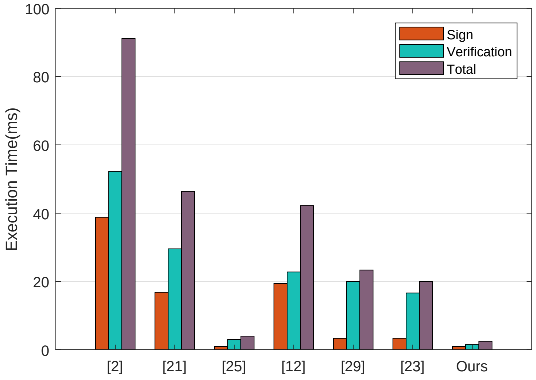


Figure 5: Computation overhead comparison

To demonstrate the computation cost comparison in the schemes, we show the performance comparison results of the total computation cost, as shown in Table 4. In the the sign phase, the computation cost requires $1T_{sm}$, the verification phase requires $4T_{sm} + 4T_{add}$, and the total process cost requires $5T_{sm} + 4T_{add} = 2.218$ms. It clearly shows that the improvement in the total computation coat of our proposal in percentage over the scheme [2] is nearly $\frac{91.152-2.218}{91.152} \times 100\% \approx 97.5\%$, similarly, over the scheme [21] is about 95.2%, over the scheme [25] is about 28.5%, over the scheme [12] is almost 94.7%, over the scheme [29] is nearly 90.5% and over the scheme [23] is nearly 88.9%. Furthermore, we give a comparison through the bar graph to indicate the execution time of various phases in the schemes, as shown in Figure 5. Specifically, in our scheme, Figure 6 and Figure 7 show the relationship between the running time and the number of computations in the verification and signature algorithm, respectively. The results indicates the relationship between running time of each phase and the number of computations is linearly increasing. In addition, we also compare the total running time of algorithm with the schemes [2, 12, 21, 23, 25, 29], as shown in Figure 8. Obviously, the schemes [2, 12, 21] use bilinear pairing operation in the verification phase, which leads to more running time each time than our presented scheme.

Therefore, as shown in the above comparison, the results demonstrate that our proposal requires less execution time than other related schemes. Also, as the number of calculations increases, it is obvious that our scheme reduces the computational overhead and is more suitable for the multiple IoT devices communication.

Besides, in Table 5, we present the properties comparison at the aspect of "Public Key Replacement", "KGC Impersonation", "Revocation" and "Security Assumption", respectively. Here, "*yes*" represents the schemes is secure against the attacks in public key replacement and KGC impersonation, "*no*" states the schemes cannot against the attacks, the symbol "×" shows that the character we define is not fulfilled and the symbol "✓" indicates that the scheme satisfies the features we define. Specifically, we can see that the studies [12] and [23] are not secure to resist the public key replacement attack and the KGC impersonation attack. The schemes [2,12,21,29] and [23] are implemented without the revocation process. Fortunately, our proposal provides the security to against the public key replacement adversary and the KGC impersonation adversary, moreover, it is also secure against the Type III adversary and Type IV adversary defined in the security model. In addition, there is the revocation process in our presented scheme.

In summary, our proposal is secure under the ECDL assumption and requires lower computation cost than the relevant schemes.

7 Conclusion

Blockchain technology plays a significant role in the IoT lightweight application communication environment. This paper designs a secure revocable CLS scheme with-

Table 5: Comparison of E-voting schemes in requirement

Sceheme	Public Key Replacement	KGC Impersonation	Revocation	Security Assumption
[2]	*yes*	*yes*	✓	CDH
[21]	*yes*	*yes*	×	CDH
[25]	*yes*	*yes*	✓	ECDL
[12]	*no*	*no*	×	BSDH&EBSDH*
[29]	*yes*	*yes*	×	SDH
[23]	*no*	*no*	×	k-CAA**
Ours	*yes*	*yes*	✓	ECDL

*BSDH: Bilinear Strong Diffie-Hellman problem; EBSDH: Extended BSDH problem

**k-CAA: Collusion attack algorithm with k traitors

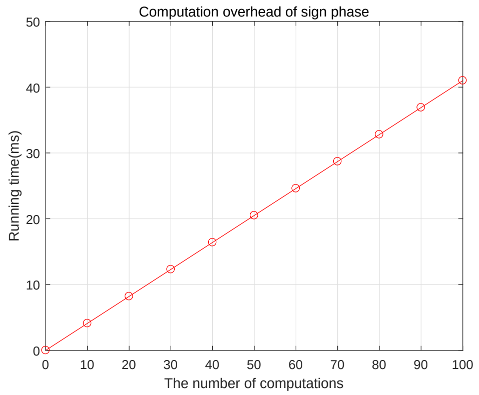


Figure 6: Computation overhead of sign phase

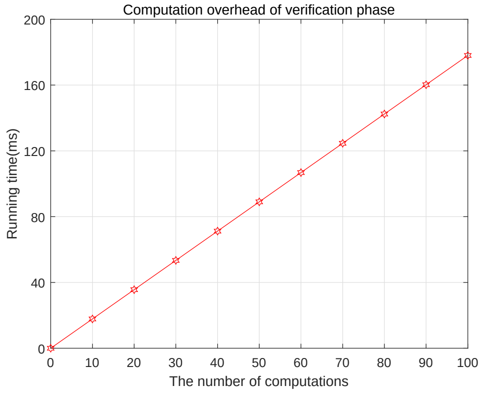


Figure 7: Computation overhead of verification phase

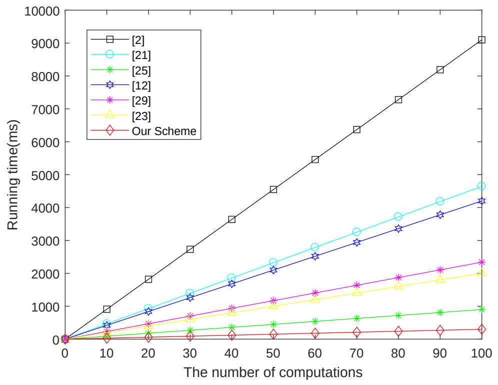


Figure 8: Comparison in computation overhead

out the bilinear pairings based on the blockchain network. On the one hand, the proposed scheme goes through a transaction without a traditional third-party platform, and we store two address lists on the blockchain to complete public verification and key management. On the other hand, the proposal efficiently updates the time key with the outsourcing cloud server during the revocation process. In conclusion, our security and performance analysis demonstrate that the presented solution is provably secure under an ECDL problem, while providing high execution efficiency in lightweight devices.

Acknowledgments

This work was supported by the National Key R&D Program of China under Grant 2017YFB0802000, the Natural Science Foundation of China under Grant 61802303, 61772418 and 61602378, the Key Research and Development Program of Shaanxi under Grant 2019KW-053, the Innovation Ability Support Program in Shaanxi Province of China under Grant 2017KJXX-47, the Natural Science Basic Research Plan in Shaanxi Province of China under Grant 2019JQ-866, 2018JZ6001 and 2016JM6033, the Research Program of Education Bureau of Shaanxi Province under Grant 19JK0803, the New Star Team of Xi'an University of Posts and Telecommunications under Grant 2016-02, the Fundamental Research Funds for the Central Universities under Grant GK201903005, and Guangxi Cooperative Innovation Center of Cloud Computing and Big Data under Grant YD1903.

References

[1] S. S. Al-Riyami and K. G. Paterson, "Certificateless public key cryptography," in *International Conference on the Theory and Application of Cryptology and Information Security*, pp. 452–473, 2003.

[2] S. Canard and V. C. Trinh, "An efficient certificateless signature scheme in the standard model," in *International Conference on Information Systems Security*, pp. 175–192, 2016.

[3] Y. C. Chen, R. Tso, W. Susilo, X. Huang, and G. Horng, "Certificateless signatures: Structural extensions of security models and new provably secure schemes," *IACR Cryptology ePrint Archive*, pp. 193, 2013. Corpus ID: 11704828.

[4] K. Christidis and M. Devetsikiotis, "Blockchains and smart contracts for the internet of things," *IEEE Access*, vol. 4, pp. 2292–2303, 2016.

[5] H. Du, Q. Wen, and S. Zhang, "A provably-secure outsourced revocable certificateless signature scheme without bilinear pairings," *IEEE Access*, vol. 6, pp. 73846–73855, 2018.

[6] H. Du, Q. Wen, S. Zhang, and M. Gao, "A new provably secure certificateless signature scheme for internet of things," *Ad Hoc Networks*, vol. 100, pp. 102074, 2020.

[7] N. B. Gayathri, G. Thumbur, P. R. Kumar, M. Z. U. Rahman, and P. V. Reddy, "Efficient and secure pairing-free certificateless aggregate signature scheme for healthcare wireless medical sensor networks," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 9064–9075, 2019.

[8] Z. Guo, "Cryptanalysis of a certificateless conditional privacy-preserving authentication scheme for wireless body area networks," *International Journal of Electronics and Information Engineering*, vol. 11, no. 1, pp. 1–8, 2019.

[9] D. He, J. Chen, and R. Zhang, "An efficient and provably-secure certificateless signature scheme without bilinear pairings," *International Journal of Communication Systems*, vol. 25, no. 11, pp. 1432–1442, 2012.

[10] B. C. Hu, D. S Wong, Z. Zhang, and X. Deng, "Key replacement attack against a generic construction of certificateless signature," in *Australasian Conference on Information Security and Privacy*, pp. 235–246, 2006.

[11] X. Huang, W. Susilo, Y. Mu, and F. Zhang, "On the security of certificateless signature schemes from asiacrypt 2003," in *International Conference on Cryptology and Network Security*, pp. 13–25, 2005.

[12] A. Karati, S. K. H. Islam, and M. Karuppiah, "Provably secure and lightweight certificateless signature scheme for iiot environments," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 8, pp. 3701–3711, 2018.

[13] J. Li, Y. Ji, K. K. R. Choo, and D. Hogrefe, "Cl-cppa: Certificateless conditional privacy-preserving authentication protocol for the internet of vehicles," *IEEE Internet of Things Journal*, vol. 6, no. 6, pp. 10332–10343, 2019.

[14] R. Li, T. Song, B. Mei, H. Li, X. Cheng, and L. Sun, "Blockchain for large-scale internet of things data storage and protection," *IEEE Transactions on Services Computing*, vol. 12, no. 5, pp. 762–771, 2018.

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

[16] J. K. Liu, M. H. Au, and W. Susilo, "Self-generated-certificate public key cryptography and certificateless signature/encryption scheme in the standard model," in *Proceedings of the 2nd ACM Symposium on Information, Computer and Communications Security*, pp. 273–283, 2007.

[17] Y. Lu, "A p2p anonymous communication scheme in iot based on blockchain," *International Journal of Network Security*, vol. 23, no. 1, pp. 49–56, 2021.

[18] B. Lynn, "PBC library–the pairing-based cryptography library," *Version 0.5*, vol. 11, 2007.

[19] A. Shamir, "Identity-based cryptosystems and signature schemes," in *Workshop on the Theory and Application of Cryptographic Techniques*, pp. 47–53, 1984.

[20] G. Sharma, S. Bala, and A. K. Verma, "An improved rsa-based certificateless signature scheme for wireless sensor networks," *International Journal of Network Security*, vol. 18, no. 1, pp. 82–89, 2016.

[21] K. A. Shim, "A new certificateless signature scheme provably secure in the standard model," *IEEE Systems Journal*, vol. 13, no. 2, pp. 1421–1430, 2018.

[22] Y. Sun, Z. Zhang, and L. Shen, "A revocable certificateless signature scheme without pairing," in *International Conference on Cloud Computing and Security*, pp. 355–364, 2016.

[23] J. L. Tsai, "A new efficient certificateless short signature scheme using bilinear pairings," *IEEE Systems Journal*, vol. 11, no. 4, pp. 2395–2402, 2015.

[24] J. L. Tsai, N. W. Lo, and T. C. Wu, "Weaknesses and improvements of an efficient certificateless signature scheme without using bilinear pairings," *International Journal of Communication Systems*, vol. 27, no. 7, pp. 1083–1090, 2014.

[25] H. Xiong, Q. Mei, and Y. Zhao, "Efficient and provably secure certificateless parallel key-insulated signature without pairing for iiot environments," *IEEE Systems Journal*, vol. 14, no. 1, pp. 310–320, 2019.

[26] Y. Yang, H. Cai, Z. Wei, H. Lu, and K. K. R. Choo, "Towards lightweight anonymous entity authentication for iot applications," in *Australasian Conference on Information Security and Privacy*, pp. 265–280, 2016.

[27] Y. Yang, L. Wu, G. Yin, L. Li, and H. Zhao, "A survey on security and privacy issues in internet-of-things," *IEEE Internet of Things Journal*, vol. 4, no. 2, pp. 1250–1258, 2017.

[28] D. H. Yum and P. J. Lee, "Generic construction of certificateless signature," in *Australasian Conference on Information Security and Privacy*, pp. 200–211, 2004.

[29] Y. Zhang, R. H. Deng, D. Zheng, J. Li, P. Wu, and J. Cao, "Efficient and robust certificateless signature for data crowdsensing in cloud-assisted industrial iot," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 9, pp. 5099–5108, 2019.

Biography

Yushuang Chen received the B.S. degree from the Xi'an University of Posts and Telecommunications in 2018. She is currently pursuing the M.S. degree with the National Engineering Laboratory for Wireless Security, Xi'an University of Posts and Telecommunications, China. Her research interests include blockchain, information security, and the Internet of Things (IoT).

Dong Zheng received the Ph.D. degree from Xidian University in 1999. He joined the School of Information Security Engineering, Shanghai JiaoTong University. He is currently a Professor with the Xi'an University of Posts and Telecommunications, China. His research interests include information theory, cryptography, and information security.

Rui Guo received the Ph.D. degree from the State Key Laboratory of Networking and Switch Technology, Beijing University of Posts and Telecommunications, China. He is an associate professor at National Engineering Laboratory for Wireless Security, Xi'an University of Posts and Telecommunications. His research interests include attribute-based cryptograph, cloud computing, and blockchain technology.

Yinghui Zhang received the Ph.D degree in Cryptography from Xidian University, China. Currently, he is a research fellow at Singapore Management University. His research interests include cloud security, public key cryptography and wireless network security.

Xiaoling Tao received the M.S. degree in computer application technology from Guilin University of Electronic Technology. She is a professor at the school of computer science and information security, Guilin University of Electronic Technology. Her research interests include cloud computing and security and network security.

Robust Training for Injection Attacks Detection in Web-based Applications

Benjamin Appiah, Zhiguang Qin, Owusu A. Kwabena, and Muhammed A. Abdullah
(Corresponding author: Benjamin Appiah)

School of Information and Software Engineering, University of Electronic Science and Technology of China
Chengdu, 610054, China
Email: 1746627510@qq.com

(Received Aug. 3, 2020; Revised and Accepted Apr. 24, 2021; First Online Oct. 19, 2021)

Abstract

Existing machine learning classifiers to detect mutants of injection attacks against web-based applications are gradually becoming ineffective. This paper proposes an algorithm that integrates a matrix learning loss term to the original neural network's objective function. The new loss serves as a regularization term that projects traffic data into a Euclidean space where distance can be directly used to measure the similarity of genuine traffic and malicious traffic and bring similar traffics close together and push dissimilar traffic far away from their false classes. Furthermore, we propose a projected gradient descent noise introduced into the training phase to produce a more generalized classifier. The experimental results on the CIC-IDS2018 dataset show that training neural network classifier with matrix learning loss term and introducing our projected gradient descent noise into the machine learning training phase proves to be more robust in detecting tiny mutants of existing injection attacks.

Keywords: Adversarial Detection; Adversarial Training; Doc2vec; Injection Attacks; Metric Learning

1 Introduction

Studies have shown that tiny mutants of existing attacks are misclassified by machine learning or neural network classifiers classifiers [3,4,6,7,9,15,18,19,24]. This issue is more challenging in Web-based applications where there is frequent and evolving attacks such as structured query language injection (SQLi) and cross-site scripting (XSS) attacks [23]. The bottom layer in Figure 1, illustrates how mutants of injection attack samples make machine learning or neural network-based intrusion detection system vulnerable.

Given a Testing Data $x_0 \in D$ with true class labels y on the left is an ordinary traffic data which could be benign or malicious (SQLi or XSS) traffics. However, the Perturbed Data on the right $(x_0^* = x_0 + \triangle x)$ is crafted by adding tiny perturbations $\triangle x$ to the original x_0 that forces a particular machine learning or neural network classifier (F) to falsely label malicious traffic into false class $(F(x^*) \neq y)$ and the amount of perturbation is maximized $(max||x^* - x||_p \ s.t \ F(x^*) \neq y)$, where $||\cdot||_p$ is the L_p ($i.e.$ L_1, L_2 or L_∞) norm defining the amount of perturbation. Researchers have shown that these transformations are effective in the physical world [8,15,17,24,30]. More specifically, considering the vulnerability of machine learning or neural network, an attacker can make twist to the inputs of most classifiers and cause Web-based intrusion detection systems to behaved abnormally by designing these adversarial traffics.

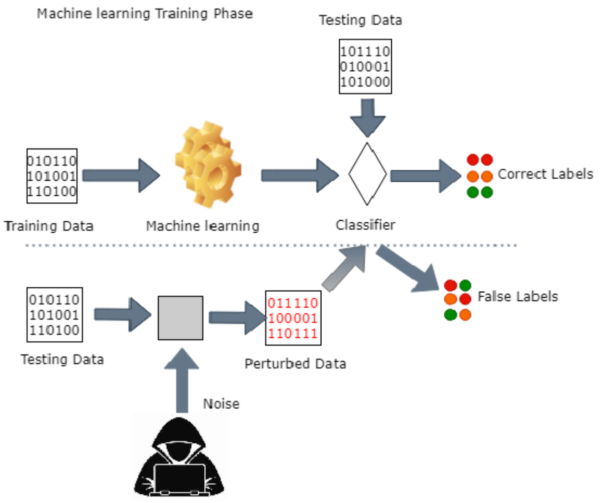


Figure 1: An adversarial machine learning. The upper layer represent the traditional machine learning or neural network training phase and the bottom part is adversarial attack phase.

Considering the challenges in existing classifiers, this paper aims to detect mutant of injection traffics from normal ones. A review of the literature shows that most previous studies on adversarial attacks detection have focused on altering the architectures of existing classifiers in order for them to distinguish between normal and adversarial samples [8, 17, 30]. However, we will not alter any architecture but rather add a regularization term to the neural network learning processes using metric learning method. Our metric learning method regularizes the classifiers representation space with N-pair loss term to learn a feature representation that detects malicious and benign traffics similarity and bring these similar traffics close to their original class and push dissimilar traffics away from their false classes.

The main limitation of the standard N-pair loss [13] is how to select representative triplets, which are made up of three examples from two different classes and jointly constitute a positive pair and a $(N-1)$ negative pairs and an anchor (*i.e.* where N is the cardinality of the set of triplets used in the training process). In this work, we select and modify the components of the N-pair loss with our projected gradient descent noise as mutant of existing injection attacks to enlarge the distance between attack examples and examples with different labels in the embedding space during the training phase. We then add this modified N-pair loss to the training process.

Contibutions: We propose to integrate a modified N-pair loss regularization term into neural network classifier to produce a more robust model against injection attacks in Web-based applications. We further propose a robust training method against mutants of these injected attacks to boost most on-the self machine learning classifiers ability to classifier these mutant attacks. Thus we inject projected gradient descent noise into the training data intentionally and randomly to achieve a more generalizable learning model. Our experimental findings on the CIC-IDS2018 datasets [10] demonstrate the superior performance of training neural network classifier with our regularizer in combination with our projected gradient noise proves to be more robust in detecting adversarial attacks, hence achieving a Detection rate over 98% in non-adversarial settings and 85% in adversarial settings.

We describe the layout for the remaining of this paper. In the next section, we will briefly review related work. Our proposed approach is described in Section 3. Section 4, shows experimental results applying the our method to the multiple datasets. Finally, we conclude the paper in the last Section 5.

2 Related Work

Sever-Side solutions defense measures [21, 22, 27] have been proposed to address the challenges Web-based applications face with injection attacks. The authors in [17] proposed two mitigation stages: Taint analysis and feedback message, to the developer with information concerning the vulnerabilities, their fixes and classification distributions. The authors in [22] propose a detection mechanism for detecting injection vulnerabilities looking Web page source angle. These defense mechanisms are incorporated on the application's servers, or set up as reverse proxy based on filtering rules and output escape, can avoid XSS attacks to some extent, however, these techniques are extremely time-consuming and error-prone to manually discover the keyword combination rules for new XSS attack statements [25]. Dynamic and static statistical analysis or a hybrid has also been discussed to address the injection attack problem [2, 11, 20]. The authors in [20] explores the use of taint analysis to flag candidate vulnerabilities and data mining to predict the existence of false positives. The work in [11] put forward a dynamic solution in the Java virtual machine, warning external attacks and recording specific taint propagation path. Their model adopted aspect-oriented programming (AOP) technology to insert monitor codes for efficiency improvement. The authors in [2], compared the parse tree of the SQL statement at run time before forwarding user inputs to the database server. However, a disadvantages of dynamic and static statistical analysis or a hybrid techniques are high over heads and precision lacking in identifying injection vulnerabilities [26].

Machine Learning techniques have also been adopted to address injection attacks [5, 12, 14, 28, 29]. The authors in [12] applied unsupervised machine learning algorithms KNN and affinity propagation to the detection of malicious Web pages. The work in [14] applied Naive Bayes, Decision Trees, and Multi-Layer Perceptron (MLP) to classify injections attacks. In [28], Decision Tree, Support Vector Machine and Naive Bayes algorithm were adopted to improve Web-based application security. The Naive Bayes algorithm achieved high classification accuracy and its calculation cost was relatively small, a limitation with the Naive Bayes algorithm is that, its based on the premise that each feature is independent of each other, and the characteristics of Cross-Site Scripting attack statement are usually closely related. Therefore, this will affect the recognition accuracy of Naive Bayes algorithm for Cross-Site Scripting attacks to a certain extent. Support Vector Machine out performed Naive Bayes algorithm in injection attacks recognition [29]. However, better accuracy and recall rates have been achieved by [5] using the AdaBoost algorithm. The conventional Machine learning approaches have achieved a certain degree of accuracy in the detection of injection attacks in Web-based applications to some extent, however, they are incapable of detecting tiny mutants of existing malicious (adversarial) attacks. Our method can not only detect injection attacks but also mutants of these attacks with higher margin.

3 Methodology

This section explains our proposed method (Figure 2) for injection detection in Web-based applications. Firstly, we collect a dataset in pcap format containing a number of instances representing each of the classes of interest, *i.e.*, benign and malicious, then we read the pcap file using CICFlowMeter network analysis tool [16]. The CICFlowMeter process pcap formatted data and select similar flow features into a standard output. We prelearned on these features using Doc2vec method and get an embedded representation for each of the features. Finally, we apply metric learning method to learn on the extracted features to produce similar representations for both inputs, in case the inputs are similar, or distant representations for the two inputs, in case they are dissimilar.

3.1 Dataset Preprocessing and Feature Selection

We rely on recent public and labeled dataset of network traffic that include SQL Injections (SQLi) and Cross-site Scripting (XSS) related attacks, CIC-IDS2018 [10]. As this dataset involve different types of attacks in pcap format, we only focus on these portion that include SQLi and XSS traffics.

We utilized CICFlowMeter network analysis tool to process pcap files and extract similar flow features. The CICFlowMeter network analysis tool is an open-source network traffic analyzer to analyze network activity log files from the pcap files. These logs describe the captured traffic and contain essential information for our feature selection. Determining the input packet size, each flow is formed by a sequence of up to 784 ($x = 784$) packets and each is made up of 16 features. We elected 16 most relevant features (*i.e.*. listed in Table 1) for feature representation. The features are collected from each flow independently. The selected features were stored as a sequence of comma separated values (CSV) files, each consists of 1,044,354 instances (Injections flows= 283,429, Legitimate flows=760,824) with 16 relevant features. We randomly split the dataset into two separate sets: 70% of samples is used for training and adjusting weights and biases and 30% for testing the models.

Table 1: List of selected features

1. Bwd Pkt size	9. Fwd IAT Min
2. Bwd Pkt Len Std	10. Bwd Pkt Len Mean
3. Bwd Pkt Len Std	11. ACK Flag Cnt
4. Bwd Pkt Len Max	12. Init Bwd Win Byts
5. Pkt Len Max	13. Init Bwd Win Byts
6. Flow IAT Min	14. PSH Flag Cnt
7. Flow Duration	15. Dst Port
8. Fwd IAT Tot	16. Flow IAT M

3.2 Feature Extraction with Doc2vec

Our Doc2vec uses a 3 layered feed forward neural network to gauge the context of the document and relate similar context phrases together and is based on the same distributional hypothesis, which signifies that the meaning of a sentence can be gauged by its context. Thus, if two sentences occur in the same position in two paragraphs, they are very much related either in semantics or syntactic in the same way.

This function is important in our work for new traffic classification. Because, new traffic might include many new words which might be in the form of SQLi and XSS queries.

In order to apply Doc2vec, we consider the network traffic as paragraphs with separate words. Besides the 16 features discussed above, a paragraphs also consists of Fully Qualified Domain Name (FQDN) and path. We separate FQDN and path by the delimiters such as dot, slash, question mark, equal, and so on. Our method also extract variable names and values from query string, which are used in running program to a server. These words constructs a document to train our Doc2vec model.

Given a document vector matrix D is a vector matrix for all paragraphs, each paragraphs is mapped to a unique vector that is represented by a column in matrix D, whereas each word is mapped to a unique vector that is represented by a column in matrix P.

Doc2vec constructs a vector space from the document and converts each paragraph in the document into vectors with the labels as defined in Equation (1). These labeled vectors are training data for the classifier.

$$y = b + vh(P_1, ..., P_k; P, D), \qquad (1)$$

here y is the output value of the Doc2vec, b is the bias terms between the hidden and output layers, v is weight matrix between the hidden and output layers, h is the concatenation for context words, m is the window size for preserving the contextual information and P is word embedding matrix.

3.3 Proposed Architecture

We reformulate the injection detection using metric learning method. We adopt the metric learning method akin to [13] the N-pair loss also termed as Multi-class triplet loss, consisting 3 different inputs samples, an anchor (x_a), a positive (x_p), and $\{x_i\}_{i=1}^{N-1}$ negative samples, where x_a and $\{x_i\}_{i=1}^{N-1}$ are from different class and x_a and x_p are from the same class. Given the batch inputs $\{x_a, x_p, x_n\}$, the objective of N-pair (Multi-class) loss is to push away the negative point x_n from the anchor x_a by a distance margin $\alpha > 0$ compared to the positive x_p:

$$||h(x_a) - h(x_p)||^2 + \alpha \leqslant ||h(x_a - h(x_n)||^2 \qquad (2)$$

where $x \in D$ is the input, and α is a hyper-parameter for margin. The standard N-pair loss function is defined as:

$$L_{MLM} = \frac{1}{N}\sum_{i=1}^{N} log\ (1 + \sum_{j \neq 1} exp(||h(x_a) - h(x_p)||_2^2 + \alpha - ||h(x_a - h(x_n)||_2^2)). \qquad (3)$$

here N is the cardinality of the set of triplets used in the training process.

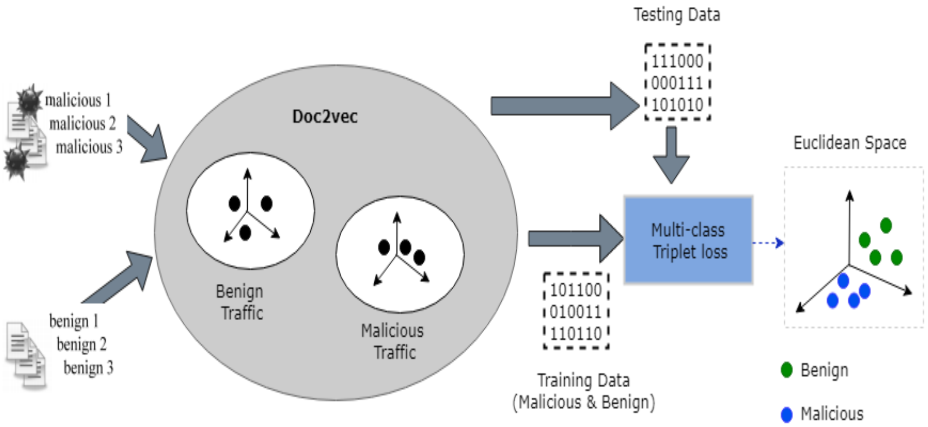


Figure 2: An overview of the proposed method

3.3.1 Robust Training Against Mutant of Injection Attacks

To induce stronger robustness in the proposed classifier, we propose a projected gradient noise which is introduced into the training input. By applying our robust training technique, we aim to improve the classifiers's generalization, *i.e.* predictions for samples outside of our training set. Generalization makes our classifier less sensitive to small perturbations, and therefore also more resilient to adversarial examples by finding the optimal parameter θ^* such that $\theta^* = argmin_{||x^*-x||}\ L(\theta, \triangle x, y)$, loss function $L(.)$ is used to train the model.

Given a constraint set $D \in R^k$, starting from training data x_i, mutants of injection attacks $x^{agm} = x_i + \triangle x_i$ are introduced into the classification model, generated through our projected gradient noise

$$\triangle x_i = P_D\ (x_i - \varepsilon sign\rho(\nabla F(x_i, y), +\lambda_i \nabla g(x_i))), \quad (4)$$

where ∇F is the unbiased stochastic gradient of F, ε is the perturbation size, the parameter $\rho \geq 0$ is the step size, λ corresponds to Lagrangian multiplier, λ corresponds to Lagrangian multiplier, i is the iteration counter and $g(x) \leq 0$ is a constraint function . P_D is a single projected operator and its optimization is represented by

$$P_D(x_0) = argmin_{x \in D} \frac{1}{2} ||x^{agm} - x_0||_2^2. \quad (5)$$

Generally speaking, as step size (ρ) increases, we assume that the accuracy of mutant samples will increase, while the accuracy of normal samples will decrease. It is worth noting that y is the predicted class produced by the classifier $F(x)$ rather than the true class labels of the input x_0 since accessibility to true class labels in the real world is impossible contrary to Adversarial Training [4]. We also selected predicted class label since we try to find a point $x \in D$ which is closet to the perturb input traffic x^{agm}.

We train all classifiers on these augmented samples $x^* = x + x^{agm}$. Our proposed method's training process is done under the supervision of Softmax Cross-Entropy loss (L_{SCE}) and N-pair loss (L_{MLM}). The final training objective is presented as

$$L_{all} = L_{SCE}(f(x^*), y) + \alpha L_{MLM}\{f(x_a), f(x_p^*), f(x_i^*)\} \quad (6)$$

here α controls the strength of the training stability, $f(\cdot)$ is the output of the last fully connected layer of our Doc2vec. The training process is different from the standard N-pair loss training [13] where components are clean samples, in this settings, we inject the adversarial noise into the training batch and chose the anchor (the center of the positive samples $x_a = 1/N_p \sum^{N_p} h(x_p^*)$) to serve as the decision boundary between the "true" class and the "false" class. We chose the center of positives to avoid the complex construction of triplets and hard sample mining mechanism as seen in the standard N-pair loss training [13]. Using the cosine similarity distance, we select negative examples as the nearest samples from he center. As a result, our model is able to learn to enlarge the boundary between the adversarial examples and their closest negative examples from the other classes. Figure 3 shows our modified N-pair Triplet loss for injection training.

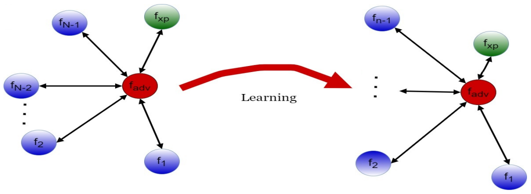


Figure 3: Illustration of our modified N-pair (Multi-class Triplet) Loss for injection attack detection training with $(N-1)$ triplets. The negative examples (blue), from a different class to the positive examples (green), is the closest input to the anchor in feature space. Our proposed loss learns to pull the anchor and positive examples from the true class closer, and push the $(N-1)$ negative examples of false classes apart, based on their similarity to the anchor example.

4 Experimental Results

We compare the performance of our proposed method with algorithms that have been employed in state-of-the art researches and have demonstrated their applicability to the task of identifying injection attacks in Web-based applications such as Support Vector Machine (SVM) in [5], Naive Bayes (NB) in [14], Decision Trees (DT) in [28], Multi-layer Perceptron (MLP) in [14], AdaBoost in [5] and KNN in [1]. We use MLM to denote our approach in Section 3.

Experiments were conducted on a Windows PC with Intel Core i7-2600 and a 16GB memory. MLM method was implemented using TensorFlow whiles the baselines were implemented using the Scikit-learn Python library. We evaluated baselines on the dataset discussed Section 3.2, whiles MLM method evaluation includes the application layer contents mentioned in Section 3.3. We trained the MLM network using the typical mini-batch stochastic gradient descent with momentum and used a grid search on a subspace of the hyper-parameter to select the ones which result in the best performance. The best value found for the hyper-parameter are $\alpha = 1.0$.

The objective function for MLM is defined in Equation (6). We evaluate our proposed model and baselines under two fold objectives,

- the performance of our proposed model and baselines in non-adversarial settings.
- performance under adversarial settings whiles considering the effect of our defense against adversarial attacks.

4.1 Performance in Non-Adversarial Settings

We measure the performance of our approach and baseline in non-adversarial settings through the Recall (Detection Rate), Precision and F1-score, which are computed as follows:

$$Precision = \frac{TP}{TP + FP} \quad (7)$$

$$Recall = \frac{TP}{TP + FN} \quad (8)$$

$$F1{-}score = 2 \times \frac{Precesion \times Recal}{Precesion + Recal} \quad (9)$$

where TP represent true positive, FP represent false positive and FN represent false negative. In our setting, we consider a positive to be an SQLi or XSS attack sample.

4.2 Non-Adversarial Settings Evaluation

We begin by determining which classifier reach a performance that complies with real-world requirements (F1-Score, Precision and Recall). Thus, we train and test our method and baselines considered in this paper against the SQLi and XSS injections included in the dataset.

Table 2: Distribution of Precision, Recall and F1-Score for all classifiers on CIC-IDS2018 dataset in non-adversarial settings

Classifiers	Precision	Recall	F1-score
SVM	98.13	98.38	98.03
NB	99.33	99.46	99.14
DT	99.24	98.31	98.95
MLP	97.32	98.41	98.52
AdaBoost	99.67	99.27	99.18
KNN	98.73	98.21	98.95
MLM	99.36	99.07	99.47

Figure 4, provide the curves of the test errors rate w.r.t training time in seconds. Note that the MLM loss induce faster convergence rate.

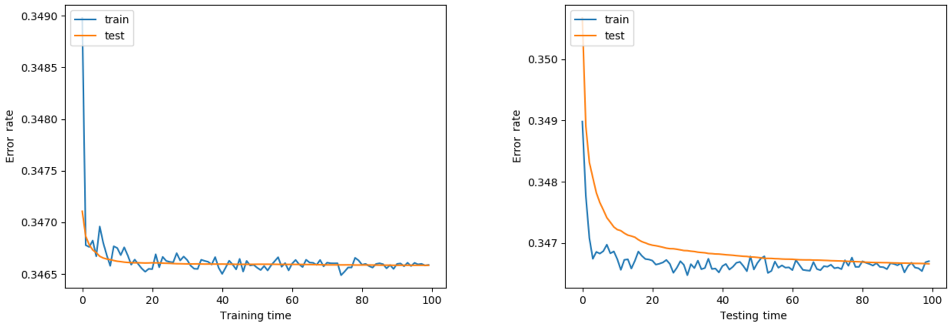
Aggregated experimental results obtained by our model and baselines classifiers are outlined in Table 2. Rows in this table indicate a specific classifiers, while columns illustrate the value of Precision, Recall and F1-Score evaluation metrics. Each cell contains the average value of a given evaluation metric. We can see that the MLM model achieve good detection performance, comparable to the baselines. Results are particularly promising for our proposed model, where F1-score is very near to 1. We highlight that our approach can be successfully applied to real network environments.

4.2.1 Performance in Adversarial Settings

To evaluate our model and baselines on a realistic adversarial attacks, we generate datasets of mutant injection examples $x^{'}_{*}$ using already established attacks generation tools with different combinations of the attacking parameters: the perturbation and iteration steps. The following tools were adopted in our work;

- The Iterative attack (PGD) proposed by [19], generates the adversarial inputs based on $x^{*}_{i-1} = \bigtriangledown_{x}\mathcal{L}(F(x); y)$.
- The Carlini & Wagner (C&W) attack proposed by [3], finds the minimal l_2 distortion by applying a binary search mechanism on the model parameters. C&W attack generate x^* by minimizing the loss $g(x) := max_x \ (max_i z(x)_i : i \neq t - z(x)_t, -t)$, where $z(x) = logit \ (h_{n-1} \ (x))$ and t controls the confidence on adversarial examples.
- The Basic Iteration Method (BIM) attacks proposed by [15], is an extension of FGSM by applying it multiple times with small steps, where the update formula at the i-th step is: $x^* = clip_{x,\epsilon} \ (x^{'}_{i-1} + \alpha \cdot sign \ (\bigtriangledown_{x^{'}_{i-1}}\mathcal{L}(F(x^{'}_{i-1}); y))$.

We only consider the SQLi and XSS attacks samples in our dataset and modify their features using the adversar-



(a)

(b)

Figure 4: (a) Test error rates on CIC-IDS2018 in non-adversarial settings w.r.t training time in seconds. (b) Test error rate on CIC-IDS2018 in adversarial settings w.r.t training time in seconds.

ial attack tools discussed above. We consider the untargeted $(L_\infty) = 0.8$ bounded attack settings during the attacks generation. The obtained adversarial datasets are then used to evaluate MLM and baselines. The effectiveness of these attacks are measured using the Detection Rate metrics defined in Equation (8).

Table 3: Average distribution of Detection Rate for all classifiers under 20 step PGD with random start of 20, 30 step C&W and 40 step BIM attacks. Adversarial perturbation level $\epsilon = 0.8$. Note that the Detection Rate of all classifiers reduces. High scores are indicated in bold.

Methods	20 PGD$_{20}$	C&W$_{30}$	BIM$_{40}$
SVM	68.62	69.57	67.24
NB	67.94	68.41	67.51
DT	71.72	74.86	69.31
MLP	45.94	62.32	43.23
AdaBoost	68.81	74.06	68.27
KNN	66.76	72.31	61.61
MLM	**77.47**	**81.21**	**72.36**

4.2.2 Adversarial Settings Evaluation

We now evaluate our baseline classifiers in the considered adversarial settings. We generate the adversarial samples and test MLM and baselines against these samples by following the procedure explained in Section 4.2.1.

Aggregated experimental results are outlined in Table 3. This table compares the average Detection Rate (Recall) of the classifiers with the Recall obtained on the adversarial samples. We focus on the Detection Rate since it reflects the number of adversarial samples that the classifiers are able to identify.

The lower values shown in Table 3 compared to Table 2 shows the effectiveness of adversarial attacks in invading classifiers. From these two tables we can see that the effects of adversarial samples change significantly for different classifiers.

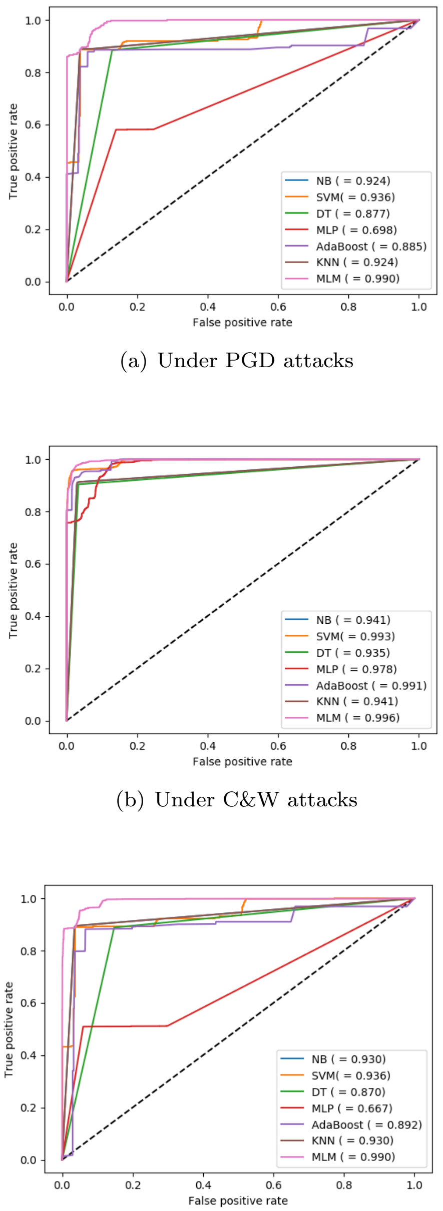
Table 4: Average distribution of Detection Rate for all classifiers under robust training our projected gradient descent noise. Note that the Detection Rate of MLM is on the top which shows MLM achieves higher detection efficiency for adversarial examples.

Methods	20 PGD$_{20}$	C&W$_{30}$	BIM$_{40}$
SVM	74.62	84.97	72.24
NB	73.51	85.41	72.94
DT	76.72	85.86	75.31
MLP	60.94	86.32	59.23
AdaBoost	75.27	86.06	73.81
KNN	73.76	84.31	72.61
MLM	**85.47**	**88.21**	**82.36**

As an example, for the SVM, DT, AdaBoost, KNN and NB classifiers Detection Rate fall by an average percentage of 29.88, 28.35, 28.89, 31.32, 29.50 respectively. The Detection Rate is even worse for MLP, the Recall drops by about an average of 44.58%. Compared with other classifiers, MLM performed considerably well with Detection Rate falling by an average of 22.19%.

Next, we assess the effectiveness of our defensive strategy based on projected gradient descent noise injection into the training process. To this purpose, we perform the robust training on MLM and baselines, as described in Section 3.3.1, and evaluate their Recall. Experimental results are summarized in Table 4. By comparing Tables 3 and 4 it is clear that our robust training improve the detection performance.

We also compare the ROC-curves and AUC score of MLM and baselines detecting mis-classified examples as shown in Figure 5 under a lower perturbation level $\epsilon = 0.3$, the ROC-curves and AUC score demonstrate that our approach is superior in adversarial example detection. Compared with baselines under the same robust training, MLM improves the AUC score by up to 2.3%. It can be seen in that our robust training method improved the adversarial attack detection rate across board. We



(a) Under PGD attacks

(b) Under C&W attacks

(c) Under BIM attacks

Figure 5: The ROC curve and AUC scores of detecting mis-classified examples. We test MLM and baselines on adversarial samples generated with a lower perturbation level $\epsilon = 0.3$ under 20 step PGD with random start of 20, 30 step C&W and 40 step BIM on all datasets. The numerical results for AUC score are shown in the legend. Note that the ROC curves of MLM is on the top and the AUC score of MLM is the highest, which shows MLM achieves higher detection efficiency for adversarial examples.

conclude that, for strong adversarial attacks requires an additive added noise to lead to a better empirical results, however, high step size might result in lower non-adversarial scenario accuracy and high accuracy in adversarial settings.

5 Discussion and Conclusion

Injection attacks present a great challenge in a web-based application environment. Though several attempts have been made by scholars to develop an efficient model for detecting and preventing these attacks, there is still the need to establish a more robust model to avert the injection attacks. Based on this premise, this study incorporated a modified N-pair term into a neural network classifier to generate a more robust model against injection attacks. The study again proposed a projected gradient descent noise to further improve the ability of the detection model to defend against tiny mutants of the existing injection attacks. The dataset from the state-of-the-art methods (literature) was utilized to examine neural network injection attack detection and prevention techniques. When the approach was assessed against untargeted and state-of-the-art adversarial attacks coupled with iterative attack (PGD), basic iterative method (BIM), and Carlini & Wagner (C&W), it was found that the combination of regularization method and projected gradient descent noise led to high accurate adversarial classification compared to state-of-the-art detection and techniques. Another merit of the approach adopted in this study is that it does not require modification to the neural network or machine learning model architecture, and as such, can improve the robustness of most-off-the-shelf neural networks and machine learning models with an additional overhead during training. It is suggested that future studies can incorporate an advanced defensive approach to improving the robustness of classifiers in adversarial settings.

Acknowledgments

The authors wish to thank the reviewers and editors for their invaluable suggestions that helped in improving the quality of the paper. This work was supported by the School of Information and Software Engineering, University of Science and Technology of China, Chengdu.

Declaration of Competing Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

References

[1] G. Aceto, A. Montieri, A. Pescapè, D. Ciuonzo, "Mobile encrypted traffic classification using deep learning: Experimental evaluation, lessons learned, and challenges," *IEEE Transactions on Network and Service Management*, vol. 16, no. 2, 2019.

[2] G. Buehrer, W. B. Weide, P. A. Sivilotti, "Using parse tree validation to prevent SQL injection attacks," in *Proceedings of the 5th International Work-*

shop on Software Engineering and Middleware, pp. 106-113, 2005.

[3] N. Carlini, A. Athalye, N. Papernot, W, Brendel, F. Rauber, D. Tsipras, A. Kurakin, "On evaluating adversarial robustness," *Machine Learning*, 2019. arXiv:1902.06705.

[4] N. Carlini, D. A. Wagner, "Adversarial examples are not easily detected: Bypassing ten detection methods," *Machine Learning*, 2017. arXiv:1705.07263.

[5] Y. J. Franceschi, A. Fawzi, and O. Fawzi, "Robustness of classifiers to uniform ℓ_p and gaussian noise," in *The 21st International Conference on Artificial Intelligence and Statistics (AISTATS'18)*, 2018. (`http://proceedings.mlr.press/v84/franceschi18a/franceschi18a.pdf`)

[6] K. Grosse, N. Papernot, P. Manoharan, M. Backes, P. McDaniel, "Adversarial perturbations against deep neural networks for malware classification," *Cryptography and Security*, 2016. arXiv:1606.04435.

[7] K. Grosse, N. Papernot, P. Manoharan, M. Backes, P. McDaniel, "Adversarial examples for malware detection," in *European Symposium on Research in Computer Security*, pp. 62-79, 2017.

[8] F. Guo, Z. Qingjie, L. Xuan, K. Xiaohui, Z. Jianwei, H. Yahong, T. Yuan, "Detecting adversarial examples via prediction difference for deep neural networks," *Information Sciences*, vol. 501, pp. 182–192, 2019.

[9] W. Hu, and Y. Tan, "Generating adversarial malware examples for black-box attacks based on GAN," *Machine Learning*, 2017. arXiv:1702.05983.

[10] S. Iman, L. M. Arash, and G. A. Ali, "Toward generating a new Intrusion detection dataset and intrusion traffic characterization," *The 4th International Conference on Information Systems Security and Privacy (ICISSP'18)*, Jan. 2018. (`https://www.scitepress.org/Papers/2018/66398/66398.pdf`)

[11] Z. Jingling, Q. Junxin, Z. Liang, C. Baojiang, "Dynamic taint tracking of Web application based on static code analysis," in *The 10th International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing*, pp. 96-101, 2016.

[12] H. B. Kazemian, and S. Ahmed, "Comparisons of machine learning techniques for detecting malicious webpages," *Expert Systems with Applications*, vol. 42, no. 3, pp. 1166-1177, 2015.

[13] S. Kihyuk "Improved deep metric learning with multi-class N-pair loss objective," NIPS, 1849-1857, 2016. (`https://www.nec-labs.com/uploads/images/Department-Images/MediaAnalytics/papers/nips16_npairmetriclearning.pdf`)

[14] S. Krishnaveni, K. Sathiyakumari, "Multiclass classification of XSS web page attack using machine learning techniques," *International Journal of Computer Applications*, vol. 74, no. 12, pp. 36-40, 2013.

[15] A. Kurakin I. J. Goodfellow, S. Bengio, "Adversarial examples in the physical world," in *The 5th International Conference on Learning Representations (ICLR'17)*, 2017. arXiv:1607.02533.

[16] A. H. Lashkari, G. Draper-Gil, S. I. M. Mohammad and G. A. Ali, "Characterization of tor traffic using time based features," in *Proceeding of The 3rd International Conference on Information System Security and Privacy*, 2017. DOI:10.5220/0006105602530262.

[17] B. Liang, H. Li, M. Su, X. Li, W. Shi, X. Wang, "Detecting adversarial image examples in deep neural networks with adaptive noise reduction,"*IEEE Transactions on Dependable and Secure Computing*, vol. 18, no. 1, pp. 72-85, 2018.

[18] Y. Liu, X. Chen, C. Liu, D. Song, "Delving into transferable adversarial examples and black-box attacks," *Machine Learning*, 2017. arXiv:1611.02770.

[19] A. Madry, A. Makelov, L. Schmidt, D. Tsipras, A. Vladu, "Towards deep learning models resistant to adversarial attacks," in *The 6th International Conference on Learning Representations (ICLR'18)*, 2018. arXiv:1706.06083.

[20] I. Medeiros, N. F. Neves, M. Correia, "Automatic detection and correction of web application vulnerabilities using data mining to predict false positives," in *ACM 23rd International Conference on World Wide Web*, pp. 63-73, 2014.

[21] I. Medeiros, N. Neves, M. Correia, "Detecting and removing web application vulnerabilities with static analysis and data mining," *IEEE Transactions on Reliability*, vol. 65, no. 1, pp. 54-69.2, 2016.

[22] M. Mohammadi, B. Chu, H. R. Lipford, "Detecting cross-site scripting vulnerabilities through automated unit testing," in *IEEE International Conference on Software Quality Reliability and Security (QRS'17)*, pp. 364-373, 2017.

[23] OWASP Top Ten Web Application Security Risks. (`https://owasp.org/www-project-top-ten`)

[24] N. Papernot, P. McDaniel, I. Goodfellow, S. Jha, Z. B. Celik, A. Swami, "Practical black-box attacks against machine learning," in *Proceedings of the ACM on Asia Conference on Computer and Communications Security*, pp. 506-519, 2017.

[25] J. Shanmugam and M. Ponnavaikko, "XSS application worms: New internet infestation and optimized protective measures," in *The Eighth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing (SNPD'07)*, vol. 3, pp. 1164-1169, 2017.

[26] L. K. Shar, H. B. K. Tan, "Auditing the XSS defence features implemented in web application programs," *IET Software*, vol. 6, no. 4, pp. 377–390, 2012.

[27] G. Shashank, B. B. Gupta, "XSS-immune: A Google chrome extension-based XSS defensive framework for contemporary platforms of web applications," *Security and Communication Networks*, vol. 9, no. 7, 2016.

[28] B. A. Vishnu, and K. P. Jevitha, "Prediction of cross-site scripting attack using machine learning al-

gorithms," in *Proceedings of International Conference on Interdisciplinary Advances in Applied Computing*, 2014. DOI:10.5120/12940-0033.

[29] R. Wang, X. Jia, Q. Li, "Machine learning based cross-site scripting detection in online social network," in *IEEE International Conference on High Performance Computing and Communications, IEEE 6th International Symposium on Cyberspace Safety and Security, IEEE 11th International Conference on Embedded Software and System*, pp. 823-828, 2014.

[30] W. Xu, D. Evans, Y. Qi, "Feature squeezing: Detecting adversarial examples in deep neural networks," *Computer Vision and Pattern Recognition*, 2018. arXiv:1704.01155.

Biography

Benjamin A. is currently a Ph.D. candidate at University of Electronic Science and Technology of China, Chengdu, China. His research interests include machine learning and deep learning, data mining, big data analysis.

Zhiguang Q. is currently a Full Professor with the School of Information and Software Engineering, University of Electronic Science and Technology of China (UESTC), where he is also the Director of the Key Laboratory of New Computer Application Technology and the UESTC-IBM Technology Center. His research interests include medical image processing, computer networking, information security, cryptography, information management, intelligent traffic, electronic commerce, distribution, and middleware.

Owusu-A. K. received the M.Sc. degree from Coventry University. He is currently pursuing the Ph.D. degree with the School of Information and Software Engineering, University of Electronic Science and Technology of China. His research interests include machine learning, data mining, big data analysis, applied cryptography, blockchain technology, and medical image processing.

Muhammed A. A. received the B.Sc. degree in computing-with-accounting from University for Development Studies–Navrongo, in 2017. He is currently pursuing the master's degree with the School of Information and Software Engineering, University of Electronic Science and Technology of China. His research interests include deep learning, Network security and the Internet of Things.

Access Control Model of Industrial Control System Based on Multi-attribute Decision Making

Rui-Hong Dong, Tong-Tong Xu, and Qiu-Yu Zhang
(Corresponding author: Tong-tong Xu)

School of Computer and communication, Lanzhou University of Technology
No. 287, Lan-Gong-Ping Road, Lanzhou 730050, China
Email: xutongtong1007@163.com

(Received Aug. 5, 2020; Revised and Accepted Apr. 24, 2021; First Online Oct. 19, 2021)

Abstract

For solving the problem that the access rights cannot be dynamically adjusted when the environment and task status change in the industrial control system, and in the collaborative environment, there are many problems, such as permissions switching, permissions changing frequently, which make it inability to carry out fine-grained access control. This paper puts forward a kind of access control model based on multiple attribute decision-making. This model firstly evaluates multi-attribute factors such as environment, resources, and tasks in access control by introducing entropy TOPSIS and dynamically reflects the risk values in the process of access control. Then, based on the user's historical access record, an algorithm to calculate the user's trust value is proposed, which is used to adjust the user's access permission dynamically. Furthermore, the model integrates access control with the industrial control system's organizational structure and task attributes. Finally, a natural gas pipeline access control data set published by the University of Mississippi is used to verify the effectiveness of the proposed user trust value algorithm and entropy weight TOPSIS method for user trust value adjustment and task state decision. The experimental results show that the model can meet the requirements of dynamic permission adjustment and fine-grained access control in the industrial control system environment and has high security.

Keywords: Access Control; Dynamic Authorization; Entropy TOPSIS Method; Multi-attribute Decision-Making

1 Introduction

The normal operation of various types of tasks in industrial control systems requires the coordination of task allocation and resources of various departments. Operation users will constantly change, and the user's access rights should also change when the context in which the user accesses changes [3,8,11,15]. At the same time, the description of authority and authorization management in industrial control systems are affected by multiple attributes such as task execution environment and task state [10]. Due to the diverse types of tasks and complex processes in the industrial control system, its security control is relatively complex, so reducing the complexity of authorization is also an issue to be considered [2]. Therefore, the study of access control model based on multi-attribute decision making in industrial control system has been widely concerned by scholars and has very important theoretical research and application value.

In recent years, many researchers have made rich achievements in access control for industrial control system environments. For example, task-based access control model combines access rights with tasks to propose task-based access control model, which can solve the problem of dynamic assignment of permissions in workflow, but does not separate roles from tasks. The work-based Access Control (WBAC) model [21] can better meet the security requirements in the industrial control workflow environment, but it is weak in the ability of dynamic management authority. Team-based Access Control (TMAC) [1, 17]is extended from the aspect of user organization structure to improve the descriptive ability of user authorization and reduce the tedious degree of authorization in the process of user sub authorization. Although task-role-based access control model [19] can adapt to access control in task collaboration environment, it is static authorization and the permission inheritance method is not flexible enough. Moreover, it cannot real-time monitor users in the process of task execution, so it cannot meet fine-grained access control requirements in task collaboration environment. The basic idea of the Organization Based 4 Level Access Control (OB4LAC) model [16]is that in the authorization process, information related to roles and positions is utilized. However, the OB4LAC authorization is static, and the model's constraints consist only of responsibility constraints and cardinality constraints, lacking finer grained constraints. The policy library of the Attribution-based Access Control model (ABAC) [6, 14] can be decentralized storage according to the actual situation, and can accurately describe the Access Control policy to realize fine-grained Access Control. However, in the system with a large number of subjects and resources, ABAC has too many access control rules. When the main body, resource attribute and environmental conditions increases, the number of rules will obviously increasing state, and an increase in the number of rules is likely to cause conflict strategy problem, at the same time the rapid increase of the number is likely to cause strategies library expansion, serious when even lead to normal operation of the system, is difficult to guarantee the system security and stability.

Literature [20] combined with ABAC integrates security level constraints into users, access behaviors and structured documents to realize multi-level access control mechanism of structured documents. It only conducts access control for a specific resource without considering the mutual constraints between tasks in workflow. The authorization method of the traditional access control model is still the authorization at the technical level. When the system is huge and complex, too many roles and permissions need to be managed, which increases the difficulty and complexity of authorization management, seriously degrades system performance, and even leads to authorization chaos [5, 9, 13]. The traditional access control model cannot be directly used in the workflow environment of industrial control systems, so a more fine-grained access control model is needed to divide specific tasks corresponding to specific resources, and realize dynamic control of permissions and real-time assessment of task risks.

Therefore, in order to solve the problems existing in the above research work, improve the flexibility of separation of responsibilities in the traditional access control model, and authorize the dynamic authorization in the process of task execution in real time. Combined with the industrial Control system, this paper presents a MATRBAC Model (Multi Attribute Task-based Access Control Model), which is more applicable to the fine-grained Access Control under the dynamic environment of the industrial Control system. The main contributions of this paper are as follows:

1) The proposed access control model based on multi-attribute decision making (MATRBAC) solves the problem that authorization is not flexible enough under the dynamic environment of industrial control system to realize real-time monitoring and management of users during task execution process.

2) In order to distinguish user credibility effectively and provide important basis for dynamic assignment of authority, an algorithm for calculating user trust value is proposed based on the user's historical access records.

3) By analyzing the data set of natural gas pipelines and comparing the characteristics of multi-attribute decision making under the industrial control system by linear weighting, analytic hierarchy process and Similarity by entropy Preference to an Ideal Solution, it is proved that the entropy TOPSIS method is more suitable for task state decision making under this environment.

The rest of the paper is organized as follows: Section 2 analyzes the factors associated with access control in the industrial control system and introduces the basic concepts of the MATRBAC model and the authorization process. Section 3 is the detailed design and introduction of two important modules of the MATRBAC model: the calculation of user trust value and the calculation of task transition risk value. In Section 4, we use a data set of a natural gas pipeline to analyze the effectiveness of the user trust value calculation method proposed by MATRBAC modeling, and compare the advantages and disadvantages of entropy weight TOPSIS algorithm and the same type of multi-attribute decision-making algorithm. Finally, the functional comparison and security analysis of MATRBAC model and other related access control models are carried out. In Section 5, the work of this paper is summarized.

2 Design of MATRBAC Model

2.1 Factors Related To Access Control in Industrial Control Systems

An abstract model of the factors associated with access control in an industrial control system is shown in Figure 1, which includes users, organizations, business roles, tasks, business processes, and business rules. In general, users belonging to an organizational structure perform assignments based on their jobs or business roles. Some tasks comprise business processes with special access control requirements. Many task rules and the user, environment, resource and other attributes involved in the process of task execution involve access control of various businesses. The access control model of industrial control systems should not only support finer grained access control but also ensure the separation of responsibilities and the principle of least privilege [12]. The specific attribute information involved in each factor in the access control process is analyzed below.

As shown in Figure 1 , the specific attribute information involved in each factor in the access control process mainly includes the following factors:

Task attributes: Include task type, task status, and task dependencies. From the perspective of access control, literature [4] can classify tasks in industrial control systems according to whether they can be inherited or whether they can be accessed passively: class P (Private), class S (Supervision), class

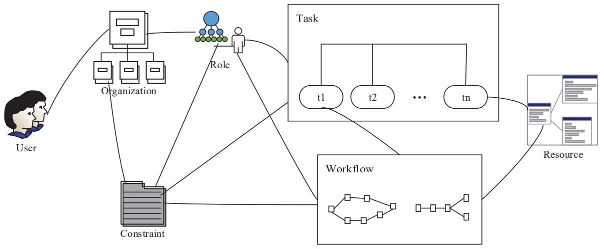


Figure 1: Abstract model of factors related to access control in an industrial control system

W (Workflow) and class A (Approval for activity). The task classification is shown in Table 1.

Table 1: Task classification in industrial control system

Pattern\Inheritance	non-inheritable	inheritable
active access	Workflow tasks(W)	approval tasks(A)
passive access	Private tasks(P)	Supervision (S)

Task states have four states: Ready, executing, suspended, and revoked. Task dependencies include the tasks that must be completed before the current task and the next task after the current task is completed.

User attribute: Subject specifically to a user in the industrial control system, each user has attributes associated with it, and these properties represent the main body status and characteristics of the included user ID, user roles, login name and password, the credibility of the user, system will be based on user history operation situation for real-time update, the concrete will be introduced in Section 4.2.

Resource attribute: [7] categorizes resources under the Internet of Things environment. In this paper, resources in industrial control systems are divided into device resources and data resources. The device can have the following attributes: device ID, device type, device location, device life, emergency alarm turn. Data resource attribute definition: device ID, information data, information data type, information data encoding, *etc.*

Operational attribute: Actions represent actions performed by the user on a resource, such as opening, closing, reading, writing, deleting, *etc.*

Environment attribute: In industrial control systems the environment refers primarily to the current time and date.

2.2 Structure of the MATRBAC

After the expansion of MATRBAC on the basis of ABAC and T-RBAC, the key improvements of the MATRBAC model are:

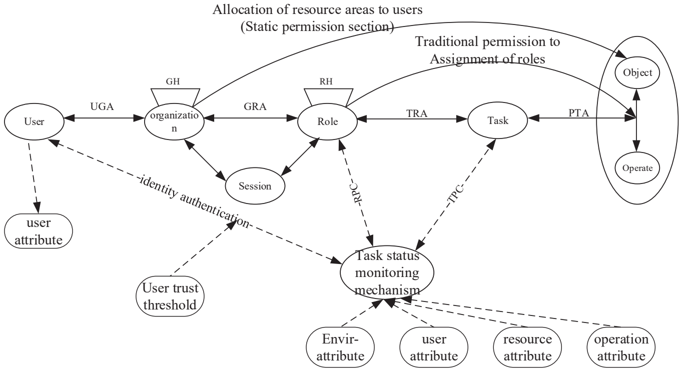


Figure 2: Structure of the MATRBAC model

1) Incorporating time and environmental constraints into the model so that the authority of the system is likely to change following the dynamic changes of time and environment;

2) The original four-tier structure of user-role-task-permission has been transformed into a five-tier structure of user-organization-role-task-permission. The introduction of organizational components on the basis of user and role components makes the model and industrial control application scenario more in line with reality. Moreover, when authorization is needed during the operation of the same task, the organization can automatically authorize through the assignment of roles, effectively reducing the pressure of administrator authorization;

3) A supervisory mechanism is added on the basis of roles and task components, which enhances the constraint correlation between roles and tasks and enhances the dynamic authorization between roles and tasks. Put an end to information leakage caused by malicious use of roles;

4) In the process of distributing the trust constraint is introduced into the dynamic permissions, the user's trust by user attribute and environmental attribute calculation, the user trust as a task execution status under dynamic decision-making authority of an attribute, along with the environment attribute, role attribute, time attribute and the task attribute calculation value at risk of the currently executing task, decision whether if the current task to terminate or permission to withdraw. The multi-attribute access control characteristics of the MATRBAC model are shown in Figure 2.

User: A user of the system or an actor of a task, and the set of users is denoted as $USERS$.

Organization: The corresponding organization department or position in the system, and the set of organization is denoted as $ORGANIZATION$.

Role: The permissions needed to implement a function or complete a business. A role is a set of permissions to perform a task. Users belonging to a role have the right and relevant permissions to perform the corresponding task. The set of roles is denoted as $ROLES$.

Session: Mapping between the organization and a subset of the set of roles that the organization owns. The organization $USES$ a session to activate a role, and the activated role has basic static permissions for parts of that role. A session is an active process of a role and sessions represents a set of $SESSIONS$.

Task: A logical unit of work that is indivisible and must be performed completely. A task is not a task in the real world, but an abstract representation of a class of tasks in the real world. Each task corresponds to a number of resource access permissions necessary to perform the task. The set of tasks is represented by $TASKS$.

Object: The content of a resource accessed by a user and also a protected object.

Operation: An executable program that is used by the user to perform operations on a resource (for example, read, write, update, delete).The set of all operations is represented by OPR. Any action the user performs on an object is completely defined within the permissions.

Permission: The collection of all permissions, represented by $PERMS$, is a combination of objects and operations.

2.3 Access Control Procedures for Models

In the model of MATRBAC access control described in this paper, the control of user rights assignment is realized dynamically through the interaction of user trust mechanism and task status supervision mechanism. The access control logic of the MATRBAC model is shown in Figure 3.

As shown in Figure 3, the authorization process of the MATRBAC access control model may be divided into the following five steps:

Step 1: The user logs in and issues an organization activation application. The organization controls the activation state of the user organization by judging the legitimacy of the user. Assign the organization when the user is legal, otherwise reject the user request.

Step 2: The organization assigns a role to the user. After the role is activated, it applies for permission to execute the task.

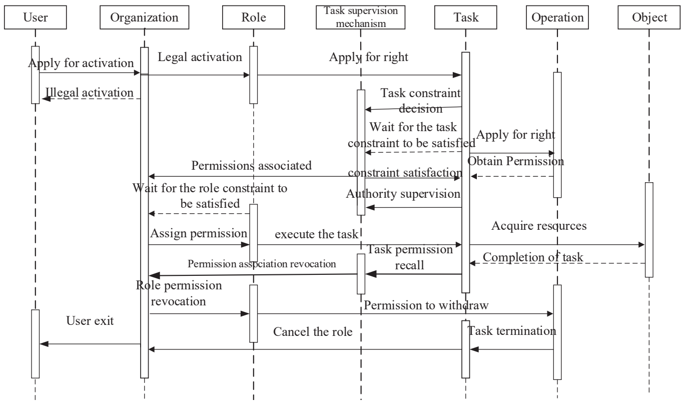


Figure 3: Access control logic of the MATRBAC model

Step 3: Task state supervision mechanism determines the task permission constraint. If the constraint does not meet, the task will be suspended and wait for the task to meet the constraint. When the task meets the constraint, the task will obtain the execution permission.

Step 4: Dynamically monitor whether constraints are met during task execution. If not, the task status monitoring mechanism revokes the permission and the task is suspended or terminated. If so, the task will continue to perform as normal until the permission is automatically revoked at the end of the task.

Step 5: After the normal completion of the task, reclaim the role permission, revoke the role task association, revoke the permission organization association, and the user exits.

3 Implementation the MATRBAC Model

The dynamic authorization and fine-grained authority management of the MATRBAC model is realized through the task status supervision module and the user trust determination module. Multi-attribute decision making is an important application in task state monitoring module. A user trust value determination algorithm is proposed, which can distinguish the user credibility effectively by the user's historical operation record and reputation record, and has the time attenuation factor. The following will describe in detail the multi-attribute decision is the task status monitoring module and the user trust determination module.

3.1 Task Status Monitoring Module

Task status supervision mechanism is the user activated after a character is in the process of performing tasks on the supervision and management of a function module, through the user properties, environment, resources, task attribute of the constraints of dynamic judgment, the role of permissions granted to and revoked the real-time dynamic monitoring, can strengthen the roles are endowed with the strength of the rules to follow, improve the security in the process of task execution. By introducing the task state dynamic management module to decouple role authorization and role behavior, the system's ability to resist malicious attacks can be enhanced.

The factors affecting the industrial control system environment analyzed in Section 2.1 include user attribute, environment attribute, resource attribute and task attribute. The multi-attribute decision making algorithm is used to calculate the risk value of task state. Finally, a reasonable threshold value is obtained through systematic testing. If the calculated risk value is larger than the set threshold value, the execution state of the task will be changed, and after risk screening, the task state can enter the ready state and continue. Task status monitoring mechanism is shown in Figure 4. Its core components include task status monitor, authorization structure, policy decision point, policy information point, *etc.*, which are defined as follows:

Task status monitor: Real-time monitoring of task execution state, acquisition of current task attributes, and calculation of current task risk value. If the risk value is greater than the set threshold value, the authorization structure can be suspended or revoked directly, and the current environment attributes, user attributes and operation attributes can be returned to the decision information point.

Authorization structure: It refers to the set of resources required to complete a task or operation, including the entrusted user and the operation set. The entrusted user is the set of personnel to complete the current task or operation, and the operation set refers to the set of permissions required to complete this part of operation.

Policy Decision Point: It refers to the attribute information provided to the policy information point, gives the permission decision, and creates the authorization structure.

Policy information points: Collect and save key attribute information, including principal attribute, action attribute, environment attribute, and resource attribute.

Operation steps of task status monitoring mechanism:

Step 1: Set the subject attribute, operation attribute, resource attribute and environment attribute in advance.

Step 2: Policy information points collect, filter, and store these attribute values.

Step 3: The policy decision point requests relevant attribute information from the policy information point according to the task attribute.

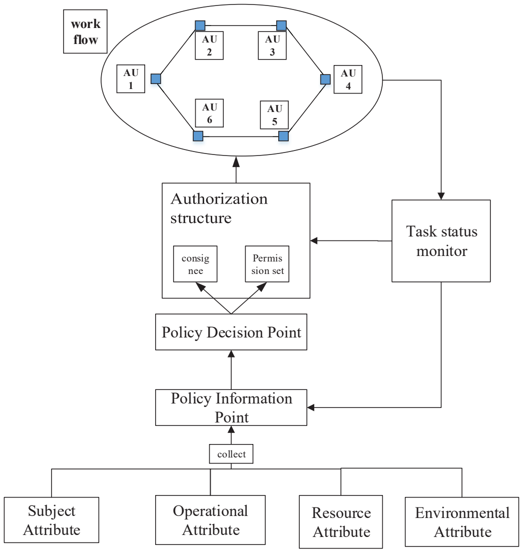


Figure 4: Task status monitoring mechanism

Step 4: The policy decision point creates the authorization structure through the collected attribute information and task attribute information.

Step 5: Authorize the structure to implement the action on the task.

Step 6: The task status monitor monitors the task status in real time.

Step 7: If the task status risk value is greater than the set risk threshold, the task status monitor will adjust, suspend or revoke the state of the authorization structure as the case may be, and send the collected attribute information to the policy information point to provide the basis for the next decision.

Step 8: The authorization structure is revoked immediately after it completes the task.

Common multi-attribute decision making algorithms that can be used to calculate task risk are

1) Simple weighting (SAW);

2) AHP;

3) Entropy weight TOPSIS.

The simple weighting algorithm is simple and easy to understand, but it requires that the values of all attributes are constant and can be compared, and that there is no important complementarity between attributes. Considering that the attributes related to access control in industrial control systems are complementary, and that this method cannot reflect the prominent influence of some attribute indexes, thus resulting in the distortion of evaluation results, this paper does not adopt this method. A comparison of the three methods is given in Section 4 with an example. In this paper, TOPSIS method is used to calculate the risk value of task status.

TOPSIS (Technique for Order Preference by Similarity to an Ideal Solution) [5] was first proposed by C.L.Wang and K.oon in 1981. TOPSIS is a sorting method based on the proximity of a finite number of evaluation objects to the ideal target, and evaluates the relative merits of existing objects. The basic principle is to sort the evaluation object by detecting the distance between the evaluation object and the optimal solution and the worst solution. If the evaluation object is the closest to the optimal solution and the furthest away from the worst solution, it is the best.Otherwise it's not optimal. Among them, each index value of the optimal solution achieves the optimal value of each evaluation index.Each index value of the worst solution reaches the worst value of each evaluation index. The process of TOPSIS Algorithm 1 is as follows: A positive ideal solution is a hypothetical optimal case in which each attribute value achieves the best of the alternatives. The negative ideal solution is the worst imaginable, with each attribute value reaching the worst of the alternatives. The ranking rule of schemes is to compare alternatives with ideal solution and negative ideal solution. If one of the alternatives is most close to ideal solution, but at the same time far away from negative ideal solution, the scheme is the best scheme among alternatives. The evaluation criteria of positive ideal solution and negative ideal solution are generally divided into three types:

Algorithm 1 TOPSIS

```
Begin
Initialize the Raw data set $R$ and the weight of each index $w = w_1, w_2, ..., w_n$
The index attributes in the original data set are converted into $B$ in the same direction.
Set up the weighted normalization matrix $Z$.
for each $z_j \in Z$ do
    The $z_j$ dimension of the optimal scheme $Z^- \leftarrow A^+$ Element minimum
    The $z_j$ dimension of the optimal scheme $Z^+ \leftarrow A^-$ Element maximum
end for
for each $z_j \in Z$ do
    Degree of proximity between $z_j$ and the optimal scheme $d_i^+$. (Formula 7)
    $z_j$ proximity to the worst $d_i^-$. (Formula 7)
    Degree of closeness between $z_j$ and the optimal scheme $C_i$. (Formula 8)
end for
Sort by size $C_i$
TOPSIS evaluation results of each data sample
```

Very small indicators: the smaller the expected index value is, the better (such as morbidity and mortality).

Intermediate index: the expected index value should be neither too large nor too small, and the appropriate intermediate value should be the best (such as the PH value of water quality assessment).

Interval index: The best value of expected index should be in a certain interval (such as body temperature). The ideal solution of dynamic decision making based on multi-attribute access control authority in industrial control system belongs to intermediate index.

Assumption in the industrial control system standard values for a property, as an index of the maximum possible value, m as an index of possible value of the minimum value, then the properties of expectations index value computation formula is as follows:

$$x' = \left\{ \begin{array}{ll} 2\frac{x-m}{M-m}, & m \leq x \leq \frac{1}{2}(M+m) \\ 2\frac{M-x}{M-m}, & \frac{1}{2}(M+m) \leq x \leq M \end{array} \right. \quad (1)$$

In this article for each attribute weights calculation using the entropy weight method, entropy method is more objective and can better explain the results, the use of information between the variability (*i.e.*, diversity) for empowerment, but need to have some sample size when using this method, through the sample to determine the weights, as determined by a weight on the analysis of the new things. In the specific application process, the entropy weight method calculates the entropy weight of each index according to the variation degree of each index by using the information entropy, and then modifies the weight of each index through the entropy weight, so as to obtain the more objective index weight.

Step of task status risk value calculation:

Step 1: Analysis task status by the risk value by the user attribute,environment attribute, multiple attribute decision factors such as task attribute, attributes expressed with Q, build property set:Q. The task set to be analyzed is represented by P. The attribute of task $Pi$for the Qj corresponding attribute values expressed in rij $(i = 1, 2, ..., m; j = 1, 2, ..., n)$.

Step 2: Each attribute value of each task to be analyzed is represented by rij to obtain the multi-attribute decision matrix R:

$$R = (r_{ij})_{m\times n} = \begin{bmatrix} r_{11} & r_{12} & \cdot & \cdot & \cdot & r_{1m} \\ r_{21} & r_{22} & \cdot & \cdot & \cdot & r_{2m} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ r_{n1} & r_{n2} & \cdot & \cdot & \cdot & r_{nm} \end{bmatrix} \quad (2)$$

Step 3: The weight of each attribute is calculated by entropy weight method: Calculate the proportion of the attribute value of the i-th sample method in the j-th attribute p_{ij}:

$$p_{ij} = r_{ij} / \sum_{i=1}^{m} r_{ij} \quad (3)$$

Calculate the entropy value e_{ij} of the j-th index:

$$e_{ij} = -k \sum_{i=1}^{m} p_{ij} \cdot \ln p_{ij} \qquad (4)$$

with $k = 1/\ln m$. Calculate the entropy weight of the j-th index w_j:

$$\omega_j = (1 - e_j) / \sum_{j=1}^{n} (1 - e_j) \qquad (5)$$

The relative importance of each factor is expressed as $w_1, w_2, ..., w_n$ represents and meets the normalization condition:

$$\sum_{j=1}^{m} w_j = 1 \qquad (6)$$

Step 4: The decision matrix $R = [r_{ij}]_{m*n}$was standardized to obtain the matrix $B = [b_{ij}]_{m*n}$.

Step 5: Establish the weighted standardized matrix $Z = [z_{ij}]_{m*n}$, with $z_{ij} = b_{ij} w_j$.

Step 6:] Determine the positive and negative ideal solutions $A^+ = (z_1^+, z_2^+, ..., z_n^+)$and $A^- = (z_1^-, z_2^-, ..., z_n^-)$, $A^+$and A^-represent the most ideal and the least ideal solutions respectively. Where:
$z_j^+ = \max z_{ij}, z_j^- = \min z_{ij}$ $j \in$ Efficiency measure
$z_j^+ = \min z_{ij}, z_j^- = \max z_{ij}$ $j \in$ Cost type measure

Step 7: Compute the Euclid distance d_i^+ and d_i^+ between each solution and the positive and negative ideal solution:

$$d_i^+ = \sqrt{\sum_{j=1}^{m} \left(z_{ij} - z_j^+\right)^2}, d_i^- = \sqrt{\sum_{j=1}^{m} \left(z_{ij} - z_j^-\right)^2} \qquad (7)$$

Step 8: Calculate the paste progress of each task sample and positive ideal scheme:

$$C_i = \frac{d_i^+}{d_i^+ + d_i^-} \qquad (8)$$

Step 9: According to the characteristics of the system attribute value, set the risk threshold and compare it with C_i. If C_i is less than the set threshold, it will be judged as a risk task and the task-related permissions will be adjusted.

3.2 User Credibility Determination Module

When a user is granted an organization and the corresponding role, he/she already has some basic permissions, but during the execution of the task, Is granted other advanced permissions according to the different tasks. In order to prevent the abuse of permissions in the process of task execution, every time the user opens a task, a judgment is made on the credibility of the user, and then compared with the preset threshold, so as to determine whether the user has the authority to execute the task.

User credibility is related to three parts:

1) Historical credibility,

2) Reputation and credibility,

3) Time attenuation factor.

Historical trust value calculation: When the user requests to access system resources, the system will obtain the user's operation record and record it in the database. Firstly, the historical operation record of the user is obtained. The number of legal operations of the user and the resource is denoted as n. The illegal interactive access times of users are recorded as m, V_{ni} represents the trust value after each legal access, and V_{mi} represents the trust value after each illegal access. When the user accesses legally, $n = n + 1$, Otherwise, $m = m + 1$, V_k is used to represent the trust value after the end of the current access behavior, which can be calculated according to Formula (1).

$$V_k = \frac{\sum_{i=0}^{n} V_{ni}}{n} - \frac{\sum_{i=0}^{m} V_{mi}}{m} \qquad (9)$$

Calculate the behavior trust value before the current access V_k', and then calculate the actual behavior trust value V_h based on the current and previous trust values.

$$V_h = \frac{V'k + V_k}{2} \qquad (10)$$

When a user accesses the resources in the cloud platform for the first time, due to the lack of historical data, it is impossible to accurately calculate the actual trust value of the user. So his actual behavioral trust value is equal to the current behavioral trust value.

$$V_h = V_k \qquad (11)$$

To sum up, the actual trust value of the user's historical behavior is:

$$V_h = \begin{cases} \frac{V'k + V_k}{2}, & (\text{ Non } - \text{ first access }) \\ V_k, & (\text{For the first time to access}) \end{cases} \qquad (12)$$

Reputation is a measure of a user's trust by the other. In a system, the user is not only the visitor of the resource but also the owner of the resource. Access is a two-way process, during which mutual trust values are formed. The trust value cannot be passed, but by listening to the wishes of another user, the trust value can be adjusted to better meet the requirements of the system. It is not comprehensive to control the user's behavior by relying solely on the historical trust value, and it may face the problem of malicious access. Therefore, the trust of the third party is increased to monitor the user's behavior, that is, to conduct the trust evaluation based on the evaluation information of the user by others. When user u accesses other resources, other resources also have a trust value store for user u. Then the reputation trust value of the user is:

$$V_{ck} = \frac{\sum_{k=0}^{q} c_k}{q} \tag{13}$$

Where, V_{ck} represents the reputation trust value of the user, C_k represents the trust value of the third party to the current user, and q represents the number of the given third party to the user. Trust value decays over time: If a user has not accessed a resource for a long time, his trust value for that resource also decays over time. The trust value plus the time decay factor, the trust value between people in the real world will be disconnected for a long time and the trust value will decrease with time. The decay rate of the trust value is not constant and is related to time. The time decay factor is expressed by a(t), and the calculation formula is:

$$\mathrm{a(t)} = e^{-k(t-t_0)} \tag{14}$$

Where k is the set attenuation coefficient, t is the current time, t_0 is the last access time.

According to the above analysis, the reliability calculation formula of users is as follows:

$$\mathrm{T}\left(u_t\right) = \mathrm{a(t)}\left(\alpha V_h + \beta V_{ck}\right) \tag{15}$$

Where $\alpha ? \beta$ are the weights of historical trust value and reputation trust value respectively, the value range of and is [0,1], and $\alpha + \beta = 1$.

4 Experiment and Safety Analysis

4.1 Introduction of Experimental Data Set and Experimental Environment

The data set used in this experiment is the data set obtained from the laboratory-scale natural gas pipeline system of the University of Mississippi [18]. There are a total of 274,628 records in the data set, of which 210,528 records have incomplete characteristics. If the missing data is filled with average value or other methods, it is difficult to reflect the real state of the system for the multi-attribute decision making scheme studied in this paper. In view of this, records with missing features and attributes that are not relevant to the access control schemes studied in this article are removed, and some attributes related to access control mentioned in this article are appropriately added based on the results of the data set. The dataset contains four types of attacks: response injection, detection, distributed denial of service injection, and command injection. Response injection is divided into simple malicious injection (NMRI) and complex malicious injection (CMRI). Command injection can be divided into malicious status command injection (MSCI), malicious parameter command injection (MOCI) and malicious functional code injection (MFCI) attacks. Detect attacks to collect control system network information, draw network architecture, identify equipment characteristics, such as manufacturer, model, database information, *etc.* Table 2 lists the access-control-related attribute selected for this paper and their descriptions.

Table 2: Access control-related attribute and descriptions selected for this article

Attribute	Descriptions
address	MODBUS slave device workstation address
taskcode	Function code
usertrust	User trust value
time	Time to perform tasks
PIDgain	PID gain
PIDresetrate	PID reset rate
PIDrate	PID rate
pressure	Pressure measurement
deadband	PID dead band
cycletime	PID cycle time
cmdresponse	Command or response
Attack category	Category of attack
specific result	Specific attack

The experimental hardware platform was Intel(R) Core(TM) I5-7300HQ CPU, 2.50ghz, and 8GB of memory. The experimental software platform is: operating system bit windows 10, development tool is Python 3.7.

4.2 Experimental Methods and Results Evaluation

Dynamic adjustment of user trust value: given two users, the initial trust value is 0.5, and the trust time decay coefficient is 0.2. The change of trust value is calculated in the interaction process of 50 tasks selected respectively to see whether it is consistent with the actual situation. Figure 5shows the change in trust values between the two users over 50 tasks. As can be seen from Figure 5, at the beginning, the trust values of users u1 and u2 both showed

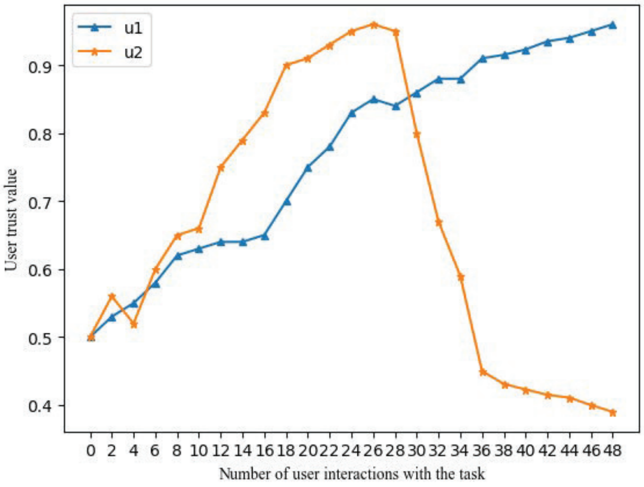


Figure 5: Fluctuation curve of user trust value

a fluctuating rising trend. u1 rose more slowly than u2, but u2's trust value steadily rose, and u1 first gained the highest trust value in the interaction. From the 27th task, u2 trust value gradually decreased. According to the data analysis, u2 trust value decreased rapidly due to the influence of time decay factor 30 days between the 27th task and the 28th task. It conforms to the attenuation law of fast and slow decline of trust value. The experiment shows that the proposed method is effective. This method can accurately reflect the influence of user's behavior on the trust value under different task scenarios.

Calculation of task status risk value: In this experiment, 1000 pieces of data from the above attributes and corresponding data set were selected for risk value calculation by simple weighting method, analytic hierarchy process and entropy weight TOPSIS method respectively. The data includes both the normal and the attack. This paper evaluates the three multi-attribute decision making algorithms in terms of the degree of differentiation of calculated risk values, and then selects the algorithm more suitable for the system's multi-attribute access control risk value calculation. The matching degree between the high risk state calculated by entropy method, analytic hierarchy process and TOPSIS method and the actual situation was compared respectively. The lower the value of risk calculated under normal circumstances, the better. The higher the value of risk calculated under abnormal circumstances, the better. The entropy weight method is used to calculate the weight points of 11 attributes, as shown in Figure 6. As can be seen from Figure 6, the weights of the 11 attributes selected in this paper calculated by information entropy method are respectively 0.01359, 0.02313, 0.2604, 0.1142, 0.10146, 0.0752, 0.048, 0.0634, 0.1536, 0.0427 and 0.10432. Among them, the weight of user trust value is the largest and the weight of address is the smallest, which conforms to the actual rules in the system.

The 1000 samples were calculated according to the simple weighting method, analytic hierarchy process and entropy weight TOPSIS method, and then sorted into 20 small samples to calculate their mean value, as shown in Table 3. Among them, samples 17-20 were samples with abnormal task status. It can be seen from the table that the task risk value calculated by simple weighting method and hierarchical analysis is not clearly distinguished due to the small difference in sample data, while the entropy TOPSIS method can be clearly seen in the table that the risk value changes significantly between samples 16 and 17. In order to show the risk differentiation degree of the three algorithms in the system data set more clearly, the comparison diagram of the risk differentiation degree of simple weighting method, analytic hierarchy process and entropy weight TOPSIS in this sample is drawn, as shown in Figure 7. It is obvious that the TOPSIS method has a high degree of distinction. It can be seen from Figure 7 that the entropy weight TOPSIS method changes significantly in the risk values of sample 16 and sample 17 compared with the other two algorithms, which can intuitively reflect the attacked samples of the system, indicating that the entropy weight TOPSIS is more in line with the multi-attribute decision making of task risk under the system environment.

Table 3: Three algorithms respectively calculate the mean risk of their corresponding samples

Sample	SLW	AHP	Entropy TOPSIS
Sample1	13.333	13.960	0.976
Sample2	13.236	13.873	0.972
Sample3	13.229	13.570	0.96
Sample4	13.206	13.533	0.944
Sample5	13.179	13.313	0.942
Sample6	12.996	13.168	0.929
Sample7	12.839	12.987	0.89
Sample8	12.774	12.878	0.887
Sample9	12.760	12.675	0.857
Sample10	12.746	12.636	0.844
Sample11	12.723	12.431	0.839
Sample12	12.660	12.130	0.829
Sample13	12.644	12.091	0.826
Sample14	12.630	12.067	0.817
Sample15	12.612	12.032	0.812
Sample16	12.584	12.001	0.803
Sample17	11.683	11.700	0.426
Sample18	11.349	11.697	0.093
Sample19	11.023	11.601	0.076
Sample20	11.005	11.389	0.075

4.3 Security Analysis of the Model

In general, access control models in an industrial control system environment must meet the following requirements:

1) Fine-grained access control.

2) With stronger expansibility.

3) The authorization process is simple.

4) Active and passive access control coexist.

5) Dynamic detection and risk quantification of users' behaviors.

The model of MATRBAC achieves finer grained access control. The traditional role-based access control model assigns permissions to the role level. The access control model based on the task role correlates the task with the role, and the permission assignment realizes the more granular assignment of the permission at the task level to some extent. Following the role based access control model, MATRBAC retains the access control mechanism of the task role-based access control model, In addition, in

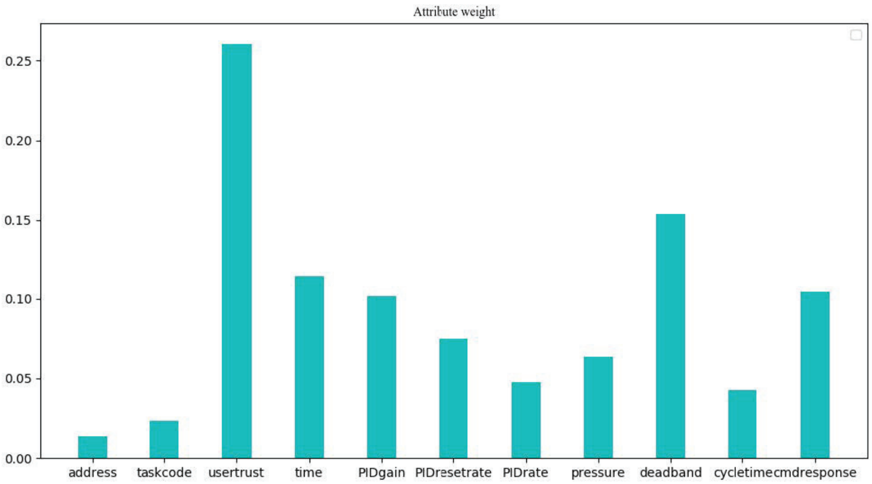


Figure 6: Weight distribution of 11 attributes in the dataset

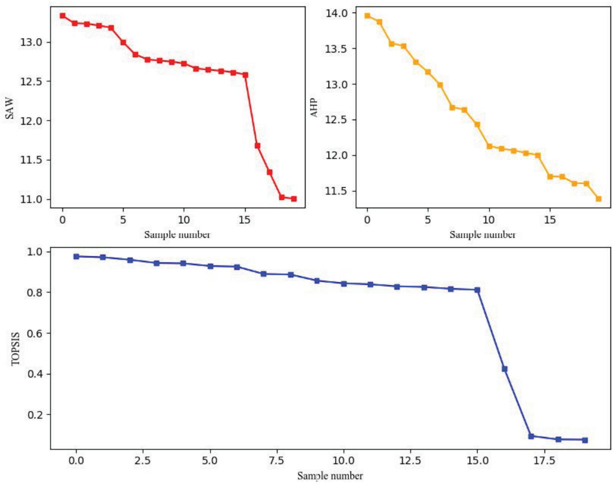


Figure 7: Weight distribution of 11 attributes in the dataset

the process of task execution, by collecting and saving key attribute information, including subject attribute, operation attribute, environment attribute and resource attribute, task execution is subject to environmental constraints, which effectively increases the ability of access control model to describe fine-grained constraints. The entropy weight TOPSIS is introduced to evaluate the risk of task state, which can adjust the access control strategy in real time according to the specific task running state and make access granularity finer.

The MATRBAC model is more scalable. The introduction of organization and role in the MATRBAC model is more consistent with the structure of industrial control system in reality, especially the classification of tasks, making different types of tasks subject to different access control strategy constraints and more consistent with the multi-task scenario in industrial production.

The MATRBAC model simplifies the authorization process. Through the introduction of organization, all and part of authority inheritance in the organization structure can be realized through the association relationship between organization and role and the hierarchical inheritance relationship between organization and role when the roles need to delegate authority to complete the same task. Permissions can change quickly through the association between the organization and the role, without the need for administrators to operate alone, and the introduction of the organization reduces the complexity of the authorization process.

The MATRBAC model integrates active and passive access control. The classification of tasks shows that the MATRBAC model is an active access control task for type A activity approval task and type W workflow task. Private tasks of class P and supervisory tasks of class S belong to passive access control. The MATRBAC model distinguishes between active and passive access control by the task.

The MATRBAC model integrates the determination of the user's trust value. Through the user's historical access record, the evaluation record of other resources to the user. And the user's access time interval to a specific resource, the user's trust value is calculated comprehensively. The user's trust value is an important reference index for assigning roles and tasks to users. As well as one of the indexes for task status risk assessment. This design makes the assignment of permissions more reasonable and reflects the dynamic assignment strategy of MATRBAC access control model to permissions.

4.4 The Comparison of the Model with Other Models

Table 4 provides a comprehensive comparison and analysis of the MATRBAC model introduced in this paper and the existing access control model, In the table, "$\surd$" means that the model proposed in the literature can meet the requirements of a security feature; " × " means that the model proposed in the literature cannot meet the requirements of a security feature.

As shown in Table 4, the access control model MATRBAC in this article provides better security features and more functional attributes compared to other relevant access control models. Therefore, the MATRBAC model is more applicable to the industrial control system environment.

5 Conclusions

This paper presents a MATRBAC model based on multi-attribute decision making. This model introduces the concept of organization and can adapt to the hierarchical structure of industrial control system and reduce the complexity of authorization effectively. Constraints such as subject attribute, operation attribute, resource attribute and environment attribute are added to enhance the ability of describing permissions. An algorithm for calculating user trust value by using user history access record is proposed to improve the model's adaptability and security to permissions. The multi-attribute decision making algorithm is applied to the task execution, and the risk value of the task is calculated in real time through the change of each attribute value during the task execution. By using user trust value judgment module and task supervision module, the authority can be dynamically adjusted during task execution, which can adapt to the special situation to achieve both coarse granularity and fine granularity. The experiment demonstrates the use of entropy weight TOPSIS method for multi-attribute decision making is more discriminative than that of simple weighting method and analytic hierarchy process, which can more intuitively show the difference between normal task and risk task, and demonstrate the availability of MATRBAC model for authority dynamic adjustment in industrial control system environment. Compared to other access control models, MATRBAC has more security features and extensibility.

Acknowledgments

This study was supported by the National Science Council of Taiwan under grant NSC 95-2416-H-159-003. The authors gratefully acknowledge the anonymous reviewers for their valuable comments.

References

[1] F. T. Alotaiby and J. X. Chen, "A model for team-based access control (TMAC 2004)," in *International Conference on Information Technology Coding and Computing*, pp. 450–454, Sep. 2004.

[2] S. Y. Belim and N. F. Bogachenko, "User authorization in a system with a role-based access control on the basis of the analytic hierarchy process," *Dynamics of Systems Mechanisms and Machines*, vol. 66, no. 08, pp. 1–5, 2017.

[3] N. Chistokletov and Y. Vavilin, "Safety management system of machine-building production," *Engineering Review*, vol. 38, no. 2, pp. 226–231, 2018.

[4] K. Haefner, B. Bezawada and I. Ray, "Securing home IoT environments with attribute-based access control," *Proceedings of the Third ACM Workshop on Attribute-Based Access Control*, vol. 86, no. 34, pp. 43–53, 2018.

[5] X. Hei, R. Akkaoui and C. Guo, "RBAC-HDE: On the design of a role-based access control with smart contract for healthcare data exchange," in *IEEE International Conference on Consumer Electronics - Taiwan (ICCE-TW'19)*, pp. 156–163, Sep. 2019.

[6] D. R. Kuhn, V. C. Hu and D. F. Ferraiolo, "Attribute-based access control," *Computer*, vol. 48, no. 2, pp. 85–88, 2015.

[7] E. S. Lee, H. S. Shih, H. J. Shyur, "An extension of topsis for group decision making," *Mathematical Computer Modelling*, vol. 5, no. 7, pp. 801–813, 2007.

[8] X. Lian, Y. Chen and D. Yu, "Exploring shodan from the perspective of industrial control systems," *IEEE Access*, vol. 99, no. 1, pp. 1–1, 2020.

[9] S. Long and L. Yan, "Racac: An approach toward rbac and abac combining access control," in *IEEE 5th International Conference on Computer and Communications (ICCC'19)*, pp. 16–25, Nov. 2019.

[10] S. Lu, F. A. Bhuyan and R. Reynolds, "A security framework for scientific workflow provenance access control policies," *IEEE Transactions on Services Computing*, vol. 99, no. 32, pp. 1–1, 2019.

[11] S. Mikko, D. Gregory and P. Seppo, "Toward a unified model of information security policy compliance," *MIS Quarterly*, vol. 42, no. 3, pp. 285–311, 2018.

[12] S. Oh and S. Park, "Task-role-based access control model," *Information Systems*, vol. 23, no. 6, pp. 533–562, 2003.

[13] B. Ou, S. Y. He and X. Liao, "Access control model of role matching in distributed workflow environment," *Computer Science*, vol. 45, no. 7, pp. 129–134, 2018.

[14] D. Servos and S. L. Osborn, "Current research and open problems in attribute-based access control," *ACM Computing Surveys*, vol. 49, no. 4, pp. 1–45, 2017.

[15] J. Shevchenko and David, "Joint management systems for operations and safety a routine-based perspective," *Journal of Cleaner Production*, vol. 194, no. 1, pp. 635–644, 2018.

Table 4: Comparison of security features

feature∖Model	$RBAC$[7]	$T-RBAC$[15,24]	$ABAC$[17,18]	$WBAC$[12]	$AMAC$[19]	$MATRBAC$
role concepts	√	√	√	√	√	√
task concepts	×	√	√	√	×	√
organizational concepts	×	×	×	×	×	√
Supporting environment	×	×	√	×	√	√
Support the tense	×	×	√	×	√	√
role units	√	√	√	√	√	√
task units	×	√	×	√	×	√
role inheritance	√	√	×	√	√	√
active access control	×	√	√	√	√	√
passive access control	√	√	×	√	√	√
separation of powers	√	√	√	√	√	√
dynamic authorization	×	√	√	√	√	√
Fine-grained access control	×	√	√	√	√	√

[16] Y. Song, Y. Peng and H. Ju, "OB4LAC: An organization-based access control model for e-government system," *Applied Mathematics Information Ences*, vol. 8, no. 3, pp. 1467–1474, 2014.

[17] R. K. Thomas, "Team-based access contural (TMAC): A primitive for applying role-based access controls in collaborative environments," in *International Conference on access control security*, pp. 13–19, Aug. 1997.

[18] Z. Thornton, T. H. Morris and I. Turnipseed, "Industrial control system simulation and data logging for intrusion detection system research," in *Annual Southeastern Cyber Security Summit*, pp. 1–14, June 2015.

[19] P. Wang and L. Y. Jiang, "Task-role-based access control model in smart health- care system," in *In MATEC Web of Conferences*, pp. 1–11, Mar. 2015.

[20] J. B. Xiong and Z. Q. Yaoand, J. F. Ma, "Behaviour-based multi-level access control for structured documents," *Journal of Computer Research and Development*, vol. 7, no. 4, pp. 53–62, 2013.

[21] H. Yang, A. Mohamed and M. Koien, "Access control model for cooperative healthcare environments: Modeling and verification," in *IEEE International Conference on Healthcare Informatics*, pp. 18–26, June 2016.

Biography

Dong Rui-hong biography. 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.

The Second Author biography. Xu Tongtong In 2017, Xu Tongtong obtained his bachelor of engineering degree from Xi'an University of Technology, he is studying for his masters degree at Lanzhou University of Technology. His research focuses on the industrial control network security.

The Last Author biography. It is required by the IJNS to have maResearcher/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.

Location Privacy Protection Algorithm Based on PageRank and Differential Privacy in Internet of Vehicles

Peng-Shou Xie, Xin Wang, Hao-Xuan Yang, Liang-Xuan Wang, Tao Feng, and Yan Yan
(Corresponding author: Xin Wang)

School of Computer and Communications, Lanzhou University of Technology, 287 Lan-gong-ping Road, Lanzhou,
Gansu 730050, China
Email: 415711979@qq.com

(Received Aug. 10, 2020; Revised and Accepted Apr. 24, 2021; First Online Oct. 19, 2021)

Abstract

The Internet of Vehicles location service has been widely used, but it still faces many security issues. Researchers have proposed some Location-Based Service privacy protection algorithms for IoV. However, These methods are easily attacked by attackers with background knowledge. Therefore, the privacy protection algorithm of the Internet of Vehicles based on differential privacy is improved by the PageRank algorithm and forms the PR-Diff algorithm. The algorithm first calculates the sensitive attribute value of the IoV user's location through the PageRank algorithm and then allocates a privacy budget. Then, different noise levels are added using the Laplacian mechanism according to the privacy budget of each area to protect the private information of the entire track. Finally, the availability and security of the PR-Diff algorithm are verified through experiments on real data sets. As a result, the algorithm can better resist the attacker's background knowledge attack.

Keywords: Differential Privacy; Internet of Vehicles; PageRank Algorithm; Privacy Protection; Trajectory Data

1 Introduction

The Internet of Vehicle (IoV) is an application of the Internet of Things in road traffic and is also an important part of the Intelligent Transport System (ITS) [15]. With the development of V2X(Vehicle to Everything, V2X) communication [10] and cloud computing, more and more security problems of Internet of Vehicles appear. The increasing number of Location Based Service (LBS) applications in the IoV facilitates users' lives. It also enables service providers to collect a large number of vehicle users' trajectory data. Trajectory data may expose users' private information to attackers, such as hobbies, social relations, physical conditions, *etc.* This brings serious threats to users' lives [7]. For example, through the analysis of a certain trajectory, the attacker may obtain the private information such as the home address, work location, and behavior patterns of the IoV users based on their background knowledge [12].Therefore, the protection of trajectory privacy has become a common concern of IoV users and researchers.

Differential privacy (DP) technology is supported by mathematical theories and has a strict mathematical definition of privacy protection. Its basic principle is to disturb the original data and network structure, including adding, deleting, exchanging, *etc.*, to make the disturbing data different from the original data, *i.e.*, protecting original data through publishing the disturbed data [8]. It has two main characteristics: Not being affected by the background knowledge of the attacker and not being affected by changes in a specific piece of data. Differential privacy technology has been widely used in location privacy protection since it was proposed. [22] proposed a location privacy method based on k-anonymity to prevent privacy disclosure in LBS constrained in incomplete data collection [22]. [18] combines the concepts of differential privacy and k-anonymity to propose the notion of differentially private k-anonymity (DPkA) for query privacy in LBS [18]. But adding too much noise may lead to poor data availability and large errors. Therefore, in response to the noise reduction problem, [16] proposed a differential privacy publishing method for trajectory data (CLM). A correlated Laplace mechanism is presented by CLM, which let Gauss noises pass through a specific filter to produce noise whose auto-correlation function was similar with original trajectory series [16]. [21] proposed a deferentially privacy location release mechanism (DPLRM) that considers the temporal correlation to protect the trajectory privacy of users. Specifically, It model the temporal correlation between user's true locations by Markov chain matrix, and define the DPLRM as an optimization problem by minimizing an objective function based on the total distance between the true location and possible released location [21].

Although differential privacy mechanism has been widely studied and applied as a random perturbation method. It was originally proposed to solve the privacy leakage problem of static data sets composed of mutually independent data. It protects privacy by adding independent noise to the data that needs to be protected. Therefore, the current trajectory publishing method based on the differential privacy mechanism is to process the trajectory data as an independent sequence. It protects privacy by adding an independent noise sequence to the trajectory sequence that needs to be released. However, when the location data is not independent (relevant), the independent noise methods will still leak privacy. The following problems are faced in the existing trajectory differential privacy protection methods.

1) Privacy budget allocation. Differential privacy is a strictly defined privacy protection model, which is independent of any background knowledge of the attacker. The attack capability from differential privacy protection can be measured by the privacy budget ε. The smaller the ε, the greater the noise added, and the greater the degree of privacy protection provided by the algorithm. However, the unreasonable allocation of the privacy budget ε will destroy the privacy protection performance of the algorithm. Then the algorithm loses its meaning.

2) Low data availability. The trajectory privacy protection method based on differential privacy is mainly realized by adding disturbing noise, so that the attacker cannot identify the real information of the user from the released trajectory data. Due to the correlation of trajectory data, the weight of the sensitivity of differential privacy will increase. In order to achieve the same privacy protection strength, more noise needs to be added to the trajectory data, but more noise will reduce the usability of the data.

Based on the shortcomings of the existing differential privacy protection algorithm for the IoV trajectory, a PR-Diff algorithm that uses PageRank algorithm to improve the location privacy protection algorithm of IoV based on differential privacy is proposed in this paper. The algorithm first uses background knowledge [24] to predict the current trajectory position [3]. The sensitive attribute value of the IoV user location is calculated through the PageRank algorithm, and the privacy budget is allocated through the sensitive attribute value. Then Laplace noise is added through the privacy budget of each area. Finally, the purpose of protecting the privacy information of the entire track is realized. The difference between the work done in this paper and the current research is: First of all, In order to allocate the privacy budget reasonably, this paper uses the PageRank algorithm to calculate the sensitive attribute value of the position on the track.Then allocate the privacy budget according to the sensitive attribute value. Secondly, in order to improve data availability, the PR-Diff algorithm replaces protecting the entire trajectory by protecting the position points on the user's trajectory, and adds noise of different intensities to pass different privacy budget values. The PR-Diff algorithm improves data utilization while protecting information from being leaked.

2 Theoretical Basis and Definition

2.1 PageRank Algorithm

The basic idea of PageRank algorithm: If there is a link to web page A on web page T, it means that the owner of T considers A to be more important, and thus assigns a part of the importance score of T to A. The importance score value is: PR(T)/C(T). Where PR(T) is the PageRank value of T and C(T) is the number of outgoing links of T, then the PageRank value of A is the accumulation of a series of T's page importance scores [14].

2.2 Differential Privacy

The basic idea of differential privacy: Any two data sets P and Q, $P = d_1, d_2, \cdots, d_{i-1}, d_i, d_{i+1}, \cdots, d_n$,$Q = d_1, d_2, \cdots, d_{i-1}, d_i, d_{i+1}, \cdots, d_n$, If and only if one data record di is different between them, then P and Q are called neighbor data sets.For a given privacy query function f, the query output results on any two neighboring data sets P and Q represent $f(P)$, $f(Q)$, the set of output results is expressed as 0, and the probability that the output result is 0 is expressed as Pr. If and only if the function f satisfies the condition of Equation (1), it satisfies ε-differential privacy [20].

$$\frac{Pr[f(P) \in O]}{Pr[f(Q) \in O]} \leq exp^{\varepsilon} \tag{1}$$

Among them, ε is a differential privacy budget parameter, which represents the strength of privacy protection.

2.3 Global Sensitivity

For the function $f : D \rightarrow R^d$ satisfying Equation (2), then Δf is called the global sensitivity [4], which is the maximum range of the output value change of the specific query function f when querying on all possible adjacent data sets D and $D^{'}$. The measurement is the distance between the two. It has nothing to do with the data set D and is determined by the query function f. The global sensitivity directly affects how much noise is added. The greater the global sensitivity of f, the more noise that needs to be injected under the same conditions of ε. And replaced by smaller local sensitivity in special scenes.

$$\Delta f = max_{D,D'} ||f(D) - f(D')||_1 \tag{2}$$

2.4 Laplace Mechanism

For the query function $f : D \to R^d$, if the output result of the function f satisfies the Equation (3), then f satisfies the ε-differential privacy protection [13].

$$K(D) = f(D) + Lap(\lambda). \tag{3}$$

Among them, $Lap(\lambda)$ is the noise following the Laplace distribution,λ is the scale parameter and the calculation of λ is shown in Equation (4).

$$\lambda = \frac{\Delta f}{\varepsilon} \tag{4}$$

3 Differential Privacy Budget Allocation

3.1 Problem Description

The privacy budget allocation strategy of differential privacy currently only adapts to specific spatial data applications. It cannot choose an appropriate allocation method according to the data privacy protection requirements, and therefore cannot reasonably allocate the privacy budget. Therefore, the location nodes on the trajectory in the PR-Diff algorithm use the PageRank algorithm to allocate the privacy budget, which can further improve the privacy protection effect.

3.2 PageRank Value of Each Position Point is Obtained

PageRank is an algorithm for evaluating the quality of a web page, which can be considered as a probability representation. The main ideas of PageRank algorithm are as follows: If the number of links to a web page is higher, the importance of this page is higher; The higher the quality of the page linked to the webpage, the higher the importance of the page. The PageRank method can be applied to any entity set with mutual reference characteristics. It can be used to sort the importance of location nodes in the road network environment.

Based on the above ideas, the basic steps for calculating the sensitivity of a location point are as follows:
In this paper, the sensitive attribute value of the IoV user's location is calculated using the PageRank algorithm.Intuitively speaking, if a location and its neighboring location have high sensitivity, the location is more sensitive. The algorithm calculation is shown in Equation (5).

$$PR(L_i) = \frac{(1-d)}{N} + d \sum_{L_j \in M(L_i)} \frac{PR(L_j)}{Out(L_j)} \tag{5}$$

Where $PR(L_i)$is the sensitive attribute value of the position to be evaluated L_i,N is all location nodes,$M(L_i)$ is the set of positions where L_i has a chain, $Out(L_j)$ is the number of outgoing chains at position L_j, d is the damping factor, $1 - d$ can be understood as the probability of randomly jumping to other positions.

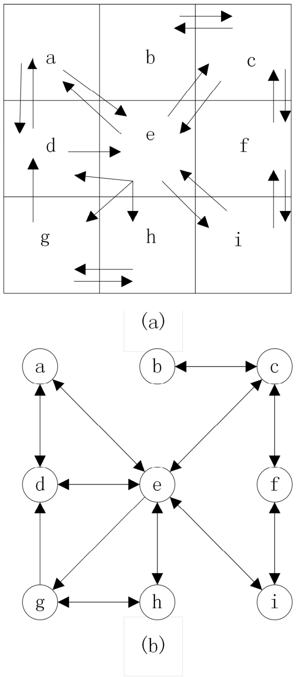


Figure 1: Vehicle inflow and outflow in the adjacent area

3.3 Mesh

Division is an effective ways to protect location information. Based on the background knowledge [6], this paper divides the IoV user activity area into grids. The location nodes that the IoV user may reach can be inferred based on background knowledge, and then these location nodes are formed into a directed graph. Figure 1 shows the inflow and outflow diagram of vehicles in the adjacent area.Then calculate the sensitivity of each location point [17].

3.4 Define Adjacency Matrix A

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots A_{1n} \\ A_{21} & A_{22} & \cdots A_{2n} \\ \vdots & \vdots & \vdots \\ A_{n1} & A_{n2} & \cdots A_{nn} \end{bmatrix} \tag{6}$$

$$A_{ij} = \begin{cases} \frac{1}{2} T_{ij}, i = j \\ T_{ij}, i \neq j \end{cases} \tag{7}$$

Where A_{ij} represents the weight of the edge of the super location point, that is, A_{ij} is the sum of the edge weights of all the pair of location nodes connecting the area i to the area j, so A_{ij} is the total number of edges of the two regions.T_{ij} is the total number of connected edges of area i and area j. When $i = j$, it is the same area.as shown in Equation (6) and Equation (7).

3.5 Calculate the Structural Coefficient of Each Area

The influence of the region on the importance of the location point is measured according to the structural characteristics of the location and the edge of the other area and the number of edges in the other region.If the number of internal edges of the location point is large, the user tends to move within the area. If there are a large number of external connections at a location, users tend to move to other communities. Therefore, this paper regards the connection between the interior and the region as the regional structure coefficient.The specific calculation method is shown in Equation (8) and Equation (9).

$$K_{in}^{n} = \frac{1}{2}\sum_{i} A_{ij}\delta_n(C_i, C_j), \tag{8}$$

$$K_{out}^{n} = \sum_{i} A_{ij}\sigma_n(C_i, C_j). \tag{9}$$

When $i = j$,$\delta_n(C_i), C_j) = 1$; When $i \neq j$,$\delta_n(C_i), C_j) = 0$; When $i \neq j$,$\sigma_n(C_i), C_j) = 1$; When $i = j$,$\sigma_n(C_i), C_j) = 0$. Where K_{in}^{n} is the sum of the internal degrees of the region C_n,K_{out}^{n} is the sum of the external degrees of the region C_n, simplify the above formula, they have Equation (10) and Equation (11) for the adjacency matrix A.

$$K_{in}^{n} = A_{nn} \tag{10}$$

$$K_{out}^{n} = -A_{nn} + \sum_{n} A_{nn} \tag{11}$$

3.6 Calculate Area Structure Coefficient Vector

K_{out}^{n} is actually the value obtained by separately summing each row of the adjacency matrix and subtracting the internal degree.The meaning it represents is the total number of outer edges of all connected areas C_n. Calculate the internal degree and external degree of the area C_n separately, and calculate the area structure coefficient vector $I = \left(I^{(1)}, I^{(2)}, \ldots, I^{(n)}\right)^T$, where each component $I^{(n)}$ is the area structure coefficient of each area C_n, which can be calculated according to Equation (12).

$$I^{(n)} = \frac{K_{in}^{n}}{K_{out}^{n} + K_{in}^{n}} InK_{out}^{n} - \frac{K_{out}^{n}}{K_{out}^{n} + K_{in}^{n}} InK_{in}^{n} \tag{12}$$

The vector I is calculated by the number of internal and external connections in an area, so the area I can indicate the tightness of the internal and external connections of an area and the flow of messages inside and outside the area, thereby measuring the connectivity of a certain area.The internal and external connections of the area will affect the transmission capacity of an area.Therefore, such a regional structure coefficient can be used to measure the status of a region in the map.

3.7 Privacy Level Value

The PageRank value of each location point is normalized, and then the sensitivity of the location point is calculated according to Equation (13).

$$P_T = I^{(n)} + \frac{\sum_N PR(L_i)}{N} \tag{13}$$

Which is expressed as the sensitivity of the location point.

3.8 Differential Privacy Budget Parameters

Because of the subjectivity of privacy, it is unreasonable to make equal disturbances at every location. Not only the sensitivity of different locations is different, but also the sensitivity of different users to the same location is different [23]. Therefore, the unreasonable allocation of the privacy budget ε will destroy the privacy protection performance, and the algorithm will lose its meaning.Therefore, this paper introduces PageRank algorithm to improve the utilization of privacy budget.For locations with higher sensitivity, a smaller privacy budget is allocated, that is, greater disturbance noise is added to obtain a higher degree of privacy protection.

ε is a differential privacy budget parameter, which represents the strength of privacy protection. The smaller the value of ε, the more noise is added, the higher the degree of privacy protection, and the lower the data utilization rate.When it is close to 0, the result output of the query function on the two data sets is basically the same, and any position information of the data set will not be disclosed at this time.The privacy budget ε is a data volume drawn up based on the differential privacy model to indicate the level or degree of privacy protection. The ε setting of each position is different, the greater the sensitivity, the smaller the ε setting [11] .

Equation (14) was used to calculate differential privacy budget parameter.

$$\varepsilon_i = \frac{\varepsilon}{P_T} \tag{14}$$

Where ε is the total privacy budget; ε_i is the privacy budget of area i.

Different data applications require different privacy budget allocation methods. However, traditional privacy budget allocation strategies cannot meet the need of users for flexible allocation of privacy budgets. In response to this problem, the allocation strategy of using PageRank algorithm to allocate privacy budget is proposed in this chapter. IoV users calculate the PageRank value of each location point, and then classify the location nodes to calculate the structural coefficient of each area, thereby comprehensively calculating the sensitivity of the location point. Finally, a reasonable allocation of privacy budget takes advantage of the sensitivity of location nodes [9].

4 PR-Diff Algorithm

4.1 Problem Statement

In the existing research, in the stage of adding noise, the addition of noise values of uniform scales is likely to cause excessive relative errors, causing the published data values to seriously deviate from the original data values, thereby greatly reducing the usefulness of the data.The goal of this paper is to provide a practical solution that can provide varying degrees of privacy protection for the location nodes on the vehicle network track.At the same time, in order to ensure that the trajectory data of Internet of Vehicles users meet the needs of differential privacy, the Laplace algorithm is used in this paper to add noise [2].

4.2 Noise Addition

The differential privacy protection method adds a certain degree of privacy protection noise to the output result through the privacy budget parameter ε to ensure that the output result on the neighboring data set is indistinguishable within a certain probability range. This reduces the probability of an attacker identifying the user's real information. At present, privacy protection is mainly achieved by adding Laplace noise or exponential noise.The Laplacian mechanism adds random noise that obeys the Laplacian distribution to the ε-differential privacy. It is suitable for numerical data protection.The exponential mechanism controls the output probability of each candidate based on a scoring function, which is suitable for discrete data protection [1].In this paper. uses the Laplacian plus noise mechanism.

The Laplacian noise is added according to the average budget of each location point. If the noise is too large, it will introduce too much noise and affect the availability of published data. If the noise is too small, the disturbance error is reduced, but the non-uniformity error increases, which exposes the distribution of actual data and cannot guarantee privacy [5].Therefore, this paper adds different Laplacian noises according to different privacy budgets, which can hide the distribution of actual data, thereby improving the degree of privacy protection of the algorithm.

Laplace noise generation method: The Laplace mechanism realizes differential privacy protection by adding noise that obeys the Laplace distribution to the query results [19].The extension of the mechanism to two-dimensional space is still applicable. The noise it adds satisfies the following probability density function, as shown in Equation (15).

$$P(x) = \frac{1}{2\lambda} exp(-\frac{|x-u|}{\lambda}) \tag{15}$$

Where u is the expected value. In order to meet the probability function of the noise that needs to be added to the position differential privacy, we give a privacy protection budget ε and the actual position x_0. If the position x after adding noise satisfies the distribution of Equation (16), then ε-position differential privacy protection is satisfied.

$$D(x_0)(x) = \frac{\varepsilon^2}{2\pi} exp(-\varepsilon d(x_0, x)). \tag{16}$$

As can be seen from the probability distribution function, When $d(x_0, x) > 0$, the probability of x decreases as the distance between x and x_0 increases,And the probability distribution is only related to the distance from x to x_0. In order to simplify the implementation, the function is transformed into a polar coordinate function, as shown in Equation (17).

$$D_{\varepsilon,x_0}(r,\theta) = \frac{\varepsilon^2}{2\pi} r \cdot exp(-\varepsilon r). \tag{17}$$

This function is the distribution function of x in polar coordinates, Where r represents the distance from x to x_0, θ is the angle between x and the polar axis in polar coordinates,In order to facilitate the solution, Equation (17) is decomposed into radius r and angle θ.As shown in Equation (18) and Equation (19).

$$D_{\varepsilon,R}(r) = \int_0^{2\pi} D_\varepsilon(r,\theta)d\theta = \varepsilon^2 r e^{-\varepsilon r} \tag{18}$$

$$D_{\varepsilon,\theta}(\theta) = \int_0^{+\infty} D_\varepsilon(r,\theta)dr = \frac{1}{2\pi} \tag{19}$$

According to the above formula, random noise with radius r and angle θ can be added to $x_0 = (s,t)$ to generate $x'_0 = (s + r\cos\theta, t + r\sin\theta)$, so as to achieve the purpose of privacy protection by position difference.

Integrating Equation (19) in $[0, +\infty)$ can get the cumulative probability distribution $C_{\varepsilon_{(r)}}$ of the distance r, and then $r = C_z^{-1}(z)$. As shown in Equation (20).

$$C_\varepsilon(r) = \int_0^r \varepsilon^2 \rho exp(-\varepsilon\rho)d\rho = 1-(1+\varepsilon r)exp(-\varepsilon r). \tag{20}$$

Among them,$C_z(r) \in [0,1)$, θ is a random number between $[0, 2\pi)$.

4.3 PR-Diff Algorithm

The PR-Diff algorithm is shown in Algorithm 1.

Algorithm 1 PR-Diff algorithm

```
Begin
Input:location point;$\varepsilon$
Output:$\varepsilon_i$
Calculate the PR value of each location point;
Divide the map area;
Construct the adjacency matrix A;
Calculate the structural coefficient of each area;
Calculate the privacy level value of each area;
Calculate the corresponding differential privacy budget $\varepsilon_i$ according to $\varepsilon$;
output $\varepsilon_i$;
Add Laplacian noise;
Complete location privacy protection.
```

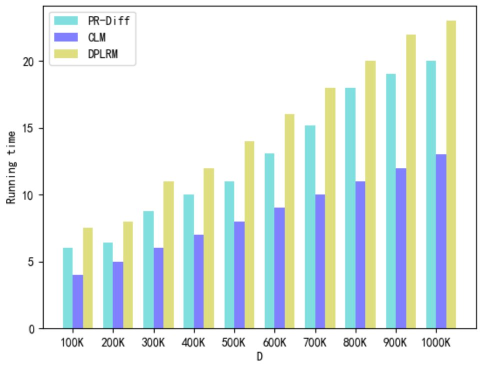


Figure 2: Algorithm performance efficiency analysis

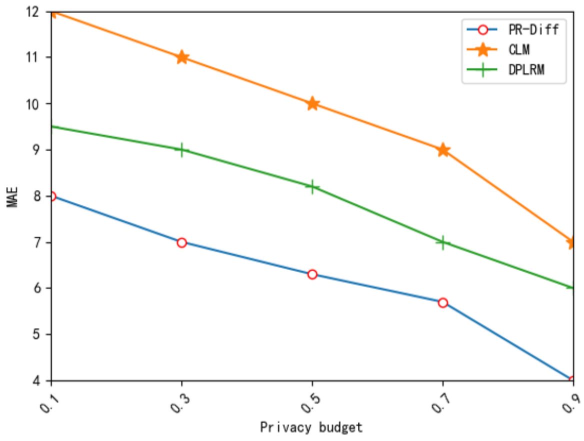


Figure 3: MSE changes when ε changes

5 Experimental Evaluation

5.1 Experimental Design and Parameter Setting

In order to verify the usability and safety of the method proposed in this paper, the experimental design is divided into two levels: horizontal and vertical. In the horizontal aspect, compare the method proposed in this paper with other methods.Set the same environment and the same parameters, compare the performance of the three methods in terms of execution efficiency, data availability and security, and analyze their advantages and disadvantages. In the vertical aspect, different parameters are set to verify the method proposed in this paper, and the performance of the method proposed in this paper is tested under different parameters to find the best performance.

All data sets for this experiment come from Microsoft Research Asia.The data set used contains 17,621 paths, with a total distance of 1,293,951 kilometers, 50,176 hours, and one sampling every 1-1.5 seconds or every 5-10 meters. This experiment uses python programming.

5.2 Perform Efficiency Analysis

In this paper, execution efficiency mainly refers to the time required for privacy protection processing. We compare the PR-Diff algorithm with the CLM algorithm and the DPLRM algorithm.

It can be seen from the experimental figure 2 that when the data set continues to increase, the running time of the three algorithms on the data set increases accordingly. This is because when k is large, the geographical space experienced by the algorithm increases, so the time required for the algorithm also increases.The execution time of the PR-Diff algorithm and the DPLRM algorithm is much longer than that of the CLM algorithm. This is because the method adds different degrees of noise parameters according to the different sensitivity of the position points on the trajectory, carries out different intensity disturbances, provides higher data availability, and therefore the algorithm takes a long time to execute.

5.3 Data Availability

The degree of accuracy can be measured by Mean Absolute Error (MAE), which is defined as shown in Equation (21).

$$MAE = \frac{1}{L}\sum_{x\in X}|x - x_0| \qquad (21)$$

Among them, L represents the length of the trajectory sequence,x is the position point after adding noise,x_0 is the original position point.The smaller the MAE, the higher the data availability.Evaluate the MAE in the PR-Diff algorithm, CLM algorithm and algorithm, and DPLRM algorithm, as shown in Figure 3.

The experimental results show that the PR-Diff algorithm is better than the other two algorithms in terms of data availability. This is because in the CLM method, all position points on the trajectory are added with the same intensity of noise information, and the entire trajectory is disturbed to the same degree, so the information loss is relatively large. Although the DPLRM method also adds different noises, But it not only considers the privacy impact of the current release location on the current moment, but also considers the privacy impact of the current release location on the previous release location. The algorithm has many limiting factors and low availability. The PR-Diff algorithm uses the PageRank algorithm to allocate a privacy budget. Under the same privacy protection strength ε, a smaller noise algorithm is needed to achieve the same privacy protection strength. Therefore, PR-Diff has higher data availability.

5.4 Security Analysis

Security mainly refers to the probability that the published trajectory data can be identified by an at-

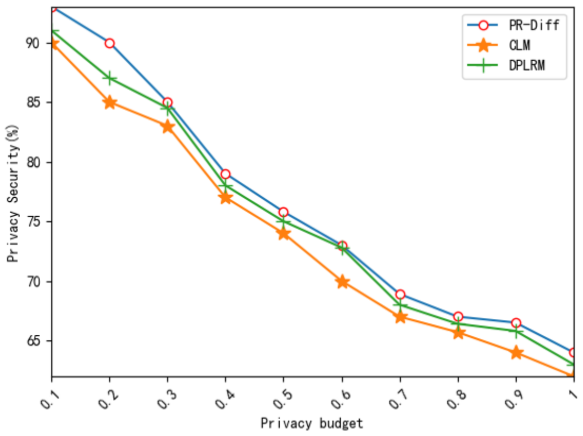


Figure 4: Analysis of privacy protection effect of three algorithms

tacker.According to the characteristics of the differential privacy protection method, the actual differential privacy budget ε directly reflects the level of privacy protection achieved.The smaller the value of ε, the more noise is added, the higher the level of privacy protection achieved, and the better the security of publishing trajectory data.

In differential privacy, ε is used as a key parameter to determine the privacy strength.According to research, when $\varepsilon = 1$ or less, data availability can reach a more appropriate level.Therefore, this paper evaluates the security of the three algorithms with ε of 0.1 1.

It can be seen from Figure 4 that as the privacy protection budget value ε increases, the level of privacy protection achieved decreases.And it can be seen that the privacy levels of the three algorithms are similar, and the PR-Diff algorithm is slightly higher than the other two algorithms. Because the three methods consider the timing correlation between the added noise sequence and the original trajectory sequence when adding noise, to a certain extent, the original trajectory sequence and the added noise sequence become unrecognizable. However, the PR-Diff algorithm uses the PageRank algorithm to calculate the differential privacy budget parameters, and the allocation of privacy parameters is more scientific, so the privacy protection effect achieved is better.

5.5 The Influence of ε on the PR-Diff Algorithm

Figure 5 compares the distribution of r under different privacy protection degrees ε. According to formula 18, the relationship between the distance r and the probability density can be seen. As r increases, the probability density first increases and then decreases.This is because within a certain radius, the larger the ε, the smaller the added noise and the smaller the degree of privacy. As the radius increases, the change of ε is not obvious. Therefore, in this paper, the cumulative probability distribution $C_\varepsilon(r)$ under different ε is compared.

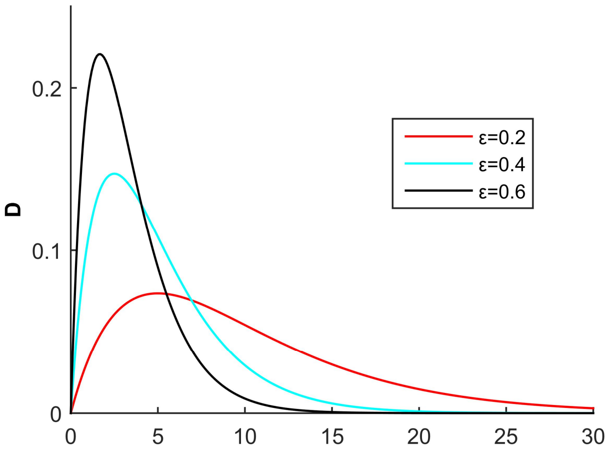


Figure 5: Comparison of R distribution under different

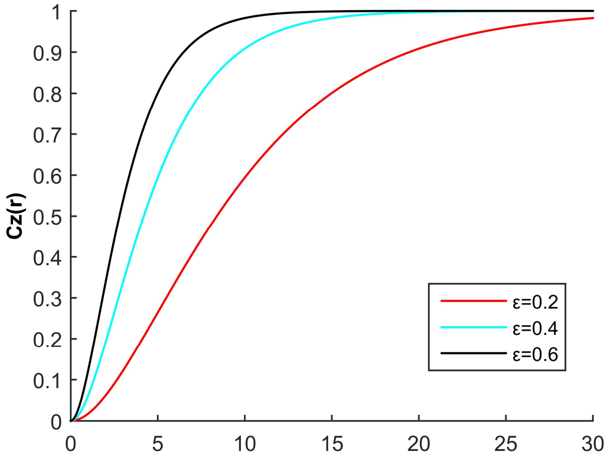


Figure 6: Comparison of cumulative probability distribution $C_\varepsilon(r)$ with different values ε

From the cumulative distribution in Figure 6,The added noise has a negative correlation with the privacy budget ε value. The smaller the ε, the higher the degree of privacy protection, and the more noise that needs to be added. It can be seen that there is a mutual restriction between the degree of privacy protection and the efficiency of location services. If you want to have a high degree of privacy protection, you must sacrifice the accuracy of the location. To ensure the accuracy of the location will cause the weakening of privacy protection.

6 Conclusion

Location privacy protection in the Internet of Vehicles has always been a research hotspot in privacy protection.Aiming at the problem of reducing data availability by protecting all points in the trajectory in existing methods, this paper proposes the PR-Diff algorithm.The PR-Diff algorithm uses background knowledge to predict the current location. The location privacy budget of the IoV user is allocated through the PageRank algorithm, and different degrees of Laplacian noise are added according to the privacy budget. Through experiments on real trajectory data sets, the usability of this method is verified. The experimental results show that compared with the existing trajectory data privacy protection methods, the method in this paper improves data availability and safety under the premise of protecting trajectory data privacy.

The disadvantage of this paper is that PR-Diff has a long running time, which will be the next problem to be solved. With the rapid growth of data in the era of big data, current research on IoV privacy protection methods can no longer keep up with the speed of data growth. Unable to meet people's urgent needs for efficiency and privacy in the information age In future research, we will consider continuing to optimize the PR-Diff algorithm to improve the availability of location data and reduce the running time of the algorithm, and to better extend it to real-time IoV location services.

Acknowledgement

This research is supported by the National Natural Science Foundations of China under Grants No.61862040, No.61762059 and No. 61762060. The authors gratefully acknow-ledge the anonymous reviewers for their helpful comments and suggestions.

References

[1] S. J. Cai, X. Lyu, and D. H. Ban, "Spatial statistic data release based on differential privacy," *KSII Transactions on Internet And Information Systems*, vol. 13, no. 10, pp. 5244–5259, 2019.

[2] D. G. Feng, M. Zhang, and Y. T. Ye, "Research on location trajectory publishing technology based on differential privacy model," *Journal of Electronics and Information Technology*, vol. 42, no. 01, pp. 74–88, 2020.

[3] Z. Q. Gao, Y. X. Sun, X. L. Cui, Y. T. Wang, Y. Y. Duan, and X. A. Wang, "Privacy - preserving hybrid K-means," *International Journal of Data Warehousing and Mining (IJDWM'18)*, vol. 14, no. 2, pp. 1–17, 2018.

[4] Z. Q. Gao, Y. T. Wang, Y. Duan, and *et al.*, "Multi-level privacy preserving data publishing," *International Journal of Innovative Computing and Applications*, vol. 9, no. 2, pp. 66–76, 2018.

[5] Q. Y. Li, H. Wu, X. Wu, and L. Dong, "Multi-level location privacy protection based on differential privacy strategy in vanets," in *IEEE The 89th Vehicular Technology Conference (VTC2019-Spring)*, Apr. 2019. DOI: 10.1109/VTCSpring.2019.8746396.

[6] W. H. Li, J. Cao, and H. Li, "Privacy protection scheme based on self-correlation of location-based service privacy," *Journal of Communications*, vol. 40, no. 5, pp. 57–66, 2019.

[7] Y. H. Li, X. Cao, Y. Yuan Y, and *et al.*, "Privsem: Differential location privacy using semantic and differential privacy," *World Wide Web-Internet and Web Information Systems*, vol. 22, no. 6, pp. 2407–2436, 2019.

[8] J. Liu and F. L. Qin, "Protection of user data by differential privacy algorithms," *International Journal of Network Security*, vol. 22, no. 5, pp. 838–844, 2020.

[9] X. D. Ma, J. F. Ma, H. Li, and *et al.*, "AGENT: An adaptive geo-indistinguishable mechanism for continuous location-based service," *Peer to Peer Networking and Applications*, vol. 11, no. 3, pp. 473–485, 2018.

[10] Z. MacHardy, A. Khan, K. Obana, and S. Iwashina, "V2X access technologies: Regulation, research, and remaining challenge," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 3, pp. 1858–1877, 2018.

[11] T. Shang, Z. Zhao, W. W. Shu, and J. W. Liu, "Big data decision tree algorithm based on arithmetic privacy budget allocation," *Engineering Science and Technology*, vol. 51, no. 02, pp. 130–136, 2019.

[12] J. Son, D. Kim, M. Z. A. Bhuiyan, R. Tashakkori, J. Seo, and D. H. Lee, "Privacy enhanced location sharing for mobile online social networks," *IEEE Transactions on Sustainable Computing*, vol. 5, no. 2, pp. 279–290, 2020.

[13] X. X. Tian, Q. Song, and F. L. Tian, "Multidimensional data aggregation scheme for smart grid with differential privacy," *International Journal of Network Security*, vol. 20, no. 6, pp. 1137–1148, 2018.

[14] B. Tidke, R. Mehta, and J. Dhanani, "Mehtimodal ensemble approach to identify and rank top-knodes of scholarly literature using twitter network," *Journal Of Information Science*, vol. 46, no. 4, pp. 437–458, 2020.

[15] C. Wang, Z. Dai, and D.Zhao, "A novel identity-based authentication scheme for IoV security," *International Journal of Network Security*, vol. 22, no. 4, pp. 627–637, 2020.

[16] H. Wang, Z. Q. Xu, L. Z. Xiong, and *et al.*, "CLM: Differential privacy protection method for trajectory release," *Journal of Communications*, vol. 38, no. 06, pp. 485–96, 2017.

[17] J. Wang, R. Zhu, S. Liu, and *et al.*, "Node location privacy protection based on differentially private grids in industrial wireless sensor networks," *Sensors*, vol. 18, no. 2, pp. 410–424, 2018.

[18] J. B. Wang, Z. P. Cai, Y. S. Li, and *et al.*, "Differentially private K-privacy with privacy in location-based services," *Personal and Ubiquitous Computing*, vol. 22, no. 3, pp. 453–469, 2018.

[19] T. Wang, Z. G. Zheng, and M.Elhoseny, "Equivalent mechanism: Distributed location data with errors through differential privacy," *Future Generation*

Computer Systems - The International Journal Of Escience, vol. 928, pp. 600–608, 2019.

[20] Y. J. Wang, Z. P. Cai, X. R. Tong, and *et al.*, "Truthful incentive mechanism with location privacy- preserving for mobile crowdsourcing systems," *Computer Networks*, vol. 135, pp. 32–43, 2018.

[21] Y. C. Wu, H. Chen, and S. Y. Zhao, "A differential privacy trajectory protection mechanism based on space-time correlation," *Chinese Journal of Computing Technology*, vol. 41, no. 2, pp. 309–322, 2018.

[22] X. D. Yang, L. Gao, J. Zheng, and W. Wei, "Location privacy preservation mechanism for location-based service with incomplete location data," *IEEE Access*, vol. 8, pp. 95843–95854, 2020.

[23] A. Y. Ye, L. Y. Meng, Z. W. Zhao, and Y. Q. Diao, "Trajectory differential privacy protection mechanism based on prediction and sliding window," *Journal of Communications*, vol. 41, no. 04, pp. 123–133, 2020.

[24] P. Zhang, C. Hu, D. Chen, and *et al.*, "Shiftroute: Achieving location privacy for map services on surveys," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 5, pp. 4527–4538, 2018.

Biography

Peng-shou Xie was born in Jan. 1972. He is a professor and a supervisor of master student at Lanzhou University of Technology. His major research field is Security on Internet of Things. E-mail: xiepsh_lut@163.com

Xin Wang was born in June 1995. She is a master student at Lanzhou University of Technology. Her major research field is Security on Internet of Things. E-mail: 415711979@qq.com

Hao-xuan Yang was born in June 1996. He is a master student at Lanzhou University of Technology. His major research field is network and information security. E-mail: 1540901751@qq.com

Liang-xuan Wang was born in June 1995. She is a master student at Lanzhou University of Technology. Her major research field is network and information security. E-mail: 2681559167@qq.com

Tao Feng was born in Dec. 1970. He is a professor and a supervisor of Doctoral student at Lanzhou University of Technology. His major research field is modern cryptography theory, network and information security technology. E-mail: fengt@lut.cn

Yan Yan was born in Oct. 1980. She is a associate professor and a supervisor of master student at Lanzhou University of Technology. Her major research field is privacy protection, multimedia information security. E-mail: yanyan@lut.cn

Universal Signature Translators

Fei Tang[1], Dong Huang[2], Fuqun Wang[3,4], and Zhenyu Chen[5]
(Corresponding author: Fei Tang)

School of Cyber Security and Information Law, Chongqing University of Posts and Telecommunications[1]
Chongqing, 400065, China
Chongqing Vocational and Technical University of Mechatronics, Chongqing 402760, China[2]
Department of Mathematics, Hangzhou Normal University, Hangzhou 311121, China[3]
Westone Cryptologic Research Center, Beijing 100071, China[4]
Chongqing Institute of Quality and Standardization, Chongqing 400023, China[5]
Email: tangfei@cqupt.edu.cn

(Received Aug. 11, 2020; Revised and Accepted Apr. 24, 2021; First Online Oct. 19, 2021)

Abstract

This work introduces the notion of Universal Signature Translators (UST). In a universal signature translator system, Bob, using a set of public parameters and some other supporting information, can create a signature translator who can translate Alice's signatures to his, where Bob chose Alice. In prior proxy re-signature schemes, signatures can only be translated if Alice and Bob use the same singing algorithm and shared parameters. A universal signature translator can translate any signature to another. Therefore, universal signature translators provide a novel opportunity for signature translation. In addition, we also define a weak version of UST called a semi-functional universal signature translator scheme which can translate any signatures to a fixed type of signatures. We give a concrete construction and security-proof semi-functional UST scheme based on indistinguishability obfuscation and some other supporting primitives.

Keywords: Indistinguishability Obfuscation; Proxy Re-signatures; Universal Signature Translators

1 Introduction

The notion of proxy re-signatures was introduced by Blaze, Bleumer and Strauss [2] and formalized later by Ateniese and Hohenberger [1]. In a proxy re-signature scheme, a semi-trusted proxy is given some information that allows it to transform Alice's (*i.e.*, delegatee) signature on a message into Bob's (*i.e.*, delegator) signature on the same message. The security of proxy re-signature schemes requires that the proxy cannot, on his own, produce arbitrary signatures on behalf of either Alice or Bob. As shown by [1, 2, 4] *et al.*, proxy re-signatures are a very useful tool for simplifying key management, sharing web certificates, authenticating network path, managing group signatures, and so on.

After the works of [1] and [2], many proxy re-signature schemes have been proposed, *e.g.*, multi-use unidirectional proxy re-signature schemes [12, 16, 20], proxy re-signature schemes in the standard model [4, 15], identity-based proxy re-signature schemes [11, 17, 19], threshold proxy re-signature schemes [21], blind proxy re-signature schemes [5, 9], certificateless proxy re-signature schemes [8], and so on. However, all of these schemes have one thing in common that they require Alice and Bob to adopt a common signature scheme and shared parameters. In practice, the common signature scheme and parameter requirements can be a large barrier to adoption. Existing users will already have established signing keys and algorithms which are hard to change. In addition, even if a user moved from a signature system to a proxy re-signature system, all previously generated signatures can not be re-signed.

To address the above-described limitation, Ateniese and Hohenberger [1] proposed an open problem: **"Whether or not proxy re-signature schemes can be built that translate from one type of signature scheme to another?"** As far as we know, [14] and [22] are the only two works to toward this problem, where [14] constructed a proxy re-signature scheme which can translate Alice's Schnorr/ElGamal signature to Bob's RSA signature and [22] constructed a proxy re-signature scheme which can translate Alice's certificate-based signature to Bob's identity-based signature. However, all schemes in [14] and [22] do not have formal security proofs.

1.1 Universal Signature Translators

In order to solve Ateniese and Hohenberger's [1] open problem, in this paper, we introduce the notion of universal signature translators (UST). In a universal signature translator system, the delegator Bob can create a signature translator $rk_{A \to B}$, *i.e.*, the proxy re-signing key, on inputs $(params, \mathsf{Sign}_B, sk_B, \mathsf{Verify}_A, vk_A)$, where $params$ is the system public parameters, Sign_B is Bob's signing algorithm, sk_B is Bob's secret key, Verify_A is Alice's verification algorithm, and vk_A is Alice's verification key, respectively. Then the translator will be given to a proxy who can translate Alice's signature $\sigma_A(m)$ on message m into Bob's signature $\sigma_B(m)$ on the same message m. In universal signature translator systems, both signatures $\sigma_A(m)$ and $\sigma_B(m)$ can be produced from any type of signing algorithms (subject to a presupposed length constraint). We call the signature schemes of Alice and Bob's are *base* scheme and *target* scheme, respectively. As in the prior proxy re-signature schemes, the delegatees of the system need not do anything special to allow re-signing; Indeed they could be unaware of the existence of such a system. The notion of universal signature translators is inspired by the notion of universal signature aggregators [10] which can aggregate a collection of signatures produced from any set of signing algorithms into one short signature.

The central challenge of proxy re-signatures is to design a way to translate a signature $\sigma_A(m)$ into another $\sigma_B(m)$ without secret key sk_B. In prior proxy re-signature schemes, it is required that two signatures reside in a common group, *e.g.*, bilinear group. However, in universal signature translators, the two signatures may be created from two different signature schemes, and thus they do not reside in a common group. It is seemingly difficult that realize universal signature translators based on classical cryptographic tools.

1.2 Our Solution

Our solution will be to overcome the above limitation by using the tool of program obfuscation [3]. At the highest level, Bob will produce an obfuscated program, *i.e.*, the translator $rk_{A\to B}$ which contains the description of $(\mathsf{Sign}_B, sk_B,\ \mathsf{Verify}_A, sk_A)$. Next, proxy takes as input Alice's signature σ_A and corresponding message, m, to the translator, if it is valid then it outputs a signature σ_B by using the embedded signing algorithm Sign_B and key sk_B. At first glance it might appear that the program obfuscation provides a straightforward solution to the universal signature translators. However, as shown by Barak *et al.* [3], such a obfuscation is impossible for general programs.

To overcome this limitation, we make use of a weak but realizable notion of program obfuscation, indistinguishability obfuscation (iO) [3]. A uniform iO for a class of programs (or circuits) guarantees that obfuscations of any two equal-size circuits that compute a same function are computationally indistinguishable. Garg *et al.* [6] provided the first candidate construction of iO.

Semi-functional UST. Based on iO, unfortunately, it is seemingly hard to construct a fully-fledged universal signature translator. In this work, we construct a semi-functional universal signature translator. In a semi-functional universal signature translator system, the base scheme can be any signature scheme, but the target scheme will be forced to be a fixed signature scheme. For proof of security, we make use of punctured programming techniques which were introduced by Sahai and Waters [18] to make the signing oracles into the obfuscated program.

The proposed semi-functional universal signature translator can be seen as the first step towards answering Ateniese and Hohenberger's [1] open problem.

1.3 Paper Organization

The rest of this paper is organized as follows. Section 2 gives the preliminaries of this work. Section 3 describes the formal definition of universal signature translator and its security model. The proposed scheme and security proof are given in Section 4. Finally, Section 5 concludes this work.

2 Preliminaries

2.1 Indistinguishability Obfuscation

Definition 1. *[6] A uniform PPT indistinguishability obfuscator, iO, for a class of circuit $\{\mathcal{C}_\lambda\}$ satisfies:*

- *Functionality preservation: for all security parameters $\lambda \in \mathbb{N}$, all circuits $C \in \mathcal{C}_\lambda$, and all inputs x,*

$$\Pr[C'(x) = C(x) : C' \leftarrow iO(\lambda, C)] = 1.$$

- *Indistinguishability: for any PPT adversary, (Samp, D), there exists a negligible function negl such that: if $Pr[\forall x, C_0(x) = C_1(x) : (C_0, C_1, \tau) \leftarrow Samp(1^\lambda)] > 1 - negl(\lambda)$, then*

$$\begin{aligned}\big|Pr[D(\tau, iO(\lambda, C_0)) = 1 :&\\ &(C_0, C_1, \tau) \leftarrow Samp(1^\lambda)]\\ -Pr[D(\tau, iO(\lambda, C_1)) = 1 :&\\ &(C_0, C_1, \tau) \leftarrow Samp(1^\lambda)]\big| \le negl(\lambda).\end{aligned}$$

2.2 Punctured Pseudorandom Functions

Definition 2. *[18] A family of punctured pseudorandom function (P-PRF) consists of three algorithms, $(\mathsf{Key}_P, \mathsf{Pun}, F)$, and a pair of computable functions $n(\cdot)$ and $m(\cdot)$, satisfying the following conditions:*

- *Functionality preserved under puncturing: any PPT adversary $\mathcal{A}$ outputs a set $S \subseteq \{0,1\}^{n(\lambda)}$, then for all $x \in \{0,1\}^{n(\lambda)} \backslash S$,*

$$\Pr[F(K, x) = F(K_S, x) : K \leftarrow \mathsf{Key}_P(1^\lambda), K_S \leftarrow \mathsf{Pun}(K, S)] = 1.$$

- *Pseudorandom at punctured points: for any PPT adversary $(\mathcal{A}_1, \mathcal{A}_2)$ such that $\mathcal{A}_1(1^\lambda)$ outputs a set $S \subseteq \{0,1\}^{n(\lambda)}$ and state τ, consider an experiment where $K \leftarrow \mathsf{Key}_P(1^\lambda)$ and $K_S \leftarrow \mathsf{Pun}(K, S)$, then,*

$$\Pr[\mathcal{A}_2(\tau, K_S, S, F(K,S)) = 1] \\ -\Pr[\mathcal{A}_2(\tau, K_S, S, U_{m(\lambda)\cdot|S|}) = 1] = negl(\lambda),$$

where $F(K,S)$ denotes $F(K,x_1)||\ldots||F(K,x_k)$, and $S = \{x_1,\ldots,x_k\}$ is the enumeration of the elements of S in lexicographic order, and U_ℓ denotes the uniform distribution over ℓ bits.

2.3 Digital Signatures

A signature scheme $\mathcal{S}$ consists of the following algorithms:

- $\mathsf{KeyGen}(1^\lambda)$ algorithm takes as input security parameter λ and outputs a signing-verification key pair (sk, vk).
- $\mathsf{Sign}(sk, m)$ algorithm takes as inputs the signing key sk and a message m and outputs a signature σ.
- $\mathsf{Verify}(vk, m, \sigma)$ algorithm takes as inputs the verification key vk and a purported signature σ on a message m and outputs either *accept* or *reject*.

Correctness: It is required that for any $(sk, vk) \leftarrow \mathsf{KeyGen}(1^\lambda)$ and message m, $accept \leftarrow \mathsf{Verify}(vk, m, \mathsf{Sign}(sk, m))$.

Definition 3. *[7] We say that a signature scheme, $\mathcal{S} = (\mathit{KeyGen}, \mathit{Sign}, \mathit{Verify})$, is existential unforgeability against chosen message attacks (EU-CMA) secure if, for any PPT adversary $\mathcal{A}$ with oracle access to Sign, the probability that, on input of a uniformly chosen verification key vk, $\mathcal{A}$ outputs a pair (m^*, σ^*) such that $\mathit{Verify}(vk, m^*, \sigma^*) = 1$ where m^* was not queried to Sign oracle, is negligible, where the probability is over vk and the randomness of the Sign oracle.*

3 Universal Signature Translators

In this section, we define the notion of universal signature translators. Let $\ell_{sig}, \ell_{ver}, \ell_{sk}, \ell_{vk}, \ell_{msg}$ and ℓ_σ be polynomials. Given a security parameter λ, ℓ_{sig} denotes a bound on the size of signing circuits, $\ell_{ver}(\lambda)$ denotes a bound on the size of verification circuits, $\ell_{sk}(\lambda)$ denotes a bound on the size of signign key, $\ell_{vk}(\lambda)$ denotes a bound on the size of verification key, $\ell_{msg}(\lambda)$ denotes a bound on the length of messages, and $\ell_\sigma(\lambda)$ denotes a bound on the size of signatures. For simplicity, we will drop λ when the context is clear. In addition, we assume that the signature schemes which are used by Alice and Bob are $\mathcal{S}_A = (\mathsf{KeyGen}_A, \mathsf{Sign}_A, \mathsf{Verify}_A)$ and $\mathcal{S}_B = (\mathsf{KeyGen}_B, \mathsf{Sign}_B, \mathsf{Verify}_B)$, respectively. $\mathcal{S}_A$ is called the base scheme and $\mathcal{S}_B$ is called the target scheme.

A universal signature translator consists of the following algorithms:

- $\mathsf{Setup}(1^\lambda)$ algorithm takes as input security parameter λ and outputs public parameters *params*.
- $\mathsf{TranKey}(params, (\mathsf{Verify}_A, vk_A, \mathsf{Sign}_B, sk_B))$ algorithm takes as input public parameters *params* and a tuple $(\mathsf{Verify}_A, vk_A, \mathsf{Sign}_B, sk_B)$ that is $(\ell_{ver}, \ell_{vk}, \ell_{sig}, \ell_{sk})$-length qualified. It outputs a translator, *i.e.*, proxy re-signing key, $rk_{A\rightarrow B}$.
- $\mathsf{Translate}(params, rk_{A\rightarrow B}, (m, \sigma_A))$ algorithm takes as input public parameters *params*, translator $rk_{A\rightarrow B}$, and a tuple (m, σ_A) that is $(\ell_{msg}, \ell_\sigma)$-length qualified. It outputs a signature σ_B on the same message m for Bob or $\perp$.

Correctness: Let $\mathcal{S}_A = (\mathsf{KeyGen}_A, \mathsf{Sign}_A, \mathsf{Verify}_A)$ and $\mathcal{S}_B = (\mathsf{KeyGen}_B, \mathsf{Sign}_B, \mathsf{Verify}_B)$ be two signature schemes.

For all $\lambda \in \mathbb{N}$, $params \leftarrow \mathsf{Setup}(1^\lambda)$, $rk_{A\rightarrow B} \leftarrow \mathsf{TranKey}(params, (\mathsf{Verify}_A, sk_A, \mathsf{Sign}_B, sk_B))$, $\sigma_B \leftarrow \mathsf{Translate}(params, rk_{A\rightarrow B}, (m, \sigma_A))$, if $(\mathsf{Verify}_A, vk_A, \mathsf{Sign}_B, sk_B)$ is $(\ell_{ver}, \ell_{vk}, \ell_{sig}, \ell_{sk})$-length qualified, (m, σ_A) is $(\ell_{msg}, \ell_\sigma)$-length qualified, and $\mathsf{Verify}_A(vk_A, m, \sigma_A) = 1$, then we require that

$$\mathsf{Verify}_B(vk_B, m, \sigma_B) = 1.$$

If $\mathcal{S}_A$ and $\mathcal{S}_B$ can be any type of signature schemes, then we say that it is a fully-fledged universal signature translator system. If $\mathcal{S}_A$ can be any signature scheme, but $\mathcal{S}_B$ is forced to be a fixed signature scheme, then we say that it is a semi-functional universal signature translator system.

3.1 Security Model of UST

We now define the security model for universal signature translators. Let $\mathcal{S}_A = (\mathsf{KeyGen}_A, \mathsf{Sign}_A, \mathsf{Verify}_A)$ and $\mathcal{S}_B = (\mathsf{KeyGen}_B, \mathsf{Sign}_B, \mathsf{Verify}_B)$ be any two $(\ell_{sig}, \ell_{sk}, \ell_{ver}, \ell_{vk}, \ell_{sig}, \ell_\sigma)$-length qualified signature schemes. Consider the following game which is played by an adversary $\mathcal{A}$ and a challenger.

- **Setup**: Challenger first runs the $(sk_A, vk_A) \leftarrow \mathsf{KeyGen}_A(1^k)$, $(sk_B, vk_B) \leftarrow \mathsf{KeyGen}_B(1^k)$ and $params \leftarrow \mathsf{Setup}(1^k)$. Next, it runs $rk_{A\rightarrow B} \leftarrow \mathsf{TranKey}\ (params, (\mathsf{Verify}_A, vk_A, \mathsf{Sign}_B, sk_B))$. Finally, it gives $(vk_A, vk_B, rk_{A\rightarrow B})$ to $\mathcal{A}$.
- **Signing queries**: $\mathcal{A}$ sends signing query (m, i), and obtains $\sigma_i \leftarrow \mathsf{Sign}_i(sk_i, m)$, where $i \in \{A, B\}$.
- **Forgery**: $\mathcal{A}$ finally outputs a forgery (m^*, σ_B^*).

We say that $\mathcal{A}$ wins the game if $accept \leftarrow \mathsf{Verify}_B(vk_B, m^*, \sigma_B^*)$ and m^* was not taken as input to the signing queries. Let $\mathbf{Adv}_{\mathcal{A}}(\lambda)$ be the probability that $\mathcal{A}$ wins the game.

Definition 4. *Let $\mathcal{S}_A = (\mathit{KeyGen}_A, \mathit{Sign}_A, \mathit{Verify}_A)$ and $\mathcal{S}_B = (\mathit{KeyGen}_B, \mathit{Sign}_B, \mathit{Verify}_B)$ be any two $(\ell_{sig}, \ell_{sk}, \ell_{ver}, \ell_{vk}, \ell_{sig}, \ell_\sigma)$-length qualified signature schemes. A universal signature translator system is secure with respect to scheme $\mathcal{S}_A$ and $\mathcal{S}_B$ if for any PPT adversary $\mathcal{A}$, $\mathbf{Adv}_{\mathcal{A}}(\lambda)$ is negligible in λ.*

We can also define a *selective* variant to the above security model where the adversary $\mathcal{A}$ is required to commit to a forgery message m^* before the setup phase.

The above security notion guarantees that if Alice and Bob are both honest, then proxy cannot create signatures for Bob unless the message was first signed by Alice. The models are relative to the notion of limited proxy of internal security given by Ateniese *et al.* [1]. In [1], they also introduced three other security notions:

1) External security: $\mathcal{A}$ cannot obtain $rk_{A\to B}$.

2) Delegatee security: If Alice is honest, then she is safe from a colluding proxy and Bob. In universal signature translator systems, Alice is always safe because the involved information about her in $\mathsf{TranKey}$ and $\mathsf{Translate}$ algorithms is only verification key vk_A, and thus if she is not safe, then this will contradict with the security of signature scheme $\mathcal{S}_A$.

3) Delegator security: If Bob is honest, then he is safe from a colluding proxy and Alice. Apparently, universal signature translators always cannot meet this security because Alice first generates a signature σ_A on a message m, by using translator $rk_{A\to B}$ which is kept by proxy, then they can compute a signature σ_B for Bob.[1]

4 Semi-Functional UST

In this section, we construct a semi-functional universal signature translator system based on indistinguishability obfuscation. In our construction, the base scheme $\mathcal{S}_A$ can be any type of signature scheme, but the target scheme $\mathcal{S}_B$ is forced to be the SW scheme [18]. The following construction is inspired by the construction of non-interactive zero knowledge proof [18] which is also based on indistinguishability obfuscation.

- $\mathsf{Setup}(1^\lambda)$ algorithm first chooses a secure indistinguishability obfuscator, iO. Next, it creates parameters $\ell_{sig}, \ell_{ver}, \ell_{sk}, \ell_{vk}, \ell_{msg}$, and ℓ_σ based on security parameter λ. Finally, it sets the public parameters to be $params = (iO, \ell_{sig}, \ell_{ver}, \ell_{sk}, \ell_{vk}, \ell_{msg}, \ell_\sigma)$.

- $\mathsf{TranKey}(params, (\mathsf{Verify}_A, vk_A, \mathsf{Sign}_B, sk_B))$ algorithm is run by Bob. He creates an obfuscation of the translating program P_t. The signature translator, *i.e.*, proxy re-signing key, $rk_{A\to B}$ is the obfuscated program $iO(P_t)$ that will be given to proxy.

- $\mathsf{Translate}(pp, rk_{A\to B}, (m, \sigma_A))$ is run by the proxy who takes as input (m, σ_A) and outputs the result.

Translating program P_t from Alice to Bob

- **Constants**: $(\mathsf{Verify}_A, vk_A, sk_B := K, \mathsf{Sign}_B := F(\cdot,\cdot))$

- **Inputs**: (m, σ_A)

  + Test if $\mathsf{Verify}_A(vk_A, m, \sigma_A) = accept$.

  + Output $\sigma_B \leftarrow F(K, m)$ if true, $\perp$ if false.

Correctness of the above universal signature translator is obvious by inspection.

4.1 Proof of Security

In this subsection, we prove the selective security of our construction. Our proof is similar to the proof of non-interactive zero knowledge in [18].

Theorem 1. *Assuming iO is a secure indistinguishability obfuscator, F is a secure P-PRF, $\mathcal{S}_A$ is a secure $(\ell_{sig}, \ell_{sk}, \ell_{ver}, \ell_{vk}, \ell_{sig}, \ell_\sigma)$-length qualified signature scheme, and $\mathcal{S}_B$ is the SW scheme, then the universal signature translator is selectively secure.*

Proof. Let $\mathcal{S}_A = (\mathsf{KeyGen}_A, \mathsf{Sign}_A, \mathsf{Verify}_A)$ be a secure $(\ell_{sig}, \ell_{sk}, \ell_{ver}, \ell_{vk}, \ell_{sig}, \ell_\sigma)$-length qualified signature scheme, $\mathcal{S}_B$ be the SW scheme, and $\mathcal{A}$ be a PPT adversary. We describe the proof as a sequence of hybrid games.

- Hyb_0 : This hybrid corresponds to the honest execution of the selective unforgeability game where the adversary $\mathcal{A}$ initially submits his challenge message m^*. We construct a challenger $\mathcal{B}$ to interact with $\mathcal{A}$:

  1) $\mathcal{B}$ first runs $(sk_A, vk_A) \leftarrow \mathsf{KeyGen}_A$. Next, it runs $K \leftarrow \mathsf{KeyGen}_P(1^\lambda)$ and chooses a one-way function f to establish the SW scheme $\mathcal{S}_B = (\mathsf{KeyGen}_B, \mathsf{Sign}_B, \mathsf{Verify}_B)$. Then, it creates a translator $rk_{A\to B}$ based on $\mathsf{Verify}_A, vk_A, \mathsf{Sign}_B := F$ and $sk_B := K$ by using iO. Finally, it gives vk_A, vk_B and $rk_{A\to B}$ to $\mathcal{A}$.

  2) $\mathcal{A}$ makes signing query on input (m, A), $\mathcal{B}$ returns back $\sigma_A = \mathsf{Sign}_A(sk_A, m)$. It also can make signing query on input (m, B), $\mathcal{B}$ returns back $\sigma_B = F(K, m)$.

  3) $\mathcal{A}$ outputs a forgery (m^*, σ_B^*) and wins the game if $\mathsf{Verify}_B(vk_B, m^*, \sigma_B^*) = 1$ and m^* was not taken as input to the signing queries.

- Hyb_1 : This hybrid is identical to Hyb_0 with the exception that we replace the real translator by an obfuscation of the program P_t^*. Programs P_t and P_t^* have a same size by appropriate padding.

[1] In some PRS schemes, such as [1,12], Bob's signatures are layered, the forgery of colluding proxy and Alice is required to be a first-level signature which cannot be generated by honest proxy and Alice.

Translating program P_t^* from Alice to Bob

- **Constants**: $(\mathsf{Verify}_A, vk_A, \mathsf{Sign}_B, K_{\{m^*\}})$
- **Inputs**: (m, σ_A)
  + Test if $\mathsf{Verify}_A(vk_A, m, \sigma_A) = accept$.
  + Output $\sigma_B \leftarrow F(K, m)$ if true, $\perp$ if false.

- Hyb_2 : This hybrid is identical to Hyb_0 with the exception that we let $z = f(F(K, m^*))$ and replace Bob's verification key $iO(P_v)$ by an obfuscation of program P_v^*. Programs P_v and P_v^* have a same size by appropriate padding.

Verification program P_v^*

- **Constants**: $(K_{\{m^*\}}, m^*, z)$
- **Inputs**: (m, σ_B)
  + If $m = m^*$, test if $f(\sigma_B) = z$. Otherwise, test if $f(\sigma_B) = f(F(K, m))$.
  + Output *accept* if true, *reject* if false.

- Hyb_3 : This hybrid is identical to Hyb_2 with the exception that $z := y$ for y is chosen uniformly at random from the range of the one-way function.

Lemma 1. *If iO is secure, then no PPT adversary can distinguish between Hyb_0 and Hyb_1.*

Proof. Note that the input and output behaviors of the programs P_t and P_t^* are identical. The only difference is that in P_t^* the P-PRF at point m^* is punctured output of the P-PRF key $K_{\{m^*\}}$. However, according to the definition of selective security, $\mathcal{A}$ always cannot obtain a valid signature on the challenge message m^*, that is to say, $F(K, m^*)$ will never get called. Therefore, if there is a PPT adversary $\mathcal{A}$ has different advantages in Hyb_0 and Hyb_1, then we can create a pair of algorithms $(Samp, D)$ to break the security of iO. $Samp$ submits two programs $C_0 = P_t$ and $C_1 = P_t^*$ to the iO challenger. Then $Samp$ will receive an obfuscation of C_0 or C_1. If the iO challenger chooses C_0, then $\mathcal{A}$ is in Hyb_0. If the iO challenger chooses C_1, then $\mathcal{A}$ is in Hyb_1. Finally, D outputs 1 if the adversary successfully forges. In conclusion, any adversary with different advantages in Hyb_0 and Hyb_1 will lead to $(Samp, D)$ as an attacker on the indistinguishability security of iO. □

Lemma 2. *If iO is secure, then no PPT adversary can distinguish between Hyb_1 and Hyb_2.*

Proof. Note that the input and output behaviors of the programs P_v and P_v^* are identical. The only difference is that P_v computes $F(K, m^*)$ by itself, whereas P_v^* is given $f(F(K, m^*))$ as a constant z. Therefore, if there is a PPT adversary $\mathcal{A}$ has different advantages in Hyb_0 and Hyb_1, then we can create a pair of algorithms $(Samp, D)$ to break the security of iO. $Samp$ submits two programs $C_0 = P_v$ and $C_1 = P_v^*$ to the iO challenger. Then $Samp$ will receive an obfuscation of C_0 or C_1. If the iO challenger chooses C_0, then $\mathcal{A}$ is in Hyb_1. If the iO challenger chooses C_1, then $\mathcal{A}$ is in Hyb_2. Finally, D outputs 1 if the adversary successfully forges. In conclusion, any adversary with different advantages in Hyb_1 and Hyb_2 will lead to $(Samp, D)$ as an attacker on the indistinguishability security of iO. □

Lemma 3. *If P-PRF is secure, then no PPT adversary can distinguish between Hyb_2 and Hyb_3.*

Proof. If there is a PPT adversary $\mathcal{A}$ who has different advantages in Hyb_2 and Hyb_3, then we can build a pair of algorithms $(\mathcal{A}_1, \mathcal{A}_2)$ to break the security of the P-PRF at punctured point m^*. $\mathcal{A}_1$ first obtains m^* from the $\mathcal{A}$. It then submits m^* to the P-PRF challenger and receives a punctured key $K_{\{m^*\}}$ and challenge value a. It runs as in Hyb_2 except it sets $z = f(a)$. If $a = F(K, m^*)$, then $\mathcal{A}$ is in Hyb_2. If it was chosen uniformly at random, then $\mathcal{A}$ is in Hyb_3. $\mathcal{A}_2$ will output 1 if the adversary successfully forges. In conclusion, any PPT adversary with different advantages in Hyb_2 and Hyb_3 will leads to $(\mathcal{A}_1, \mathcal{A}_2)$ as an attacker to break the pseudorandomness of the P-PRF. □

The above lemmas show that a succession of three hybrid games where no PPT adversary can distinguish one from the next with non-negligible advantage. We now should show that in the last hybrid Hyb_3, any PPT adversary cannot succeed with non-negligible advantage. If there is an adversary has non-negligible advantage in Hyb_3, then we can use it to break the security of the one-way function. We build a challenger $\mathcal{B}$ that first receives m^* from the adversary and y from the one-way function challenger, it then sets $z := y$. If an adversary successfully forges on m^*, then by assumption he has computed a signature σ_B^* such that $f(\sigma_B^*) = z = y$. $\mathcal{B}$ outputs σ_B^* as the solution of the given one-way function challenge instance. Therefore, if the one-way function is secure, then no PPT adversary can succeed with non-negligible advantage. Since the advantages of all PPT adversary are negligibly close in each successive hybrid, this proves the theorem. □

5 Conclusion

In order to solve the open problem of proxy re-signatures, in this work, we introduce the concept of universal signature translators. A universal signature translator system can translate a type of signature scheme to another. Furthermore, we construct a semi-functional universal signature translator system based on indistinguishability obfuscation. The existence of fully-fledged universal signature translator remains open.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (No. 61702067, 61972124), in part by the Zhejiang Provincial Natural Science Foundation of China (No. LY19F020019), in part by the key project of science and technology research program of Chongqing Education Commission of China (No. KJZD-K201803701, in part by Chongqing Natural Science Foundation (No. cstc2019jcyj-msxmX0551), in part by the science and technology research program of Chongqing Municipal Education commission (No. KJQN201903701), and in part by international standards of electronic commerce transaction product traceability (No. 2017YFF0209002).

References

[1] G. Ateniese, S. Hohenberger, "Proxy re-signatures: New definitions, algorithms, and applications," in *Proceedings of the 12th ACM Conference on Computer and Communications Security (CCS'05)*, pp. 310-319, 2005.

[2] M. Blaze, G. Bleumer, M. Strauss, "Divertible protocols and atomic proxy cryptography," in *Advances in Cryptology (EUROCRYPT'98)*, pp. 127-144, 1998.

[3] B. Barak, O. Goldreich, R. Impagliazzo, S. Rudich, A. Sahai, S. P. Vadhan, K. Yang, "On the (im) possibility of obfuscating programs," *Journal of the ACM*, vol. 59, no. 2, pp. 1-18, 2001.

[4] S. S. M. Chow, R. C. W. Phan, "Proxy re-signatures in the standard model," in *Information Security (ISC'08)*, pp. 260-276, 2008.

[5] Y. Q. Deng, "A blind proxy re-signatures scheme based on random oracle," *Advanced Materials Research*, vol. 204, pp. 1062-1065, 2011.

[6] S. Garg, C. Gentry, S. Halevi, M. Raykova, A. Sahai, B. Waters, "Candidate indistinguishability obfuscation and functional encryption for all circuits," in *IEEE 54th Annual Symposium on Foundations of Computer Science (FOCS'13)*, pp. 40-49, 2013.

[7] S. Goldwasser, S. Micali, R. L. Rivest, "A digital signature scheme secure against adaptive chosen-message attacks," *SIAM Journal on Computing*, vol. 17, no. 2, pp. 281-308, 1988.

[8] D. Guo, P. Wei, D. Yu, X. Yang, "A certificateless proxy re-signature scheme," in *The 3rd IEEE International Conference on in Computer Science and Information Technology (ICCSIT'10)*, vol. 8, pp. 157-161, 2010.

[9] D. He, "A novel blind proxy re-signature scheme," *Computer Applications and Software*, vol. 29, no. 3, pp. 294-300, 2012.

[10] S. Hohenberger, V. Koppula, B. Waters, "Universal signature aggregators," in *Advances in Cryptology (EUROCRYPT'15)*, pp. 3-34, 2015.

[11] X. Hu, Z. Zhang, Y. Yang, "Identity based proxy re-signature schemes without random oracle," in *International Conference on Computational Intelligence and Security*, pp. 256-259, 2009.

[12] B. Libert, D. Vergnaud, "Multi-use unidirectional proxy re-signatures," in *Proceedings of the 15th ACM Conference on Computer and Communications Security (CCS'08)*, pp. 511-520, 2008.

[13] K. Ramchen, B. Waters, "Fully secure and fast signing from obfuscation," in *Proceedings of ACM SIGSAC Conference on Computer and Communications Security (CCS'14)*, pp. 659-673, 2014.

[14] N. R. Sunitha, B. B. Amberker, "Proxy re-signature scheme that translates one type of signature scheme to another type of signature scheme," in *The Third International Conference on Network Security & Applications*, pp. 270-279, 2010.

[15] J. Shao, Z. Cao, L. Wang, X. Liang, "Proxy re-signature schemes without random oracles," in *Progress in Cryptology (INDOCRYPT'07)*, pp. 197-209, 2007.

[16] J. Shao, M. Feng, B. Zhu, Z. Cao, P. Liu, "The security model of unidirectional proxy re-signature with private re-signature key," *Information Security and Privacy (ACISP'10)*, pp. 216-232, 2010.

[17] J. Shao, G. Wei, Y. Ling, M. Xie, "Unidirectional identity-based proxy re-signature," in *International Conference on Communications*, pp. 1-5, 2011.

[18] S. Sahai, B. Waters, "How to use indistinguishability obfuscation: Deniable encryption, and more," *Proceedings of the Forty-Sixth Annual ACM Symposium on Theory of Computing (STOC'14)*, pp. 475-484, 2014.

[19] M. Tian, "Identity-based proxy re-signatures from lattices," *Information Processing Letters*, vol. 115, no. 4, pp. 462-467, 2015.

[20] F. Tang, H. Li, J. Chang, "Multi-use unidirectional proxy re-signatures of constant size without random oracles," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. 98, no. 3, pp. 898-905, 2015.

[21] P. Yang, Z. Cao, X. Dong, "Threshold proxy re-signature," *Journal of Systems Science and Complexity*, vol. 24, no. 4, pp. 816-824, 2011.

[22] H. Yang, J. Sun, X. Wang, J. Cui, "Proxy re-signature scheme from CBS to IBS," *Advanced Materials Research*, vol. 304, pp. 355-358, 2011.

Biography

Fei Tang is currently an associate professor of the College of Cyberspace Security and Law, Chongqing University of Posts and Telecommunications. His research interests are public key cryptography, blockchain, and privacy protection.

Dong Huang is currently a professor of the Chongqing University of Science and Technology and Chongqing Vocational and Technical University of Mechatronics. His research interest is information security.

Fuqun Wang is currently a lecturer of the Hangzhou Normal University. His research interest is public key cryptography.

Zhenyu Chen is currently a professor of the Chongqing Institute of Quality and Standardization. His research interests are cloud computing and information security.

Blockchain-based Trust Evaluation Mechanism for Internet of Vehicles Nodes

Peng-Shou Xie, Xi-Qiang Wang, Xiao-Jie Pan, Yi-Fan Wang, Tao Feng, and Yan Yan
(Corresponding author: Xi-Qiang Wang)

School of Computer and Communications, Lanzhou University of Technology
287, Lan-gong-ping Road, Lanzhou, Gansu 730050, China
Email: xiqiang0704@163.com

(Received Sept. 21, 2020; Revised and Accepted Apr. 24, 2021; First Online Oct. 19, 2021)

Abstract

Vehicle nodes share traffic information through broadcasting in the Internet of Vehicles. Vehicle behavior decisions and safe driving will be affected by false messages. Existing trust evaluation schemes require a lot of calculations to verify the trustworthiness of nodes for each communication, which has limitations and can not be applied to the Internet of Vehicles environment of instant communication. A vehicle node trust evaluation mechanism based on blockchain was proposed in order to catch malicious nodes efficiently and further ensure vehicle driving and information security. Firstly, the reliability of the node is preliminarily judged by querying the comprehensive trust value. Then the authenticity of the event was judged and the corresponding message trust value was calculated as the subsequent impact. Next, the comprehensive trust value of the vehicle node was maintained by using blockchain. The proof of trust was used to replace the proof of work in the consensus mechanism. Simulation experiments showed that the mechanism was feasible, it had obvious advantages in terms of precision rate and recall rate.

Keywords: Blockchain; Comprehensive Trust Value; Internet of Vehicles; Proof of Trust; Trust Assessment

1 Introduction

The Internet of Vehicles (IoV) is an important foundation for the intelligent transportation systems. Vehicle nodes can share traffic information in real time. There are also corresponding safety issues while IoV can improving travel efficiency [15]. The scheme based on trust value has been widely used in the research of safety management. However, vehicle nodes are moving at a high speed, the network topology is constantly changing, and the time for vehicle node communication and message inspection is very short in the Internet of Vehicles, which puts forward higher requirements for trust evaluation mechanism. How to efficiently evaluate, maintain and manage the trust of vehicle nodes is an urgent problem to be solved.

Trust evaluation is to use the concept of trust to evaluate the reliability of nodes and messages. At present, vehicle node trust evaluation schemes [9] can be generally divided into three categories: Node-based trust evaluation method [2, 5], which mainly evaluates reliability according to the behavior of vehicle nodes. Message-based trust evaluation method [3, 20], which mainly considers whether the message is reliable, and it needs to evaluate each message received. Based on the node and message evaluation method [7, 10, 13, 14, 17], the measure needs to consider both the node and message factors. Most of the above schemes implement node evaluation through previous experience, recommendations from surrounding vehicles and recommendations from central agencies, or focus on evaluating the credibility of received messages. Even the evaluation scheme that combines node behavior and message trust does not take into account the problem of untrustworthy messages caused by behavioral trustworthy nodes.

In order to efficiently implement dynamic trust evaluation, accurately catch malicious vehicle nodes, and solve the problem of untrustworthy messages caused by behavioral trusted nodes, a new trust evaluation mechanism was proposed. The comprehensive trust value was taken as the basis of trust in this mechanism, and the message trust value obtained by instant messaging serves was taken as the subsequent influence of the comprehensive trust value. The distributed vehicle node trust database was constructed with blockchain technology to maintain and manage the trust value data of vehicle nodes, and the proof of trust mechanism was used to replace the proof of work mechanism.

2 Advantages of Blockchain Used in Trust Evaluation

In recent years, blockchain has attracted widespread attention in both academia and industry field. The blockchain allows all nodes to verify the correctness of the content in the block, and has inherent advantages such as tamper resistance, transparency, decentralized and consistency. It can be applied to smart contracts, smart transportation, data management, trust evaluation and many other fields [6,8,16].

The blockchain network uses multiple copies to synchronize management, making the cost of tampering with data too high and almost impossible to achieve, which provides trust support for data management. Based on the original data information authenticity management measures, with the help of external collaboration, the verification system of vehicle trust data is improved from the outside to the inside, forming a closed loop of vehicle trust. It can make up for the shortcomings in the existing vehicle trust data management. Specifically, the storage of trust data is based on cryptographic principles and has the characteristics of immutability, which can save third-party confirmation. Each node masters all transaction information, which can solve the problem of message asymmetry.

Blockchain and related concepts provide a new direction for the security field of the Internet of Things [1,11]. Trust evaluation and security management can also learn from its core ideas and use a decentralized distributed ledger to maintain vehicle node trust. Integrating the advantages of blockchain technology, introducing it into the Internet of Vehicle architecture, and applying to the vehicle node trust evaluation mechanism can improve the security.

At present, there have been some researches on the application of blockchain technology in the trust evaluation of the IoV nodes [4,12,18,19]. Shrestha [12] proposed a mechanism for evaluating the credibility of vehicle nodes and messages in IoV. Node credibility and message credibility were stored in the blockchain ledger, and the idea of proof of location (PoL) was proposed. Yang [19] proposed a consensus algorithm based on the credibility of vehicle node messages.

3 Trust Evaluation Mechanism for Internet of Vehicles Nodes

A blockchain-based trust evaluation mechanism for IoV nodes was proposed. In this mechanism, a blockchain network was constructed by vehicle nodes and road side unit (RSU), and the trust value of vehicle nodes was regarded as a transaction in the IoV to form a new block similar to a Bitcoin transaction, and each block hash was linked in a sequential manner. Among them, the vehicle node only carried out message communication and did not undertake data calculation. RSU verified the authenticity of the event, calculated the corresponding message trust value, and then selected one of the nodes for accounting according to the consensus mechanism. All the trust values in the blockchain were accumulated to form a comprehensive trust value of vehicle nodes, which were transmitted, stored and maintained in the network. The mechanism frame was described in Figure 1.

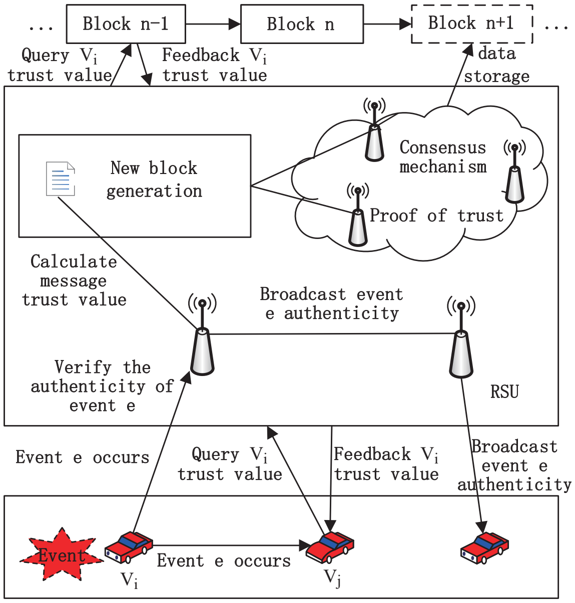


Figure 1: Blockchain-based trust evaluation mechanism frame for IoV nodes

3.1 Trust Evaluation Process for Internet of Vehicles Nodes

A trust evaluation strategy suitable for the IoV was formed according to the communication characteristics of vehicle nodes. When receiving the message from the vehicle node V_i about the occurrence of the event e, the vehicle node V_j first obtained the trust value of V_i through query, and judged its trust level. Then RSU verified the authenticity of the event, and calculated the message trust value of the sending node as the follow-up impact, which affected the comprehensive trust value of the source node. The process was shown in Figure 2.

Step 1. The node behavioral trust value was calculated during initialization as the basis for trustworthiness judgment. The vehicle node V_i found that event e has occurred, sent the event message to nearby vehicles, and reported it to the nearest RSU.

Step 2. When V_j received the message from V_i about the occurrence of event e, it first obtained the comprehensive trust value of V_i through query, and judged the credibility of the message.

Step 3. When the RSU received a message from V_i about the occurrence of event e, it quickly verified the authenticity of the event.

Step 4. Calculated the message trust value of the source node according to the authenticity of the event.

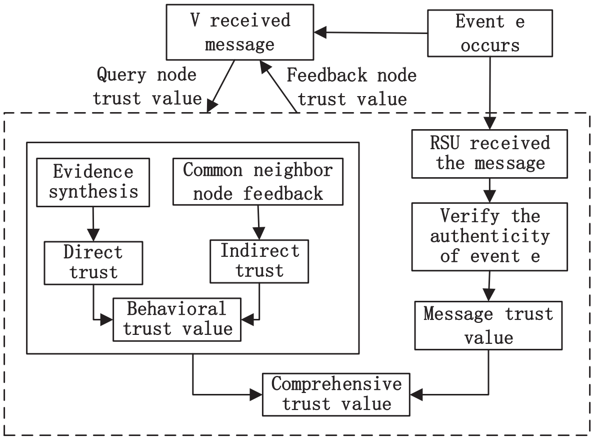


Figure 2: Trust evaluation process for Internet of Vehicles nodes

Step 5. The message trust value was treated as a transaction to form a new block and uploaded to the blockchain for storage. The message trust value was used as the follow-up impact to influence the comprehensive trust value of the node.

Reasonably quantify and classify the comprehensive trust value. The trust level TL was used to determine whether the vehicle node is trustworthy, so that the vehicle node could obtain the sender's trust level in time after receiving the message. According to the trust value, it was divided into five levels: Not credible; Uncertain credibility; Generally credible; More credible; Highly credible. As shown in Equation (1). If the trust value was less than 0, it was considered as an untrusted node and its communication would be restricted. The less trusted node was the suspicious node. Trusted nodes were divided into three categoryies: general trustworthy, relatively trustworthy, and highly trustworthy. It would only decrease but not increase after the trust value reached the maximum value of 1, that was, the trust value was still 1 when the true message was continued to be sent, and the trust value would be reduced when the false message was sent.

$$TL = \begin{cases} not \quad credible; -1 \leq T_i \leq 0 \\ uncertain \quad credible; 0 < T_i \leq 0.4 \\ generally \quad credible; 0.4 < T_i \leq 0.6 \\ more \quad credible; 0.6 < T_i \leq 0.8 \\ highly \quad credible; 0.8 < T_i \leq 1 \end{cases} \quad (1)$$

3.2 Method for Calculating Comprehensive Trust Value of IoV Nodes

(1) Method for Calculating Behavioral Trust Value

The behavioral trust value was obtained by calculating the direct trust value (DT) and the indirect trust value (IT). The direct trust value was calculated by D-S evidence theory, and the indirect trust value was fused by feedback from multiple common neighbor nodes when the system was initialized. They were weighted to obtain the behavioral trust value of the vehicle node and stored it.

The calculation of the direct trust value was calculated by D-S evidence theory. All kinds of evidence attributes reflecting node behavior characteristics were considered, and DT was calculated according to the evidence attributes related to the vehicle node. The full set Θ of research objects was expressed as 2^{Θ}, and the relationship between entities was divided into trustworthy, untrustworthy and uncertain, which were represented by {T}, {D} and {T,D} respectively.

Four typical attribute evidences were selected through the mining and analysis of node behavior characteristic attributes. Data integrity (E_1), data consistency (E_2), forwarding timeliness (E_3), and data forwarding rate (E_4), and the basic probability distribution was performed, it was shown in Table 1.

Table 1: Basic probability distribution

Trust result	E_1	E_2	E_3	E_4
{T}	m_{11}	m_{21}	m_{31}	m_{41}
{D}	m_{12}	m_{22}	m_{32}	m_{42}
{T,D}	m_{13}	m_{23}	m_{33}	m_{43}

Among them, $m_{i,j}(i = 1, 2, 3, 4; j = 1, 2, 3)$ represented the credibility probability corresponding to the four evidence attributes (E_1, E_2, E_3, E_4). The probability distribution function m represented the trust degree of the evidence attribute to the event.

Under the recognition framework Θ , for $\forall \subseteq 2^{\Theta}$, the synthesis rule of the probability distribution function $m_1, m_2 ... m_n$ was:

$$\begin{aligned} m(A) &= m_1 \oplus m_2 \oplus ... m_n \\ &= \frac{1}{K} \sum_{A_1 \cap A_2 \cap ... A_n} m_1(A_1) \cdot m_2(A_2) \cdot ... m_n(A_n) \end{aligned} \quad (2)$$

Among them, K was the normalization factor,

$$\begin{aligned} K &= \sum_{A_1 \cap A_2 \cap ... A_n \neq \emptyset} m_1(A_1) \cdot m_2(A_2) \cdot ... m_n(A_n) \\ &= 1 - \sum_{A_1 \cap A_2 \cap ... A_n \neq \emptyset} m_1(A_1) \cdot m_2(A_2) \cdot ... m_n(A_n) \end{aligned} \quad (3)$$

Then according to Equation (2), the attribute evidence was synthesized to obtain the triplet representing the direct trust of the vehicle node. DT was calculated according to the normalization method. It was shown in Equation (4).

$$DT_{i,j} = \{m_{i,j}(\{T\}), m_{i,j}(\{D\}), m_{i,j}(\{T, D\})\} \quad (4)$$

Indirect trust was that the evaluation node V_j obtained the direct trust value of the evaluated node V_i through the third-party node V_k. The trust relationship between

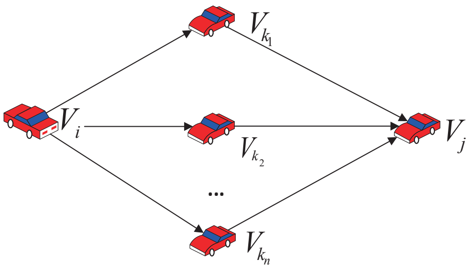


Figure 3: Indirect trust relationship diagram of vehicle nodes

nodes was shown in Figure 3. Reducing the time delay in the indirect trust evaluation process to avoid the iteration of credibility. Supposed that only the common neighbor node V_k of V_i and V_j can provide indirect trust. The direct trust value was used as the behavior trust value if there was no common neighbor node.

However, vehicle nodes are moving at high speed and their positions are constantly changing in the real Internet of Vehicles environment. So nodes V_i and V_j may have multiple common neighbor nodes such as $\{V_{k_1}, V_{k_2}, ...V_{k_n}\}$. These nodes may send forged or erroneous information to V_i for their own interests. The calculation of the indirect trust value is greatly affected in this case.

Therefore, a comprehensive analysis of the direct trust value sent by neighbor nodes was required. After V_i broadcasted the trust data requesting V_j, the common neighbor nodes V_{k_n} that met the conditions immediately respond. The direct trust value of V_j with V_i was sent to the requesting node V_i. Here, $DT_{i,k_n}$was the direct trust value of V_i to the common neighbor node V_{k_n}, and $DT_{k_n,j}$ was the direct trust value of the common neighbor node DT_{i,k_n} to V_j, DT_{i,k_n} as the reliability weight of $DT_{k_n,j}$, and n represented the number of common neighbor nodes. The IT was calculated according to Equation (5).

$$IT_{i,j} = \frac{\sum_{x=1}^{n} DT_{i,k_x} \cdot DT_{k_x,j}}{n} \tag{5}$$

The DT and IT were calculated according to the above steps, appropriate weights were introduced, and the behavior trust value was obtained by weighted calculation method. It was shown in Equation (6).

$$BT_{i,j} = \alpha DT_{i,j} + \beta IT_{i,j} \tag{6}$$

Among them, $\alpha, \beta \in (0,1)$ represented the weight of the node's direct trust and indirect trust, and $\alpha + \beta = 1$, which is determined by the node's own strategy.

(2) Method for Calculating Message Trust Value

RSU verified the authenticity of the event by synthesizing the messages sent by multiple source nodes. RSU would receive multiple source messages sent by nodes about e when event e occurs. The RSU put them in the packet $M^j = \{m_1^j, m_2^j, ...m_i^j\}$ after receiving these messages, m_i^j jrepresented the message of the vehicle node V_i about reporting the event e_j. The location proof was added to the incident report message, and this message was considered when the source sending node and the RSU were in the same area. The incident information m_i^j contained the proof of the distance between the vehicle node and the place where the incident occurred. Proof of distance was used to verify the vehicle nodes close to the event, and only the source node that sending the event message was considered. The credibility of the message content varied with the distance between the reporting node and the location of the incident. The node closest to the location of the incident usually had a higher credibility. The definition of the credibility of a certain message content was shown in Equation (7).

$$mc_i^j = b + e^{-\gamma \cdot d_i^j} \tag{7}$$

Among them, mc_i^j was the credibility of the content of the message sent by V_i about the event e_j. d_i^j is the distance between the location of the message sending node V_i and the place where the event e_j occurred. b and γ were two preset parameters, which respectively controlled the lower limit and change rate of content credibility. RSU can obtain the message credibility set mc_i^j about the even e_j through Equation (7), where $MC_i^j = \{mc_1^j, mc_2^j, ...mc_i^j\}$. Based on the message content credibility set mc_i^j, the Bayesian inference theory (Equation (8)) was used to verify the authenticity of the event e_j.

$$\begin{aligned} & p(e_j/MC_i^j) \\ & = \frac{p(e_j) \cdot \prod_{i=1}^{N} p(c_i^j/e_j)}{p(e_j) \cdot \prod_{i=1}^{N} p(c_i^j/e_j) + p(\overline{e}_j) \cdot \prod_{i=1}^{N} p(c_i^j/\overline{e}_j)} \end{aligned} \tag{8}$$

Among them, e_j was the complementary event of $\overline{e}_j$, $p(c_i^j/e_j) = c_i^j$, $p(c_i^j/\overline{e}_j) = 1 - c_i^j$, $p(e_j)$ was the prior probability of the event e_j. The authenticity of the message and the message trust of the sending node were calculated according to the authenticity of the event. The event is considered to be a real event if $p(e_j/MC_i^j)$ exceeded the preset threshold ε.

RSU judged the consistency according to the authenticity verification result of event e and the message sent by the node, and calculated the message trust value of the sending node. Give affirmation to the source node that correctly reported this event, that was, the message trust value was positive, which increased the comprehensive trust value. Give a negative evaluation to the false source node, that was, the message trust value was negative, and the comprehensive trust value would decrease.

Since the node sending messages was a continuous behavior, the counters $count_t$ and $count_u$ were used to record the feedback results of the node sending messages, $count_t$ represented the number of consecutive true messages, the calculation method was shown in Equation (9).

$$count_t = \begin{cases} count_t + 1, Message \quad is \quad true \\ 0 \qquad\qquad , Message \quad is \quad false \end{cases} \tag{9}$$

Moreover, $count_u$ represent the number of consecutive false messages, the calculation method is shown in Equation (10).

$$count_u = \begin{cases} count_u + 1, Message \quad is \quad false \\ 0 \qquad\qquad , Message \quad is \quad true \end{cases} \tag{10}$$

Rewards and punishments were implemented for sending message nodes in order to encourage vehicles to send true messages. Specifically, the vehicle node sent the true message, and the trust value increased relatively slowly in the vehicle communication process. While sending false messages, the trust value decreased very quickly, especially when the vehicle node continuously sent false messages, the trust value decreased drastically. The cost of sending false messages was far greater than the benefits of sending true messages. The message trust value was calculated by Equation (11).

$$MT_i = \begin{cases} MT_t + \varphi_t, Message \quad is \quad true \\ MT_u + \varphi_u, Message \quad is \quad false \end{cases} \tag{11}$$

$$MT_t = \frac{n}{(m+n)^2} \tag{12}$$

$$MT_u = \frac{-1}{m+n} + \frac{-n}{(m+n)^2} \tag{13}$$

Among them, MT_t was the basic trust value for sending true messages, and MT_u was the basic trust value for sending false messages. m represented the number of true messages, and n represented the number of false messages. When n was 0, the basic trust value of true messages was 0.01. φ_t and φ_u were the reward and punishment factor, respectively. The message trust value increased by multiples when the true message was continuously sent. The message trust value decreased exponentially when the false message was continuously sent, which were shown in Equation (14) and (15).

$$\varphi_t = \begin{cases} count_t \cdot MT_t, count_t \geq 2 \\ 0 \qquad\qquad , count_t < 2 \end{cases} \tag{14}$$

$$\varphi_u = \begin{cases} 2^{count_u - 1} \cdot MT_u, count_u \geq 2 \\ 0 \qquad\qquad\qquad , count_u < 2 \end{cases} \tag{15}$$

Through the above calculation, the node behavior trust value and message trust value were obtained. The comprehensive trust value was calculated the by Equation (16).

$$T_i = BT_i + MT_i \tag{16}$$

3.3 Consensus Mechanism Based on Proof of Trust

The comprehensive trust value of the vehicle node can be obtained through the trust evaluation mechanism. Due to the unmodifiable characteristics of data in the blockchain network, it was maintained by constantly adding message trust values. The working process was described as follows.

Step 1. The node broadcasted the newly generated data information, that was, the message trust value to the entire network nodes.

Step 2. Other nodes in the entire network verified the legitimacy after receiving the information. If the verification was passed, the information would be processed by the algorithm and then stored in a temporary block. The block header was calculated according to a specific format, and the block body is formed by the message data.

Step 3. Every node that received information implements a consensus algorithm for this temporary block.

Step 4. When the node that receives the information found the solution of the consensus algorithm, it was selected as the miner node and broadcasts the block to the entire network.

Step 5. Linked the block to the current longest main blockchain in chronological order to generate the latest block. That was, the new block was stored at the end of the longest chain of the current blockchain.

(1)New Block Generation

When a node in the blockchain network received a message sent by another node, it firstly checked the legitimacy of the message. Then, the verified message was processed through a hash algorithm (SHA256) and recorded in a block body in a specific way. The block data structure was shown in Figure 4. The block header contained information such as the current version number, the hash value of the previous block, the timestamp, the random number, and the Merkle root. The block body recorded the complete transaction data information of a block in the form of a Merkle tree. This structure can improve the search efficiency of the system.

The trust value message was stored in the Merkle tree. The process of adding a trust value message was described as follows. The trust value message 1 and 2 were writed into a new Merkle tree. Firstly, hash algorithm was used to obtain Hash 1 and Hash 2. Then connected them in series and performed hash again to get Hash 12. The hash algorithm was processed in this way until the last value was obtained, which was the Merkel root.

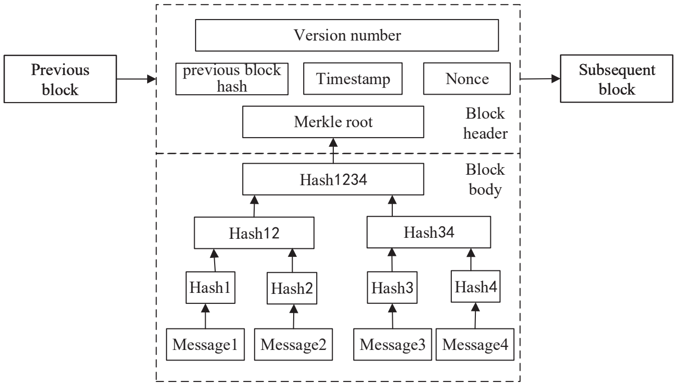


Figure 4: Block data structure

(2)Consensus Mechanism

All nodes in the blockchain network jointly maintained a comprehensive trust value, and all nodes maintained the same content. The consensus mechanism was that all nodes followed the common agreed rules and participated in the completion of asynchronous interaction. The goal was to make all nodes maintained a consistent view of the blockchain.

A consensus mechanism suitable for trust evaluation of Internet of Vehicles was proposed. It was the proof of trust (PoT) mechanism, which used trust value instead of computing power to compete for the right to ledger. Compared with common consensus algorithms, such as the proof of work (PoW) mechanism requires high-consumption computing power, the proof of staked (PoS) mechanism needed hold more shares. The PoT only needed to obtain the trust value, which can reduce computing resources and achieve better scalability.

In the Internet of Vehicles, considering that the base stations are deployed on a uniform scale, each node has roughly the same computing power. Therefore, the proof of trust mechanism was proposed, which defined the power of the node according to the trust value of the received message. The greater the trust value of the total message received by the node, the greater the power of the node is, the greater the impact on the trust status of the vehicle node is, and the significance is more profound.

At this point, the miner added a new block that records the trust value of the message to the blockchain. The update of the comprehensive trust value of the vehicle node was completed, and the latest comprehensive trust value of the vehicle node could be obtained.

4 Simulation Experiment and Performance Analysis

In this section, the performance of proposed scheme was evaluated and the simulation results were presented. The traffic simulation platform SUMO and network simulation NS-2 were used as the simulation tools in this paper. By simulating vehicle communication scenarios, a trace file of vehicle driving data was generated and loaded into NS-2. The simulation results depicted the changes in the trust value of vehicle nodes under this mechanism. And introduce precision evaluation was introduced to verify the accuracy of the mechanism. The parameters were listed in Table 2.

Table 2: Simulation parameters

Parameter	Value
Simulation area	1km × 1km
Simulation time	600s
Transmission range	200m
MAC protocol	802.11p
Number of nodes	100
Num. of malicious nodes	10/20/30
Node motion speed	10m/s-20m/s

4.1 Changes in Trust Value

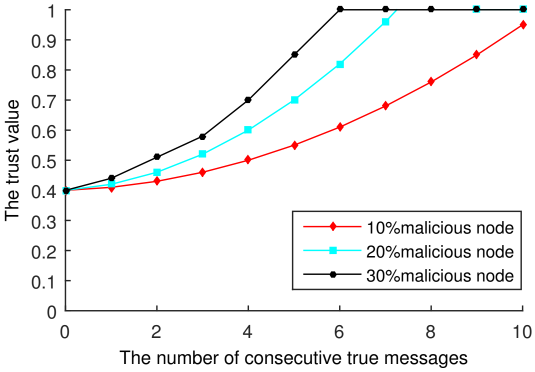


Figure 5: Changes in the trust value of consecutive sending true messages

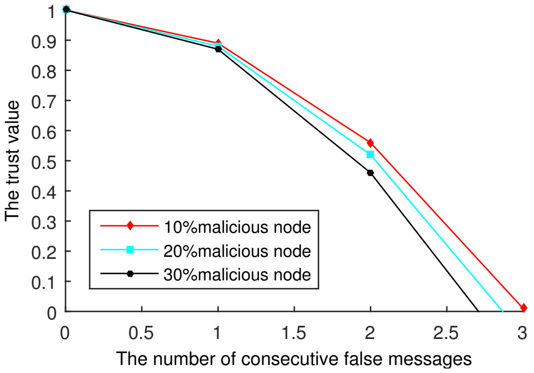


Figure 6: Changes in the trust value of consecutive sending false messages

Figure 5 showed that the trust value increased slowly as the node sends true messages multiple times under different proportions of malicious nodes. It can be seen that vehicle nodes have to send true messages multiple times in succession to accumulate higher trust values. The rate of trust value accumulation increased relatively as the proportion of malicious nodes increased. Normal nodes can increase their trust value relatively quickly, creating a larger gap between them and malicious nodes.

Figure 6 showed that the trust value dropped rapidly as the nodes continuously send false messages under different proportions of malicious nodes. After two consecutive send false messages, the trust value plummeted to the edge of general trust. Sending false information three times in a row can be judged as a malicious node. Among them, the trust value decreases faster in scenarios where malicious nodes account for a large proportion, which helps to quickly identify malicious nodes.

In order to better the trust evaluation effect of the scheme, the mechanism was compared with the trust evaluation scheme proposed in literature [17]. In the comparison algorithm, the increase in the trust value was set to 0.01. The trust value changes were consistent with the algorithm proposed in this paper when true messages were continuously sent. Figure 7 showed the comparison of changes in trust values when false messages were continuously sent. The trust values of the two algorithms had dropped rapidly, but the reduction of the algorithm in this paper was greater than the comparison algorithm. The algorithm in this paper was efficient in catch malicious nodes.

Figures 8 and 9 showed the change of trust value when the mechanism faced intermittent attacks and shock deception. Among them, Figure 8 showed the change of trust value of a node after sending a true message once and then sending false message. Figure 9 showed the change of trust value when the node sent true messages twice, and then sendt false messages twice in a row. The comparison showed that the algorithm in this paper reduced the trust value faster when the node implements were attacked intermittently, and had certain advantages in identifying malicious nodes.

4.2 Accuracy of Trust Evaluation Mechanism

The following two parameters were used to evaluate the accuracy of this scheme: Precision Rate (PR) and Recall Rate (RR), which were both widely used in machine learning and information retrieval to assess the accuracy. The precision rate refers to the ratio of the number of truly malicious nodes caught(Num_M) to the total number of untrustworthy nodes caught (Num_U). The recall rate refers to the ratio of the number of truly malicious nodes caught (Num_M) to the total number of truly malicious nodes in the network (Num_{TM}). These two parameters were defined as Equations (17) and (18).

$$PR = \frac{Num_M}{Num_U} \tag{17}$$

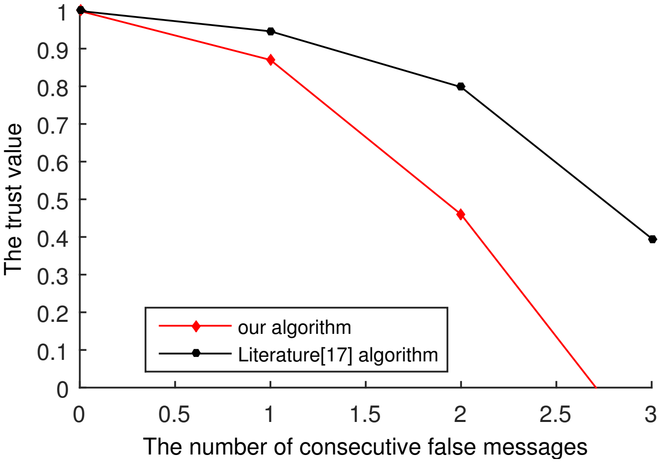


Figure 7: Comparison of changes in trust value of continuously sending false messages

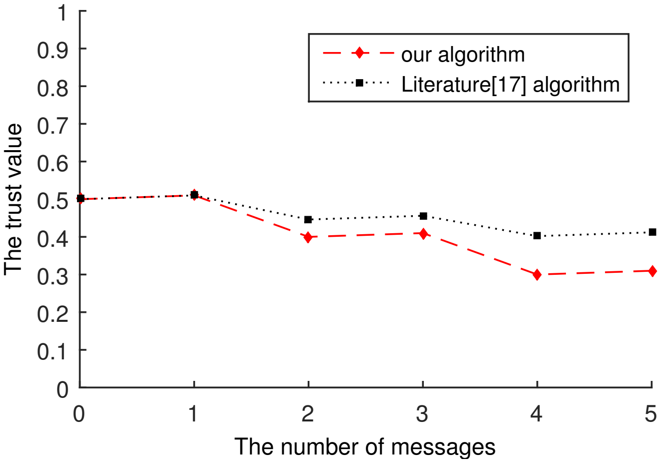


Figure 8: Changes in the trust value of sending true and false messages at intervals

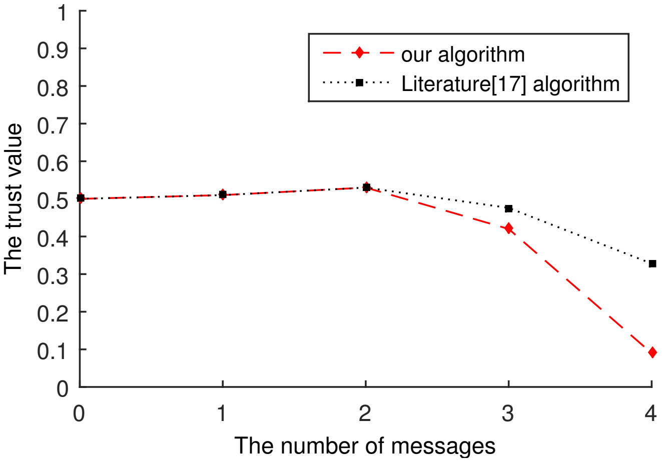


Figure 9: Changes in the trust value of sending true and false messages at twice intervals

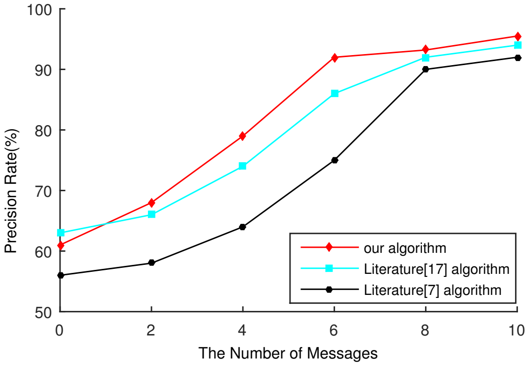


Figure 10: Changes in precision rate with the number of messages

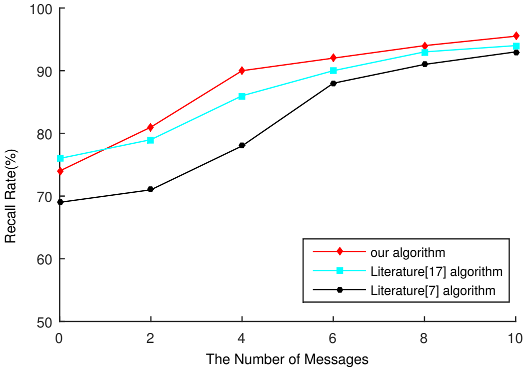


Figure 11: Changes in recall rate with the number of messages

$$RR = \frac{Num_M}{Num_{TM}} \tag{18}$$

In order to reduce the error, the experiment scene was simulated for many times to find the average value. Figure 10 showed the comparison of the three schemes in terms of PR, and Figure 11 showed the comparison of the three schemes in terms of RR. It can be seen in the three schemes that both PR and RR increase as the number of messages increases. When the node sent fewer messages, the precision rate was a relatively low value. Malicious nodes might continuously send false messages as the number of messages increased. The trust value of malicious nodes dropped sharply, which maked the precision rate higher under the action of the reward and punishment mechanism. These three schemes can achieve higher accuracy when the node sends multiple messages. Compared with the comparison scheme, our scheme had higher accuracy in terms of PR and RR, and could accurately identify malicious nodes, which had certain advantages.

5 Conclusion

Trust management is an important research direction to ensure the security of the Internet of Vehicles. The vehicle node must be able to confirm the credibility of the message before responding. How to accurately and efficiently implement trust assessment and trust maintenance is an important issue. A dynamic trust evaluation mechanism based on comprehensive trust value was proposed in this paper by comparing trust evaluation schemes and research trust evaluation strategies. The blockchain was used to store and maintain the trust value of vehicle nodes, and a consensus mechanism for proof of trust was proposed. Simulation experiments had proved that this mechanism could accurately and efficiently catch malicious nodes. To a certain extent, it improved the efficiency of trust evaluation and met the requirements of dynamic evaluation. However, only the trust change of the node sending the message was considered in the trust evaluation mechanism, and the trust change of other vehicle nodes will be studied as the next step of research.

Acknowledgement

This research is supported by the National Natural Science Foundations of China under Grants No. 61862040, No. 61762060 and No. 61762059. The authors gratefully acknowledge the anonymous reviewers for their helpful comments and suggestions.

References

[1] P. F. Fan, Y. Z. Liu, and J. Y. Zhu, "Identity management security authentication based on blockchain technologies," *International Journal of Network Security*, vol. 21, no. 6, pp. 912–917, 2019.

[2] R. W. Heijden, S. Dietzel, and T. Leinmueller, "Survey on misbehavior detection in cooperative intelligent transportation systems," *IEEE Communication Surveys and Tutorials*, vol. 21, no. 1, pp. 779–811, 2019.

[3] Y. Huo, W. Dong, and J. Qian, "Coalition game-based secure and effective clustering communication in vehicular cyber-physical system (VCPS)," *Sensors*, vol. 17, no. 3, pp. 475–482, 2017.

[4] J. W. Kang, R. Yu, and X. M. Huang, "Blockchain for secure and efficient data sharing in vehicular edge computing and networks," *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 4660–4670, 2019.

[5] C. A. Kerrache, A. Lakas, and N. Lagraa, "UAV-assisted technique for the detection of malicious and selfish nodes in vanets," *Vehicular Communications*, vol. 11, pp. 1–11, 2018.

[6] Z. C. Li, J. H. Huang, and D. Q. Gao, "ISCP: An improved blockchain consensus protocol," *International Journal of Network Security*, vol. 21, no. 3, pp. 359–367, 2019.

[7] Z. M. Li and C. X. Guo, "On joint privacy and reputation assurance for vehicular ad hoc networks," *IEEE Transactions on Mobile Computing*, vol. 13, no. 10, pp. 2334–2344, 2014.

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

[9] Z. J. Lu, G. Qu, and Z. L. Liu, "A survey on recent advances in vehicular network security, trust, and privacy," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 2, pp. 760–776, 2018.

[10] A. Mansouri and M. S. Bouhlel, "Trust in adhoc networks: A new model based on clustering algorithm," *International Journal of Network Security*, vol. 21, no. 3, pp. 483–493, 2019.

[11] W. She, Q. Liu, and Z. Tian, "Blockchain trust model for malicious node detection in wireless sensor networks," *IEEE Access*, vol. 7, pp. 38947–38956, 2019.

[12] R. Shrestha, R. Bajracharya, and S. Y. Nam, "Blockchain-based message dissemination in vanet," in *IEEE 3rd International Conference on Computing, Communication and Security (ICCCS'18)*, pp. 161–166, Oct. 2018.

[13] S. Su, Z. H. Tian, and S. Y. Liang, "A reputation management scheme for efficient malicious vehicle identification over 5g networks," *IEEE Wireless Communicat-ions*, vol. 27, no. 3, pp. 46–52, 2020.

[14] Z. H. Tian, X. S. Gao, and S. Su, "Evaluating reputation management schemes of internet of vehicles based on evolutionary game theory," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 6, pp. 5971–5980, 2019.

[15] C. G. Wang, Z. M. Dai, and D. M. Zhao, "A novel identity-based authentication scheme for iov security," *International Journal of Network Security*, vol. 22, no. 4, pp. 627–637, 2020.

[16] L. Wang, D. Zheng, and R. Guo, "A blockchain-based privacy-preserving authentication scheme with anonymous identity in vehicular networks," *International Journal of Network Security*, vol. 22, no. 6, pp. 981–990, 2020.

[17] S. B. Wang and N. M. Yao, "A rsu-aided distributed trust framework for pseudonym-enabled privacy preservation in vanets," *Wireless Networks*, vol. 25, no. 3, pp. 1099–1155, 2019.

[18] Z. S. Xu, W. Liang, and K. C. Li, "A blockchain-based roadside unit-assisted authentication and key agreement protocol for internet of vehicles," *Journal of Parallel and Distributed Computing*, vol. 149, pp. 29–39, 2020.

[19] Y. T. Yang, L. D. Chou, and C. W. Tseng, "Blockchain-based traffic event validation and trust verification for vanets," *IEEE Access*, vol. 7, pp. 30868–30877, 2019.

[20] Z. Yang, K. Zheng, and K. Yang, "A blockchain-based reputation system for data credibility assessment in vehicular networks," in *IEEE 28th Annual Internatio-nal Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC'17)*, pp. 1–5, Oct. 2017.

Biography

Peng-shou Xie was born in Jan. 1972. He is a professor and a supervisor of master student at Lanzhou University of Technology. His major research field is Security on Internet of Things. E-mail: xiepsh_lut@163.com

Xi-qiang Wang was born in Jul. 1996. He is a master student at Lanzhou University of Technology. His major research field is network and information security. E-mail: xiqiang0704@163.com.

Xiao-jie Pan was born in Feb. 1996. He is a master student at Lanzhou University of Technology. His major research field is network and information security. E-mail: 1075224210@qq.com.

Yi-fan Wang was born in Aug. 1996. She is a master student at Lanzhou University of Technology. Her major research field is network and information security. E-mail: 844782234@qq.com.

Tao Feng was born in Dec. 1970. He is a professor and a supervisor of Doctoral student at Lanzhou University of Technology. His major research field is modern cryptography theory, network and information security technology. E-mail: fengt@lut.cn

Yan Yan was born in Oct. 1980. She is a associate professor and a supervisor of master student at Lanzhou University of Technology. Her major research field is privacy protection, multimedia information security. E-mail: yanyan@lut.cn

A Note on One Privacy-Preserving Centralized Dynamic Spectrum Access System

Lihua Liu[1] and Xinyuan Cao[2]
(Corresponding author: Lihua Liu)

Department of Mathematics, Shanghai Maritime University[1]
Haigang Ave 1550, Shanghai 201306, China
School of Business, East China University of Science and Technology, China[2]
Email: liulh@shmtu.edu.cn

(Received Sept. 24, 2020; Revised and Accepted Apr. 24, 2021; First Online Oct. 19, 2021)

Abstract

Dynamic spectrum access technique is a crucial solution to mitigate the potential spectrum scarcity problem. Recently, Dou *et al.* have presented a privacy-preserving centralized dynamic spectrum access system [IEEE Journal on Selected Areas in Communications, vol. 35, no. 1, pp. 173–187, 2017], based on Paillier public-key encryption and secure multi-party computation. This note shows that the scheme fails to prevent the distributor from determining whether a target secondary user is authorized by the server and recover the user's operation data. The practical running modulus in the suggested public key encryption is 4096 bits, and the encryption should be used to blind all data, not any session key as usual. The shortcoming renders the scheme quite inefficient.

Keywords: Dynamic Spectrum Access; Paillier Encryption; Running Modulus; Secure Multi-party Computation

1 Introduction

Centralized spectrum management is a mechanism to govern the spectrum sharing between government incumbent users (IUs) and commercial secondary users (SUs). With the development of spectrum access system, privacy has become more and more serious. Since operation information of government IUs is often classified, these IUs' operation data are highly sensitive. Similarly, SUs' operation data may also be sensitive commercial secrets for their operators.

In 2013, Gao *et al.* [9] considered the location privacy in database-driven cognitive radio networks. After that, Bahrak *et al.* [1, 20] investigated the problem of location spoofing attack and its countermeasures in database-driven cognitive radio networks. Jin *et al.* [12] presented a scheme for safeguard dynamic spectrum access against fake secondary users. In 2016, Dou *et al.* [7] also presented a scheme for preserving incumbent users' privacy in exclusion-zone-based spectrum access systems.

Thakur *et al.* [17, 18] designed several frame structures for hybrid spectrum access strategy in cognitive radio communication systems, and authentication protocols for passive RFID tags. Clark *et al.* [6, 11] proposed a scalable spectrum access system for massive machine type communication. In 2019, Karimi *et al.* [3, 5, 13] have considered the problem of robust spectrum access for hybrid interweave-underlay cognitive radio systems using probabilistic spectrum access, and that of fair dynamic spectrum management in licensed shared access systems. Very recently, Pan *et al.* [4, 10, 16] have presented an enhanced secure smart card-based password authentication scheme.

Multi-party computation (MPC) allows multiple parties to jointly compute a function over their inputs, while keeping these inputs, the intermediate computation results and the outputs private. In 2009, Bogetoft *et al.* [2] discussed the practical implementation of MPC. Lindell and Pinkas investigated the possible application of MPC for privacy-preserving data mining. In 2018, Martins *et al.* [14] provided a good survey on fully homomorphic encryption (an engineering perspective). Recently, Yahyaoui and Kettani [19] designed an efficient fully homomorphic encryption scheme.

In 2017, Dou *et al.* [8] have presented a privacy-preserving centralized dynamic spectrum access system based on the Paillier public key encryption [15] and multi-party computation. It claimed that none of the IU (incumbent user) or SU (secondary user) operation data would be exposed to any snooping party. But we find the scheme is flawed, because the new entity, Key Distributor, can decide whether a target SU is authorized by the server. He can also recover the target SU's operation data. Besides, the suggested running modulus in the secure multi-party computation is of 4096 bits which renders the scheme is very inefficient.

2 Preliminaries

The involved secure multi-party computation in the scheme is based on the below variation of Paillier encryption [15].

Key generation. Choose two big primes p and q to compute $n = pq$ and $\lambda = \mathrm{lcm}(p-1, q-1)$. Pick $g \in \mathbb{Z}^*_{n^2}$ to compute

$$\mu = (L(g^{\lambda} \bmod n^2))^{-1} \bmod n, \text{where} L(x) = \frac{x-1}{n}$$

Set the public key pk as (n, g) and the secret key sk as (λ, μ).

Encryption. Given $m \in \mathbb{Z}_n$, pick a one-time random number r to compute

$$[[m]] = \mathrm{Enc}_{pk}(m, r) = g^{(m+nr)} \bmod n^2.$$

Decryption. Given the ciphertext $[[m]]$, recover the plaintext by computing

$$m = \mathrm{Dec}_{sk}([[m]]) = L([[m]]^{\lambda} \bmod n^2) \cdot \mu \bmod n.$$

3 Review of the Scheme

In the scheme, there are four entities: incumbent users (IUs), secondary users (SUs), a SAS Server $\mathcal{S}$, and a Key Distributor $\mathcal{K}$. The Server is semi-honest who can only passively monitor the executions to infer IU/SU's operation information, and cannot actively deviate from the process. The server $\mathcal{S}$ is responsible for computing spectrum allocation. The distributor $\mathcal{K}$ will not collude with the server $\mathcal{S}$ to compromise IU/SU's operation data.

The challenges for a SAS system include:

1) To ensure accurate interference management, it usually adopts complex radio propagation models, which could incur huge computation and communication overhead;

2) It should ensure that SUs' operation will not disturb any IU;

3) If an SU's spectrum access request is approved, it needs to issue a license that permits the SU to access the spectrum in a certain pattern.

The goal of the system is to realize the SAS process correctly, while preserving the IU/SU's data privacy from the semi-honest SAS server. The privacy-related parameters are summarized as follows (Table 1).

The plain scenario of spectrum access system and its secure multi-party computation scenario can be described below (Table 2). The involved operations are defined as follows, where the integers $m_1, m_2 \in \mathbb{Z}_n$ are encoded in the two's complement forms without the risk of overflow.

- Addition($\oplus$): $\mathrm{Dec}_{sk}([[m_1]] \oplus [[m_2]]) = m_1 + m_2$.
- Multiplication($\otimes$): $\mathrm{Dec}_{sk}(c \otimes [[m]]) = c \cdot m$.
- Subtraction($\ominus$): $\mathrm{Dec}_{sk}([[m_1]] \ominus [[m_2]]) = m_1 - m_2$.

Table 1: The related parameters

parameter	notation	quantization
parameter	notation	level
IU, SU location	l, j	L
IU, SU antenna height	h_I, h_S	H_I, H_S
IU, SU operating frequency	f_I, f_S	F
IU interference threshold	ζ	—
SU maximum transmit power	η	—

4 Analysis

Let $\mathrm{sign}(x) = 1$ if $x > 0$, or -1 if $x \leq 0$. The essential relations in the scheme are that

$$\begin{aligned}
\mathrm{sign}(X_b(l, h_I, f_I)) &= \mathrm{sign}(G_b(l, h_I, f_I)) \cdot \mathrm{sign}(\epsilon(l, h_I, f_I)),\\
\mathrm{sign}(Y_b(l, h_I, f_I)) &= \mathrm{sign}(X_b(l, h_I, f_I)),\\
Q_b(l, h_I, f_I) &= \mathrm{sign}(\epsilon(l, h_I, f_I)) \cdot \mathrm{sign}(Y_b(l, h_I, f_I)) - 1\\
&= \mathrm{sign}(G_b(l, h_I, f_I)) - 1\\
&= 0 \text{ or } -2,\\
\mathbf{D}_b &= \mathbf{C}_b + \sigma \sum_{l, h_I, f_I} Q_b(l, h_I, f_I).
\end{aligned}$$

If and only if $\sum_{l, h_I, f_I} Q_b(l, h_I, f_I) = 0$, *i.e.*,

$$\sum_{l, h_I, f_I} (\mathrm{sign}(G_b(l, h_I, f_I)) - 1) = 0$$

. The signature $\mathbf{C}_b$ is valid, and the user SU_b is securely authorized. The corresponding requirement in the plain scenario is that $G_b(l, h_I, f_I) > 0, \forall (l, h_I, f_I)$.

◇ *The new scenario should introduce a special entity to play the role of Distributor $\mathcal{K}$.* In the secure multi-party computation scenario, the server $\mathcal{S}$ only obtains $[[\mathbf{R}_b]], [[\mathbf{T}_i]], [[\mathbf{Y}_b]]_{pk_b}, [[\mathbf{U}_b]]$, but fails to recover the plaintexts $\mathbf{R}_b, \mathbf{T}_i$. How about the new entity $\mathcal{K}$ (the key distributor)? It suggests that [8]: "In the real world, S can be operated by some commercial third party (*e.g.*, Google) for enhanced efficiency and scalability; $\mathcal{K}$ is operated by IUs." It also specifies that: "$\mathcal{K}$ creates a group Paillier public/private key pair (pk_G, sk_G)." We want to stress that the suggestion is hard to implement practically because there are generally many incumbent users, and it is a big challenge to share the secret key sk_G among them. Moreover, it is not easy for them to answer $\mathcal{S}$'s requests dynamically and collaboratively. So, it is better to introduce a special entity to play the role.

◇ *The Distributor $\mathcal{K}$ can decide whether the secondary user SU_b is authorized by the server. Besides, he can recover the SU_b's operation data.* It specifies that [8]: "We assume $\mathcal{K}$ is trusted in keeping sk_G secret only to itself, and $\mathcal{K}$ will not collude with $\mathcal{S}$ to compromise IU/SU operation data.“ That means $\mathcal{K}$ is not fully honest. Otherwise, the server $\mathcal{S}$ can simply send $[[\mathbf{G}_b]]$ to $\mathcal{K}$, instead of its camouflage $[[\mathbf{X}_b]]$. So, the new entity $\mathcal{K}$ is assumed to have the intention to snoop IU/SU operation data. But

Table 2: Two different scenarios of SAS

The plain scenario		
SU$_b$ (secondary user)	$\mathcal{S}$ (centralized spectrum access system server)	IUs (incumbent users)
Generate $\mathbf{R}_b$. $\xrightarrow{\mathbf{R}_b}$	Compute the attenuation map $\mathbf{I} = \{I(l, j, h_I, h_S, f_I, f_S)\}$, $\mathbf{T}' = \sum_i \mathbf{T}_i$, and the interference budget matrix $\mathbf{N} = \{N(l, h_I, f_I)\}$, where $N(l, h_I, f_I) = T'(l, h_I, f_I)$ if $T'(l, h_I, f_I) \neq 0$, otherwise set $N(l, h_I, f_I) = \infty$. Compute the interference indicator matrix $\mathbf{G}_b$ by $F_b(l, h_I, f_I) = \sum_{j,h_S,f_S} R_b(j, h_S, f_S) \times I(l, j, h_I, h_S, f_I, f_S)$, $G_b(l, h_I, f_I) = N(l, h_I, f_I) - F_b(l, h_I, f_I)$. If $\exists\ (l^*, h_I^*, f_I^*)$ s.t., $G_b(l^*, h_I^*, f_I^*) \leq 0$, deny SU$_b$'s request. Otherwise, return a valid license to SU$_b$. Update $\mathbf{N}$ by $N(l, h_I, f_I) \leftarrow N(l, h_I, f_I) - F_b(l, h_I, f_I)$.	Update $\mathbf{T}_i$. $\xleftarrow{\mathbf{T}_i}$
The secure multi-party computation based scenario		
SU$_b$	$\mathcal{S}$	IUs
Generate $\mathbf{R}_b$ and encrypt it as $[[\mathbf{R}_b]]$ by pk_G. $\xrightarrow{[[\mathbf{R}_b]]}$	Compute $\mathbf{I} = \{I(l, j, h_I, h_S, f_I, f_S)\}, [[\mathbf{T}']] = \sum_i [[\mathbf{T}_i]]$, $[[\mathbf{N}]] = [[\mathbf{T}']] \oplus [[\mathbf{Z}]]$, where $\mathbf{Z}$'s entries are all set to $2^{k-1} - 1$. Compute $[[F_b(l, h_I, f_I)]]$ $= \oplus_{j,h_S,f_S} [[R_b(j, h_S, f_S)]] \otimes I(l, j, h_I, h_S, f_I, f_S)$, $[[G_b(l, h_I, f_I)]] = [[N(l, h_I, f_I)]] \ominus [[F_b(l, h_I, f_I)]]$. Choose $\alpha(l, h_I, f_I) > \beta(l, h_I, f_I) > 0, \tau(l, h_I, f_I)$, and pick $\epsilon(l, h_I, f_I) \in \{-1, 1\}$ to compute $[[X_b(l, h_I, f_I)]] = (\alpha(l, h_I, f_I) \otimes [[G_b(l, h_I, f_I)]]$ $\oplus [[\tau(l, h_I, f_I)]] \ominus [[\beta(l, h_I, f_I)]]) \otimes \epsilon(l, h_I, f_I)$. $\xrightarrow{[[\mathbf{X}_b]]}$	Update $\mathbf{T}_i$ and encrypt it as $[[\mathbf{T}_i]]$ by pk_G. $\xleftarrow{[[\mathbf{T}_i]]}$
		$\mathcal{K}$ who is practically operated by IUs, uses sk_G to decrypt $[[\mathbf{X}_b]]$. Generate $[[\mathbf{Y}_b]]$ by letting $Y_b(l, h_I, f_I) = 1$ if $X_b(l, h_I, f_I) > 0$, or $Y_b(l, h_I, f_I) = -1$ if $X_b(l, h_I, f_I) \leq 0$. Encrypt it as $[[\mathbf{Y}_b]]_{pk_b}$ by pk_b. $\xleftarrow{[[\mathbf{Y}_b]]_{pk_b}}$
	Generate $[[\mathbf{Q}_b]]_{pk_b}$ by letting $[[Q_b(l, h_I, f_I)]]_{pk_b}$ $= \big(\epsilon(l, h_I, f_I) \otimes [[Y_b(l, h_I, f_I)]]_{pk_b}\big) \ominus [[1]]_{pk_b}$. Create a spectrum license $\mathfrak{L}$ for SU$_b$. Generate a signature $\mathbf{C}_b$ of the license $\mathfrak{L}$. Pick a random integer σ to compute $[[\mathbf{D}_b]]_{pk_b} =$ $[[\mathbf{C}_b]]_{pk_b} \oplus (\sigma \otimes\ (\oplus_{l,h_I,f_I} [[Q_b(l, h_I, f_I)]]_{pk_b}))$. $\xleftarrow{\mathfrak{L}, [[\mathbf{D}_b]]_{pk_b}, [[\mathbf{F}_b]]}$	
Decrypt $[[\mathbf{D}_b]]_{pk_b}$. For $\forall\ (l, h_I, f_I)$, check whether $\mathbf{D}_b$ is a valid signature. If true, compute $[[\mathbf{U}_b]]$ by setting $[[U_b(l, h_I, f_I)]]$ $= [[F_b(l, h_I, f_I)]] \oplus [[0]]$. Otherwise, set it be $[[0]]$. $\xrightarrow{[[\mathbf{U}_b]]}$	Update $[[\mathbf{N}]]$ by $[[N(l, h_I, f_I)]] \leftarrow [[N(l, h_I, f_I)]] \ominus [[U_b(l, h_I, f_I)]]$.	

it is hard for $\mathcal{K}$ to practically eavesdrop all communications between all secondary users and the server, or that between all incumbent users and the server.

In the new scheme, $\mathcal{K}$ needs to generate $[[\mathbf{Y}_b]]_{pk_b}$ for the target user SU$_b$. So, he only needs to eavesdrop the communications between the target user and the server to obtain $[[\mathbf{F}_b]]$, $[[\mathbf{U}_b]]$. He then recovers $\mathbf{F}_b$, $\mathbf{U}_b$, and checks that

$$U_b(l, h_I, f_I) = F_b(l, h_I, f_I), \forall\ (l, h_I, f_I).$$

If true, $\mathcal{K}$ can decide that the user SU$_b$ is authorized. Clearly, he can also recover the operation data $\mathbf{R}_b$ from the tapped data $[[\mathbf{R}_b]]$. Thus, the scheme cannot truly prevent the Distributor $\mathcal{K}$ from knowing the data.

◇ *The new scheme is too inefficient to implement practically.* As we see, the computations in the plain scheme can be restricted to an upper bound w, say $w = 2^{40}$, because only the usual multiplications are involved. The new scheme needs to perform lots of modular exponentiations with the modulus n^2, where n is of 2048 bits. The working parameter is of 4096 bits. It is very time-consuming because one modular exponentiation almost takes 0.0156 second (on PC, Intel(R) Core(TM) i7-479 CPU 3.60GHz, RAM 4.00GB). Note that the working parameter for RSA cryptosystem is of 2048 bits. Moreover, RSA is only used for encrypting session keys (invoked by the subsequent symmetric key encryption, such as AES), instead of any practical message.

5 Conclusion

We show that the Dou *et al.*'s scheme fails to prevent the Key Distributor from knowing users' operation data, although it can prevent the Server from knowing the data. It is still a challenge to efficiently and systematically combine homomorphic encryption into secure multi-party computation. We would like to stress that the Paillier cryptosystem is a public key encryption which seems unsuited to directly blinding any information data because of its huge working modulus.

Acknowledgements

We thank the National Natural Science Foundation of China (project 61411146001). We are grateful to the reviewers for their valuable suggestions.

References

[1] B. Bahrak and *et al.*, "Protecting the primary users' operational privacy in spectrum sharing," in *Proceedings of IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN'14)*, pp. 236–247, Apr. 2014.

[2] P. Bogetoft and *et al.*, "Secure multiparty computation goes live," in *International Conference on Financial Cryptography and Data Security*, pp. 325–343, Feb. 2009.

[3] M. M. Butt and *et al.*, "Fair dynamic spectrum management in licensed shared access systems," *IEEE Systems Journal*, vol. 13, no. 3, pp. 2363–2374, 2019.

[4] Y. H. Chen and *et al.*, "Research on the secure financial surveillance blockchain systems," *International Journal of Network Security*, vol. 22, no. 4, pp. 708–716, 2020.

[5] S.F. Chiou and *et al.*, "Cryptanalysis of the mutual authentication and key agreement protocol with smart cards for wireless communications," *International Journal of Network Security*, vol. 21, no. 1, pp. 100–104, 2019.

[6] M. A. Clark and K. Psounis, "Trading utility for privacy in shared spectrum access systems," *IEEE/ACM Transactions on Networking*, vol. 26, no. 1, pp. 259–273, 2018.

[7] Y. Z. Dou and *et al.*, "Preserving incumbent users' privacy in exclusion-zone-based spectrum access systems: Poster," in *Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking (MobiCom'16)*, pp. 473–474, Oct. 2016.

[8] Y. Z. Dou and *et al.*, "P^2-SAS: Privacy-preserving centralized dynamic spectrum access system," *IEEE Journal of Selected Areas in Communications*, vol. 35, no. 1, pp. 173–187, 2017.

[9] Z. Y. Gao and *et al.*, "Location privacy in database-driven cognitive radio networks: Attacks and countermeasures," in *Proceedings IEEE INFOCOM*, pp. 2751–2759, Apr. 2013.

[10] L. C. Huang, C. H. Chang, and M. S. Hwang, "Research on malware detection and classification based on artificial intelligence," *International Journal of Network Security*, vol. 22, no. 5, pp. 717–727, 2020.

[11] B. A. Jayawickrama and *et al.*, "Scalable spectrum access system for massive machine type communication," *IEEE Network*, vol. 32, no. 3, pp. 154–160, 2018.

[12] X. C. Jin and *et al.*, "Safedsa: Safeguard dynamic spectrum access against fake secondary users," in *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security*, pp. 304–315, Oct. 2015.

[13] M. Karimi, S. Sadough, and M. Torabi, "Robust spectrum access for hybrid interweave-underlay cognitive radio systems using probabilistic spectrum access," *IET Signal Processing*, vol. 13, no. 9, pp. 806–813, 2019.

[14] P. Martins, L. Sousa, and A. Mariano, "A survey on fully homomorphic encryption: An engineering perspective," *ACM Computing Surveys*, vol. 50, no. 6, pp. 83:1–83:33, 2018.

[15] P. Paillier, "Public-key cryptosystems based on composite degree residuosity classes," in *International Conference on the Theory and Applications of Cryptographic Techniques*, pp. 223–238, May 1999.

[16] H. T. Pan, H. W. Yang, and M. S. Hwang, "An enhanced secure smart card-based password authentication scheme," *International Journal of Network Security*, vol. 22, no. 2, pp. 358–363, 2020.

[17] P. Thakur and *et al.*, "Advanced frame structures for hybrid spectrum access strategy in cognitive radio communication systems," *IEEE Communications Letters*, vol. 21, no. 1, pp. 410–413, 2017.

[18] C. H. Wei, M. S. Hwang, and A. Chin, "A secure privacy and authentication protocol for passive RFID tags," *International Journal of Mobile Communications*, vol. 15, no. 3, pp. 266–277, 2017.

[19] A. Yahyaoui and M. Kettani, "An efficient fully homomorphic encryption scheme," *International Journal of Network Security*, vol. 21, no. 1, pp. 91–99, 2019.

[20] K. C. Zeng, S. K. Ramesh, and Y. L. Yang, "Location spoofing attack and its countermeasures in database-driven cognitive radio networks," in *Proceedings of IEEE Conference on Communications and Network Security (CNS'14)*, pp. 202–210, Oct. 2014.

Lihua Liu, associate professor, with Department of Mathematics at Shanghai Maritime University, received her PhD degree in applied mathematics from Shanghai Jiao Tong University. Her research interests include combinatorics and cryptography.

Xinyuan Cao is currently pursuing her bachelor degree from School of Business, East China University of Science and Technology. Her research interests include E-business and marketing management.

Performance Evaluation Method of Location Privacy Protection Algorithm for Internet of Vehicles

Peng-Shou Xie, Xin-Yu Zhang, Xin Tong, Yi-Fan Wang, Tao Feng, and Yan Yan
(Corresponding author: Xin-yu Zhang)

School of Computer and Communications, Lanzhou University of Technology, 287 Lan-gong-ping Road, Lanzhou,
Gansu 730050, China
Email: kayla815@163.com

(Received Sept. 30, 2020; Revised and Accepted Apr. 16, 2021; First Online Oct. 19, 2021)

Abstract

On the Internet of vehicles environment, according to the demand of dynamic balance between privacy protection degree and location-based service quality, it is urgent to verify and optimize evaluation indexes and strategies and consider a universal performance evaluation method of location privacy protection algorithm. In this paper, we propose a performance evaluation method named DSFS-PEA based on the fusion distance, Jaccard similarity, and Hellinger divergence combined with the size of an anonymous set to quantify the performance indexes of the algorithm and the weighted idea of information entropy and private gain. Furthermore, three k-anonymity derived location privacy protection algorithms (PPA, P2P-IS-CA-HL, and SCAPGID) are re-evaluated. The experimental results show that the proposed method is better than those traditional methods in feasibility and effectiveness, which provides the theoretical basis and technical support for further researchers to select and improve appropriate privacy protection schemes.

Keywords: Internet of vehicles; Location Privacy Protection Algorithm; Performance Evaluation; Privacy Gain; Privacy Quantification

1 Introduction

With the popularization and development of the Internet and the continuous update of information technology, Internet of vehicles (IoV) technology emerges as the times require which represents the paradigm evolution from the vehicle ad hoc network (VANET) supported by cloud computing to the Internet of things. It is composed of the vehicle with the ability of perception and communication on urban road network, the roadside unit (RSU) and the back-end server. It is also an infinitely distributed VANET and characterized by high mobility of communication nodes and rapid change of network topology [11]. Location based services (LBS) is a kind of value-added service based on geographic information system platform that obtains the location information of mobile terminal users through the radio communication network of Telecom mobile operators, GPS or other external positioning methods. While enjoying the convenience it brings, the number of users, complexity and management difficulty of the whole system also increases greatly. The data security and privacy protection under the IoV environment have become the focus of attention. LBS usually require vehicle users to report the location information of continuous road sections to establish communication with RSU and obtain instant traffic information like road conditions. The vehicle is in an open physical space, and the disclosure of driver's identity information, license plate number, location and trajectory may threaten the safety of drivers and passengers' life and property [7, 14].

Ardagna proposes a LBS privacy protection system based on peer to peer (P2P) mode [15], in order to meet the requirements of precise LBS and users' privacy protection. The main concept of the system is to project the longitude and latitude of the original space into the coordinate space, and realize the position disturbance by Hilbert algorithm. Symmetric key encryption, packet fragmentation, random probability forwarding and other technologies are also used to ensure the users' location privacy.

Fu Tianxia has established a privacy protection algorithm privacy protection algorithm (PPA) based on P2P structure, k-anonymity and pseudonym technology [21], aiming at the privacy security protection problem of high-speed movement of nodes in the IoV. The network can be expanded to hide nodes more than onefold increase, which can achieve a good balance between the effect of privacy protection and the result of region selection, which can not only enhance the security intensity, but also obtain high quality services.

In order to further improve the privacy protection algorithm, the performance evaluation indexes have been analyzed to a certain extent. However, most of them are aimed at specific applications, attack models (such as replay attack, message modification and generation attack, denial attack, simulation attack, location tracking attack, *etc.*) and privacy threats [10], which are difficult to be generalized. It means that the algorithms lack universal applicability in evaluation. Hassan Sarmadi and other researchers have proposed an extensible algorithm to measure the distance between different data points in k-nearest neighbor (k-NN) classifier and k-means clustering to learn Mahalanobis distance measure from a set of labeled train-ing samples [8]. By using the principle of maximum margin, metric learning is reduced to a convex optimization problem, which keeps the semi positive definiteness of matrix variables. Bag Sujoy and other researchers have established a Jaccard similarity model [16] by considering the rating vector of all users, classified the relevant neighborhood and generated recommendations with lower operation time to improve the accuracy of the recommendation system. Paul Gardner and other researchers have applied probabilistic modeling methods to engineering applications, compared the f-divergence with integrated probability measures (IPMs) used to quantify the difference of distribution, and confirmed that Hellinger divergence is more highly interpretable than Kullback Leibler divergence [12].

For the integration of evaluation indexes, we can refer to Shannon information entropy to determine the weight of secondary privacy elements, and then calculate the quantified privacy of each record in the data set under the primary privacy elements, and use BP neural network to output the classification results of privacy data without preset measurement weights. In order to accelerate the retrieval speed and improve the query accuracy from the massive information resources, a new method has been proposed to calculate the weight of feature words based on information gain and information entropy. The results show that it can effectively improve the shortcomings of the traditional TFIDF method, and it is better than other methods of text classification in the accuracy rate, recall rate and F-measure.

To sum up, researchers have made some achievements in this field, since there is no unified performance evaluation method and standard for the privacy protection technology of Internet of vehicles, so it is necessary to develop a set of universal evaluation index and evaluation system to objectively and reasonably evaluate the anonymity technology. However, it also prove that it is feasible to introduce quantification methods based on distance (D), similarity (S), f-divergence (F) and scale (S) into the security measurement model. The privacy evaluation algorithm (PEA) which is consisted by the above four quantitative methods is named DSFS-PEA, and makes up for the shortcomings of performance evaluation model, performance index quantification method and performance evaluation algorithm of privacy protection technology of IoV.

2 Performance Evaluation Model of Privacy Protection Algorithm

Similar to the computer network architecture, the Internet of vehicles system also defines the protocols and main functions of each layer according to the hierarchical method. Because it originates from MANETs [1, 9] and has the relevant characteristics of MANETs, it also follows the three-tier architecture of perception layer, network layer and application layer, which can summarize the Internet of vehicles system and performance indexes framework, as shown in Figure 1.

- Perception layer: The sensor nodes arranged in the environment for data acquisition are easy to be modified for malicious use, and steal the location privacy and other sensitive information of communication vehicles. Usually, mechanisms such as identity authentication and security management are used to avoid privacy disclosure. Among them, multi-sensor data fusion is to verify the authenticity of data and detect abnormal nodes.
- Network layer: It is mainly divided into three modules: access, transfer and service. This paper will not discuss the security risks of malicious node attack. It is necessary to ensure the accuracy, integrity and availability of the information for the effective transmission of the collected data in transfer module; Location services such as the query of interest point belong to the category of service module.
- Application layer: It involves the storage and calculation of a large number of users' privacy information, and many kinds of services which need technical support such as big data processing and business management.

The Internet of vehicles system and performance indexes framework refer to the three-tier architecture model of the Internet of things, which can make full use of the indicators under the existing cloud service computing. It begins from the authoritative national standardization rules and network security risk assessment standards such as CC, BS7799 and ISO/IEC21827-2002 (SSE-CMM), *etc.* [2, 3] The universal performance evaluation indexes of privacy protection algorithm of Internet of vehicles,namely privacy, service quality and cost, are separated out.

3 Quantization Method of Performance Evaluation Indexes for Privacy Protection Algorithm

This chapter mainly discusses the privacy degree and service quality of the location privacy evaluation in-

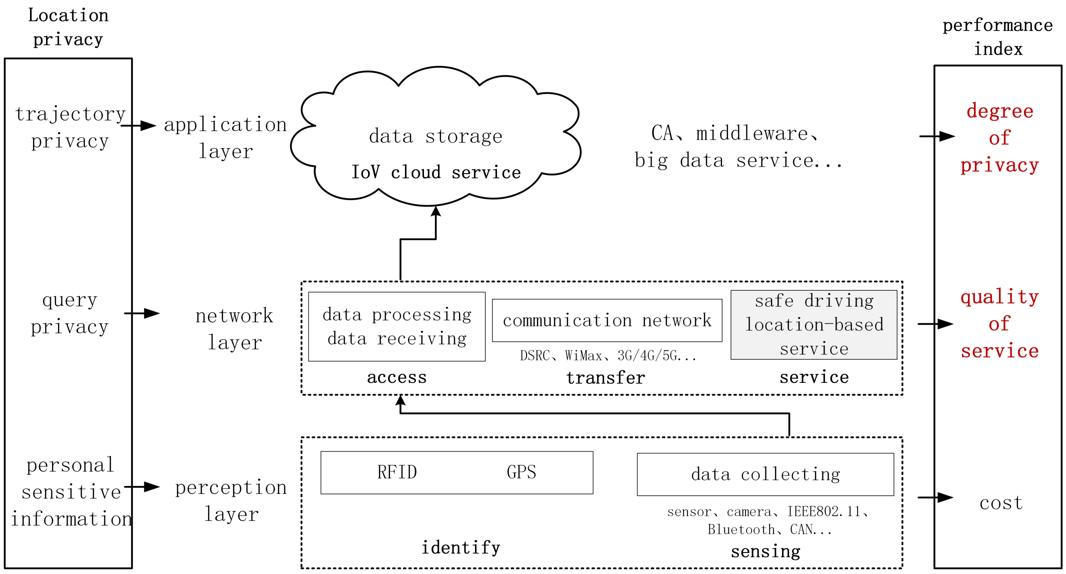


Figure 1: Framework of Internet of vehicles system and performance indexes

dexes in the above mentioned Internet of vehicles framework [20, 22]. Due to the correlation between the two indexes, the same method is proposed to be used for quantification, as shown in Figure 2, and the weight matrix is formed by distinguishing the quantized value from the forward and backward.

The quantitative methods to measure the differences among individuals mainly include the quantification based on distance, similarity, f-divergence and scale.Through comparative analysis, the fusion distance, Jaccard similarity, Hellinger divergence and average treatment method are used to quantitatively describe the differences. Thus, the accuracy, integrity, availability of data and the influence of objective factors on privacy and service quality can be quantified.

3.1 Distance Based Quantization

Mahalanobis distance [18] is defined by the statistician Mahalanobis to represent the covariance distance of data. It can effectively calculate the similarity of two unknown sample sets and is not affected by the dimension. The correlation between variables is fully considered. The Equation (1) of Mahalanobis distance from vector X to vector Y is as follows:

$$d\left(X,Y\right)=\sqrt{\left(X-\overline{Y}\right)^{T}\sum\nolimits_{Y}^{-1}\left(X-\overline{Y}\right)} \qquad (1)$$

Among them,

$$\overline{Y}=\sum_{j=1}^{n}Y_{ij}/n$$

$$\sum=Cov\left(X,Y\right)=E\left[\left(X-E\left(X\right)\right)\left(Y-E\left(Y\right)\right)\right]=$$
$$\begin{bmatrix} Cov\left(x_1,y_1\right) & Cov\left(x_2,y_1\right) & \cdots & Cov\left(x_1,y_j\right) \\ Cov\left(x_2,y_1\right) & Cov\left(x_2,y_2\right) & \cdots & Cov\left(x_2,y_j\right) \\ \vdots & \vdots & \ddots & \vdots \\ Cov\left(x_j,y_1\right) & Cov\left(x_j,y_2\right) & \cdots & Cov\left(x_j,y_j\right) \end{bmatrix} \qquad (2)$$

In Equation (2), X is the clustering sample; Y is the real location space assembly of the sample, which is also called Mahalanobis space; Y and $\sum$Y are the mean matrix and covariance matrix of sample Y respectively; when $\sum$ is the expectation matrix of covariance matrix of variable Y and it is the unit matrix, Mahalanobis distance is simplified to Euclidean distance.

$$\sum=Cov\left(X,Y\right)=E\left(X-E\left(X\right)\right)\left(Y-E\left(Y\right)\right)$$

$$\rho_{XY}=\frac{Cov\left(X,Y\right)}{\sqrt{D\left(X\right)}\cdot\sqrt{D\left(Y\right)}} \qquad (3)$$

In Equation (3), ρ_{XY} is the correlation coefficient of X and Y, the larger $|\rho_{XY}|$ is, the higher the correlation degree of X and Y is, and it is a directly proportional between ρ_{XY} and $Cov\left(X,Y\right)$.

Euclidean distance is a distance metric function. Its computational complexity is $O\left(d\right)$ and has the characteristics of space rotation invariance, Equation (4) is as follow:

$$d\left(X,Y\right)=\sqrt{\sum_{i=1}^{n}\left(x_i-y_i\right)^2} \qquad (4)$$

The Euclidean distance method is used to analyze the difference of multi-dimensional values among individuals. It is necessary to ensure that each dimension index is at the same calibration level, and X and Y are the real and fuzzy location of vehicle nodes respectively, and the Euclidean distance between location coordinates x_i and y_i

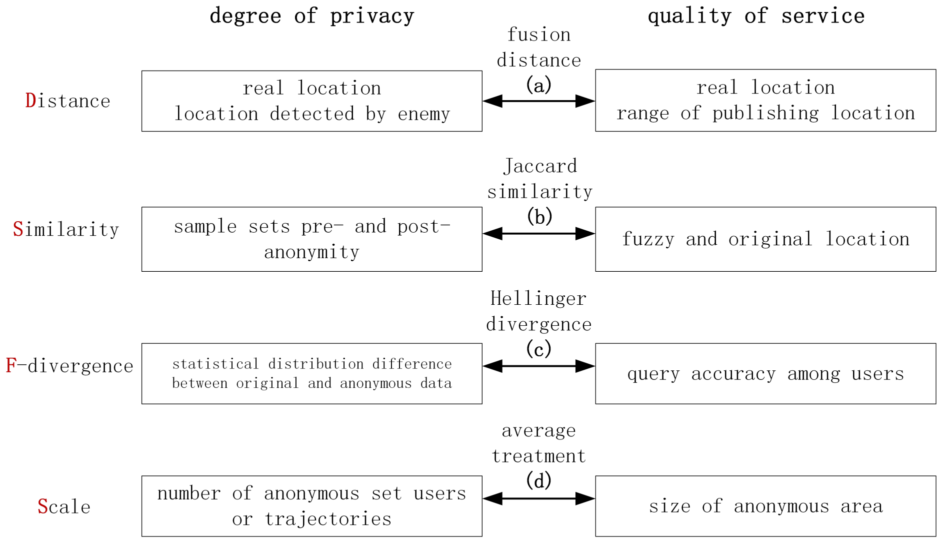

degree of privacy
quality of service
Distance
real location
location detected by enemy
fusion
distance
(a)
real location
range of publishing location
Similarity
sample sets pre- and post-
anonymity
Jaccard
similarity
(b)
fuzzy and original location
F-divergence
statistical distribution difference
between original and anonymous data
Hellinger
divergence
(c)
query accuracy among users
Scale
number of anonymous set users
or trajectories
average
treatment
(d)
size of anonymous area

Figure 2: Corresponding evaluation method of privacy degree and service quality

is defined to measure the privacy in the Internet of vehicles. In addition, in the environment of high-mobility and instantaneity of IoV, the single use of Euclidean distance as a measurement method may make the actual meaning of the dimension between variables difficult to explain, and also ignore the distribution of samples and unable to explain the correlation between multiple data.

A "fusion distance" is proposed for solving the limitation of the above two distance measurement methods, which takes into account the correlation and independence of the characteristic variables. Then the equation of the fusion distance of Mahalanobis distance and Euclidean distance is as follows:

$$d_{Mix} = \omega \times MD\left(X, Y\right) + \left(1 - \omega\right) \times ED\left(X, Y\right) \tag{5}$$

$$\omega = \sqrt{1 - |C_Y|} \tag{6}$$

In Equations (5) and (6), $MD\left(X, Y\right)$ is the Mahalanobis distance from vector X to vector Y; $ED\left(X, Y\right)$ is the Euclidean distance from vector X to vector Y; C_Y is the correlation coefficient matrix of vector Y; $|\cdot|$ is the value of matrix determinant.

3.2 Similarity Based Quantization

Data accuracy [4] means that the evaluation of anonymity method should follow the principle of minimum information loss, that is, the smaller the difference between the original data distribution and the statistical distribution of anonymous data is, the higher the similarity is, the higher the accuracy of anonymous data is. As an index to measure the extent of similarity described from the direction between individuals, cosine similarity is the cosine value of the angle between two vectors in the vector space, as shown in Equation (7):

$$sim\left(X, Y\right) = \cos\theta = \frac{X \cdot Y}{\|X\| \cdot \|Y\|} \tag{7}$$

Let X and Y be the sample sets pre- and post-anonymity, the set of multiple location nodes can be regarded as a trajectory, and the trajectory coincidence degree is used to evaluate.

Comparing the similarities and differences between original sample sets, the larger the value of Jaccard coefficient is, the higher the similarity of samples is. If we compare the Jaccard similarity coefficient of dataset X and Y, Equation (8) is as follows:

$$Jaccard\left(X, Y\right) = \frac{|X \cap Y|}{|X \cup Y|} = \frac{|X \cap Y|}{|X| + |Y| - |X \cap Y|} \tag{8}$$

Similarly, the simple use of Jaccard similarity between entities as a measure of service quality has two defects: it may be identified vehicle terminals with multiple location in formation as different individuals or ignore the weight information.

3.3 F-Divergence Based Quantization

The f-divergence based measurement method can accurately locate the probability difference of a designated record in the original dataset and the anonymous dataset through the same attribute value [13]. The former two quantization methods are described in terms of distance measurement and direction, but not applicable to the difference between statistical distributions of random variables. Therefore, the concept of f-divergence in probability theory is introduced. The commonly used discrete form of f-divergence is shown in Table 1.

Suppose that there are two probability distributions P and Q in the probability space Ω, then the f-divergence between P and Q is:

$$D_f\left(P\| Q\right) = \int_{\Omega} f\left(\frac{dP}{dQ}\right) dQ \tag{9}$$

Table 1: Common f-divergence

Name	Function	The f-divergence of discrete variables$D_f\left(P\left(x\right)\left\|Q\left(x\right)\right.\right)$
Kullback-Leibler	$x\ln x$ or $-\ln x$	$\sum\limits_{i=1}^{N} p_i \ln\left(\frac{p_i}{q_i}\right)$ or $\sum\limits_{i=1}^{N} q_i \ln\left(\frac{q_i}{p_i}\right)$
χ^2	$(x-1)^2$ or x^2-1	$\sum\limits_{i=1}^{N} \frac{(p_i-q_i)^2}{q_i}$ or $\sum\limits_{i=1}^{N} \frac{{p_i}^2-{q_i}^2}{q_i}$
Hellinger	$\frac{1}{2}(\sqrt{x}-1)^2$ or $1-\sqrt{x}$	$\frac{1}{2}\sum\limits_{i=1}^{N}\left(\sqrt{p_i}-\sqrt{q_i}\right)^2$ or $1-\sum\limits_{i=1}^{N}\sqrt{p_iq_i}$
Jensen-Shannon	$[\ln\left(\frac{2}{1+x}\right)+xln\left(\frac{2x}{1+x}\right)$	$\sum\limits_{i=1}^{N}\left[p_i\ln\left(\frac{2p_i}{p_i+q_i}\right)+q_i\ln\left(\frac{2q_i}{p_i+q_i}\right)\right]$
Jeffrey	$(1-x)\ln\frac{1}{x}$	$\sum\limits_{i=1}^{N}(p_i-q_i)\ln\frac{p_i}{q_i}$
Total variation	$\sum\limits_{i=1}^{N}(p_i-q_i)\ln\frac{p_i}{q_i}$	$\sum\limits_{i=1}^{N}\left\|p_i-q_i\right\|$

Since divergence is the general term of a series of functions, f-divergence includes Kullback Leibler divergence, total variation divergence, Hellinger divergence, Jensen Shannon divergence, Jeffrey divergence, exponential divergence, *etc.* Hellinger divergence is selected as the divergence function for evaluation by comparing the measurability, convergence and sensitivity. In practical application, the definition of f-divergence is usually transformed into discrete form. For discrete distribution of Hel-divergence, the definition Equation (10) is as follows:

$$D_{Hel}\left(P\left(x\right)\left|Q\left(x\right)\right.\right)=1-\sum_{i=1}^{N}\sqrt{p_iq_i} \tag{10}$$

3.4 Scale Based Quantization

The privacy degree and service quality of location privacy protection technology of IoV are also related to the number of anonymous users or trajectories, the size of anonymous area and other parameters. The basic indexes can be uniformly processed with the idea of linear dimensionless [19].

Select the average treatment method :

$$x_{ij}^{*}=\frac{x_{ij}}{\overline{x_j}}, i=1,2,\ldots,m; j=1,2,\ldots,n \tag{11}$$

In Equation (11), x_{ij} is the original data of the j-th index of the i-th sample; x_{ij}^{*} is the processed data, between 0 1, the numerical distribution is consistent with that before processing; $\overline{x_j}$ is the average value of the original data of the j-th index. This method can keep the whole consistency of the original data well, and also keep the information of the variation degree of the indexes while eliminating the influence of dimension and order of magnitude.

4 Privacy Protection Technology Evaluation Algorithm DSFS-PEA

The non-uniqueness of evaluation index system is determined by diversity of evaluation subject, evaluation object and evaluation scale. Domestic and international researchers have designed a number of strategies to evaluate the performance of privacy protection algorithms to determine the weight of performance evaluation indexes based on their different characteristics [17], including but not limited to: information entropy, game theory, classification confidence interval, fuzzy comprehensive evaluation and multiple nested fusion of them. The applicable scenes of privacy gain and entropy weight method are considered to model, and the information gain value and weight matrix of the index are matched with each other, finally the comprehensive evaluation values of privacy and service quality are calculated.

4.1 Process Description

According to the algorithm steps described above, the pseudo code of the algorithm DSFS-PEA is as follow:

Algorithm 1 Pseudo code for DSFS-PEA

```
Input: the real location $X$ of the user entity and location $Y$ after anonymization of evaluated privacy protection algorithm;
Output: evaluation value of privacy protection degree and service quality performance index $E_p, E_q$;
$X \leftarrow [x_1, x_2, \ldots, x_i, \ldots, x_n]$;
$Y \leftarrow [y_1, y_2, \ldots, y_i, \ldots, y_n]$;
while $x_i, y_i$ are two-dimensional position coordinates do
    calculate $eucliDist(X, Y)$ and $mhalaDist(X, Y)$;
    calculate $\omega \leftarrow \rho_{XY}$ generate correlation coefficient matrix $C_Y$;
    calculate $MixDist(X,Y)$; Jaccard similarity; Helligence divergence;
    The number of anonymous users and trajectory data, the size of anonymous area are averaged average $x_{ij}^{*}$;
        Elements a←$MixDist(X,Y)$; b←Jaccard similarity; c←Helligence divergence; d← $x_{ij}^{*}$;
        Distinguish the positive and negative indexes and construct the weight matrix $(r_{ij}{}')_{m\times n}$ by a,b,c,d;
        calculate index entropy $S_{12}$; Represent the amount of information;
        // contains the degree of privacy protection and the level of service quality
        calculate privacy gain of vehicle terminal $v_k$ in road network $IG\left(l,v_k\right)$;
    print weight based on privacy gain and information entropy $w_{ik}$;
    round sum([elements[j]*weights[j]
    for j in range(m));
end while
return $E_p, E_q$;
```

4.2 Weight Calculation

The related concepts of information entropy are defined as follows:

Definition 1. *For N messages with the same probability, and the probability of each message is $1/N$, then the amount of information carried by each message is:*

$$-\log p = \log\left(1/N\right) \tag{12}$$

Definition 2. *For a given probability distribution $P = (p_1, p_2, \ldots p_n)$, the amount of information carried by the distribution is called the entropy of P, Equation (13) is as follows:*

$$\begin{aligned} I(P) &= -(p_1 \times \log_2 p_1 + p_2 \times \log_2 p_2 + cldots \\ &\qquad +p_n \times \log_2 p_n) \\ &= -\sum_{k=1}^{n} p_k \times \log_2 p_k \end{aligned} \tag{13}$$

When P is between $(0.5, 0.5)$, $I(P)$ is 1; If P is between $(1, 0)$, $I(P)$ is 0. It can be seen that adjusting the weight equation through information entropy can overcome the defect of neglecting the distribution within class in the weight calculation

The dimensions of each index in the system are not necessarily the same, and sometimes the order of magnitude of the values is completely different. These data are difficult to be directly compared, so it is necessary to normalize the original data. The DSFS, four quantitative results, can be divided into two categories: distinguish the positive and negative indexes.

Let the original evaluation information matrix be:

$$\left(r'_{ij}\right)_{m\times n}, i = 1,2,\ldots,m; j = 1,2,\ldots,n \tag{14}$$

The normalization processing method for each index in the normative matrix is as follows:
For those positive indexes:

$$r_{ij}^{\max} = \frac{r'_{ij} - \min\limits_{j}\left\{r'_{ij}\right\}}{\max\limits_{j}\left\{r'_{ij}\right\} - \min\limits_{j}\left\{r'_{ij}\right\}} \tag{15}$$

For other negative indexes:

$$r_{\mathrm{ij}}^{\min} = \frac{\max\limits_{j}\left\{r'_{ij}\right\} - r'_{ij}}{\max\limits_{j}\left\{r'_{ij}\right\} - \min\limits_{j}\left\{r'_{ij}\right\}} \tag{16}$$

In Equations (15) and (16), $\max\limits_{j}\left\{r'_{ij}\right\}$ and $\min\limits_{j}\left\{r'_{ij}\right\}$ are the maximum and minimum values of row i in the matrix.

Calculate the index value proportion of the j object under the i index:

$$P_{ij} = \frac{r_{ij}}{\sum\limits_{j=1}^{n} r_{ij}} \tag{17}$$

The entropy value of the i-th index is calculated by entropy weight method:

$$S_i = -\alpha \sum_{j=1}^{n} P_{ij} \ln P_{ij} \tag{18}$$

In Equation (18), $\alpha = \frac{1}{\ln n}$ and regulate when $P_{ij} = 0, P_{ij} \ln P_{ij} = 0$.

On this basis, the concept of location privacy gain of IoV has been introduced. The degree of privacy improvement of original location information after anonymous processing is that the probability difference of a designated record can be accurately located by the same attribute value between the original sample set and anonymous sample set. The equation of information gain applied in this paper is as follows:

$$\begin{aligned} IG(k) &= H(L) - H(L/k) \\ &= -\sum_{l\in L} p(l)\log p(l) + p(k)\sum_{l\in L} p(l/k)\log P(l/k) \\ &\qquad + p(\bar{k})\sum_{l\in L} p(l/\bar{k}) \\ &= \sum_{l\in L}(P(l,k)\log(\frac{P(l,k)}{P(l)P(k)}) \\ &\qquad + P(l,\bar{k})\log(\frac{P(l,\bar{k})}{P(l)P(\bar{k})})) \end{aligned} \tag{19}$$

Among them, $IG(k)$ represents the information gain value of the vehicle terminal location before and after obtaining the privacy degree K of the algorithm, l represents the location coordinate class variable, L represents the location set, it also has $L = (l_1, l_2, \ldots, l_i, \ldots, l_n)$. $H\left(L\right)$ represents the entropy of the probability space where the random location of the previous anonymous entity belongs to a certain category, and $H\left(L/k\right)$ is the entropy of the probability space that the real location belongs to a certain category after the privacy degree K is obtained by using the corresponding anonymization algorithm.

Therefore, we can get the weight equation based on privacy gain and information entropy:

$$w_{ik} = IG\left(L, v_k\right) \times S_{\mathrm{i}} \tag{20}$$

Among them,

$$IG\left(\mathrm{v}_k\right) = \sum_{l\in L}\left(P\left(l, v_k\right)\log\left(\frac{P\left(l, v_k\right)}{P\left(l\right)P\left(v_k\right)}\right) + P\left(l, \overline{v_k}\right)\log\left(\frac{P\left(l, \overline{v_k}\right)}{P\left(l\right)P\left(\overline{v_k}\right)}\right)\right) \tag{21}$$

w_{ik} is the weight of the evaluation index of the location privacy protection technology of IoV, $IG\left(l, v_k\right)$ represents the privacy gain value of the vehicle terminal v_k in the road network, and S_i represents the information entropy weighting factor distributed within the class.

5 Experiment and Result Analysis

There is almost an inversely proportional relationship between accuracy of LBS and the quality of location-based privacy protection. According to the privacy protection needs of users in different contexts, combined with the performance evaluation index and evaluation strategy of the improved algorithm, PPA based on private security service model (PSSM) and LBS is compared with original P2P algorithms.

5.1 Experimental Environment

The processor is Intel64 Family 6 Model 69 Stepping 1 Genuine Intel 759Mhz, which under the 4.0GB RAM experimental environment, and enable virtual machine monitor mode is extended in firmware. Using Python3 compiling environment CW-KNN and installing Ubuntu on VMware virtual machine, the corresponding road network environment is built for experimental simulation.The vehicle movement model generated in sumo is loaded into ns-2 in the form of trace file to obtain vehicle nodes. By simulating a total of 6 two-way straight lanes with intersections, the node distribution density is adjusted within the coverage range of 1500m*1500m RSU signal, and the location privacy protection degree and location service quality are discussed. The experimental process is divided into two parts: traffic scene and network communication. The related experimental configuration parameters are shown in Table 2.

Table 2: Experimental configuration parameters

Module category	**Parameter**	**Value**
Traffic scene	Quantity of lane	6
	Link length/m	1500
	Lane width/m	3.5
	Vehicle speed/(m/s)	10
	Quantity of vehicle	100,200,300, 400,500
Network communication	Failure time of cache records Δt/s	10
	Value of K	60,70,80,90,100

Using DSES-PEA, the evaluation algorithm proposed in this paper, the algorithm performance of improved P2P privacy protection algorithm PPA based on PSSM proposed by researcher Fu Tianxia is evaluated, and the evaluation results of the effectiveness and authenticity of PPA are analyzed to verify the evaluation algorithm inversely. Peer to Peer-Information Sharing-Cloaked Area-Historical Location (P2P-IS-CA-HL) algorithm [6] and Spatial Cloaking Algorithm Based on P2P and Grid ID (SCAPGID) algorithmc [5] introduced by conferences are compared by conducting the experiment at the same time. Among them, P2P-IS-CA- HL is an algorithm for sharing node information and adjusting region in anonymous area; SCAPGID is an algorithm that divides the plane by grid, completes location anonymity through dynamic expansion of grid and cooperation of users in the same grid.

5.2 Result Analysis

In each group of experiments, we compare the number of vehicles [100, 500], the evaluation value of privacy protection degree and location-based service quality E_p, E_q are in [0, 1] interval. By analyzing the changes of the three privacy protection algorithms in Figure 3 and Figure 4, the effectiveness and usability of the evaluation algorithm DSFS-PEA are verified. It can be seen that there is a certain inverse proportional relationship between privacy protection degree and location-based service quality, that is, the optimal value cannot be reached at the same time, and the balance point can only be found according to the privacy needs of users and the amount of information contained in the indexes.

According to the privacy gain $IG\left(l, v_k\right) \approx 0.82$, when the number of vehicles is [200, 300], the comprehensive evaluation value of P2P-IS-CA-HL algorithm($E_p \in \left[0.167, 0.289\right]; E_p \in \left[0.270, 0.401\right]$), is better than that of SCAPGID($E_p \in \left[0.177, 0.197\right]; E_p \in \left[0.293, 0.325\right]$),and PPA has obvious advantages unless the sample value is very small, which basically meets the needs of improvement. The reason is that in the traffic scene with the same parameters, the more vehicles there are, the greater the number of assistance neighbor nodes in the anonymous region is, which improves the anonymous success rate of PPA algorithm. When the number of vehicles increases to an aggressive scale, the minimum anonymous area of the algorithm cannot meet the anonymous protection degree, and the anonymous area needs to be expanded by

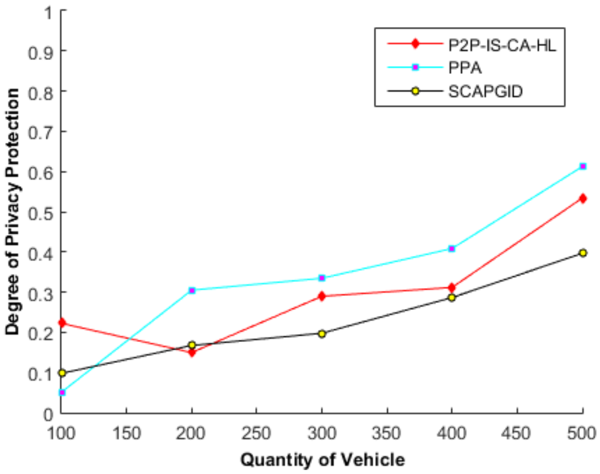


Figure 3: Degree of privacy protection varies with quantity of vehicle

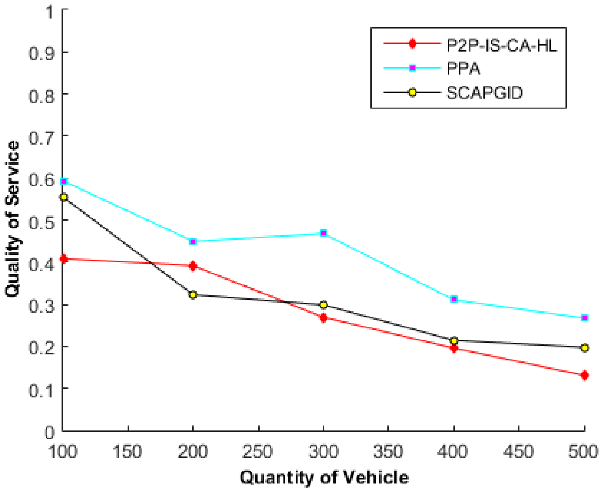


Figure 4: Quality of service varies with quantity of vehicle

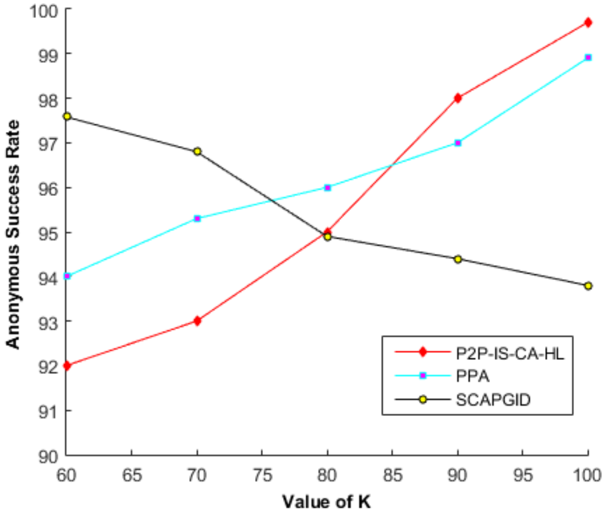


Figure 5: Anonymous success rate varies with value of K

two other types of algorithms to stabilize the whole communication network.

In addition to the number of vehicles, the anonymity success rate of privacy protection algorithm is also related to the value of anonymity degree K. Figure 5 depicts the change of anonymity success rate with the value of K, and the number of vehicles is set to 200. When K increases, the vehicle terminal needs to recruit more peers to help collect enough peer location information to make its location fuzzy. The amplification scheme of pseudonym node and anonymous region in PPA can effectively avoid network partition and other problems, so it shows a steady rising state in the figure.

To sum up, the simulation results meet the characteristics of different algorithms in a specific environment, and the evaluation algorithm DSFS-PEA has universal applicability for location privacy protection algorithm based on the generation of anonymous locations, and can provide better forward research and technical support for improving the algorithm.

6 Conclusion

A unified performance evaluation method and standard for privacy protection technology of IoV is not proposed at present, so it is necessary to develop a set of evaluation index and system to objectively and reasonably evaluate the anonymity technology. Metrics methods based on distance, similarity, divergence and scale are proposed in this paper which also uses the concepts of privacy gain and information entropy to calculate the weight of each index. The follow-up research will focus on the ment of the weight method. The noise addition method in differential privacy can be introduced to adjust the weight of the performance index, while the PPA can adapt to the privacy protection requirements of different orders of magnitude.

Acknowledgement

This research is supported by the National Natural Science Foundations of China under Grants No.61862040, No.61762059 and No.61762060. The authors gratefully acknowledge the anonymous reviewers for their helpful comments and suggestions.

References

[1] T. Alam, “Internet of things: A secure cloud-based manet mobility model,” *International Journal of Network Security*, vol. 22, no. 3, pp. 516–522, 2020.

[2] J. Antoniou, “Quality of experience and emerging technologies: Considering features of 5G, IoT, cloud and AI,” in *Innovations in Communication and Computing*, pp. 1–8, 2020.

[3] J. M. Batalla, E. Andrukiewicz, and *et al.*, "Security risk assessment for 5G networks: National perspective," *IEEE Wireless Communications*, vol. 27, no. 4, pp. 16–22, 2020.

[4] B. Wan and C. Xu, "Intelligent evaluation model of precision mobile network software evaluation accuracy," in *IEEE 3rd International Conference on Information Systems and Computer Aided Education (ICISCAE'20)*, pp. 274–277, 2020.

[5] H. Che, Y. He, and J. Liu, "Research on location anonymity algorithm based on P2P and grid ID," *Information Network Security*, vol. 12, no. 3, pp. 28–42, 2015.

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

[7] D. Das, "Improving throughput and energy efficiency in vehicular ad-hoc networks using internet of vehicles and mobile femto access points," in *IEEE Region 10 Conference (TENCON'19)*, pp. 1270–1274, 2019.

[8] S. Hassan and K. Abbas, "A novel anomaly detection method based on adaptive mahalanobis-squared distance and one-class KNN rule for structural health monitoring under environmental effects," *Mechanical Systems and Signal Processing*, vol. 104, no. 7, pp. 106–119, 2020.

[9] H. Riasudheen, K. Selvamani, and *et al.*, "An efficient energy-aware routing scheme for cloud-assisted manets in 5G," *Ad Hoc Networks*, vol. 97, no. 5, pp. 211–238, 2020.

[10] L. Li, L. Zhengjuan, T. Xiaohong, and S. Runhua, "A dynamic location privacy protection scheme based on cloud storage," *International Journal of Network Security*, vol. 21, no. 5, pp. 828–834, 2019.

[11] X. Li, H. Zhongyuan, and *et al.*, "Transfer learning based intrusion detection scheme for internet of vehicles," *Information Sciences*, vol. 547, no. 2, pp. 119–135, 2021.

[12] G. Paul, L. Charles, and J. B. Robert, "An evaluation of validation metrics for probabilistic model outputs," in *ASME Verification and Validation Symposium*, pp. 9327–9336, 2018.

[13] N. Ryo, "Source resolvability problem with respect to a certain subclass of f-divergence," *IEEE International Symposium on Information Theory-Proceedings*, vol. 7, no. 2, pp. 2234–2238, 2019.

[14] Z. Sahnoune and E. Aimeur, "Deloc: a delegation-based privacy preserving mechanism for location-based services," *International Journal of Mobile Communi-cations*, vol. 19, no. 1, pp. 22–52, 2021.

[15] K. A. Soroush, A. Danilo, and *et al.*, "Analytical composite performance models for big data applications," *Journal of Network and Computer Applications*, vol. 142, no. 2, pp. 63–75, 2019.

[16] S. Bag and *et al.*, "An efficient recommendation generation using relevant jaccard similarity," *Department of Industrial and Systems Engineering*, vol. 48, no. 3, pp. 53–64, 2019.

[17] G. Tilei, L. Tong, and *et al.*, "Research on cloud service security meas-urement based on information entropy," *International Journal of Network Security*, vol. 21, no. 6, pp. 1003–1013, 2019.

[18] V. Roizman, M. Jonckheere, and F. Pascal, "Robust clustering and outlier rejection using the mahalanobis distance distribution," in *The 28th European Signal Processing Conference (EUSIPCO'21)*, pp. 2448–2452, 2021.

[19] T. Wei, P. Zhisong, and *et al.*, "Primal averaging: A new gradient evaluation step to attainthe optimal individual convergence," *IEEE Transactions on Cybernetics*, vol. 50, no. 2, pp. 835–845, 2019.

[20] L. Wen, "Security evaluation of computer network based on hierarchy," *International Journal of Network Security*, vol. 21, no. 5, pp. 735–740, 2019.

[21] P. S. Xie, T. X. Fu, and H. J. Fan, "An algorithm of the privacy security protection based on location service in the internet of vehicles," *International Journal of Network Security*, vol. 21, no. 4, pp. 556–565, 2019.

[22] A. Z. Zulfiqar, H. Jingsha, and *et al.*, "Detection and prevention of jellyfish attacks using knn algorithm and trusted routing scheme in manet," *International Journal of Network Security*, vol. 23, no. 1, pp. 77–87, 2021.

Biography

Peng-shou Xie was born in Jan. 1972. He is a professor and a supervisor of master student at Lanzhou University of Technology. His major research field is Security on Internet of Things. E-mail: xiepsh_lut@163.com

Xin-yu Zhang was born in Aug. 1996. She is a master student at Lanzhou University of Technology. Her major research field is network and information security. E-mail: kayla815@163.com

Xin Tong was born in Aug. 1995. He is a master student at Lanzhou University of Technology. His major research field is network and information security. E-mail: 2505156603@qq.com.

Yi-fan Wang was born in Aug. 1996. She is a master student at Lanzhou University of Technology. Her major research field is network and information security. E-mail: 844782234@qq.com

Tao Feng was born in Dec. 1970. He is a professor and a supervisor of Doctoral student at Lanzhou University of Technology. His major research field is modern cryptography theory, network and information security technology. E-mail: fengt@lut.cn

Yan Yan was born in Oct. 1980. She is a associate professor and a supervisor of master student at Lanzhou University of Technology. Her major research field is privacy protection, multimedia information security. E-mail: yanyan@lut.cn

An Improved Sequence Cross Transformation Method and Its Application in Image Encryption

Chunming Xu and Yong Zhang
(Corresponding author: Chunming Xu)

School of Mathematics and Statistical, Yancheng Teachers University, P. R. China
No. 50, Kaifang Avenue, Yancheng 224002, P. R. China
Email: ycxcm@126.com

(Received Nov. 8, 2020; Revised and Accepted May 6, 2021; First Online Oct. 19, 2021)

Abstract

Image encryption technology plays an important role in the process of image transmission. In this paper, the shortcomings of sequence cross transformation are studied, and an improved sequence cross transformation is presented to overcome them. Then the improved sequence cross transformation method is used for image encryption. Finally, simulation experiments are tested on three classical images and are evaluated using the histogram, correlation analysis, entropy, number of pixel change rate (NPCR), and unified average change intensity (UACI). The experimental results show that the proposed method is effective and feasible.

Keywords: Chaotic System; Circshift Function; Image Encryption; Image Scrambling Effect; Sequence Cross Transformation

1 Introduction

Digital image is one of the most commonly used multimedia forms, which plays a important role in information transmission and is widely used in many aspects of social and economic life. However, in the military, commercial and medical fields, the security of digital image is an important issue. Therefore, image encryption technology has become a hot topic in the world [1, 4, 6–8, 10, 17].

The scrambling-diffusion based image encryption method is the current mainstream algorithm framework for image encryption [20, 21]. According to the model of scrambling-diffusion, image encryption is divided into two stages: Image scrambling and image diffusion. Through image scrambling, the position relationship of pixels can be changed and the correlation between adjacent pixels can be destroyed. However, after image scrambling, the gray values of all pixels are not changed, so the histogram of cipher image will not be changed. Therefore, it is necessary to change the pixel value through the image diffusion operation to get a better encryption effect.

So far, researchers have proposed a lot of image scrambling methods such as Arnold transformation [3, 11], Zigzag transformation [13, 19], Josephus transformation [5, 14], *etc.* In [15], the author proposes an image scrambling method based on sequence cross transformation. The image can be encrypted using sequence cross transformation by multi round scrambling. However, the problem with this method is that the sequence cross transformation operation is periodic. That is, after a certain number of scans, the resulting scrambling image will be changed to original image. Therefore, the scrambling method is vulnerable to attack and lacks security. To solve this problem, an improved sequence cross transformation based image scrambling and encryption algorithm is proposed in this paper. The proposed method has better scrambling effect and better security.

The rest of the paper is organized as follows. The sequence cross transformation method and its improved algorithm are given in Section 2. The proposed image encryption and decryption scheme are introduced in Section 3. Section 4 presents the experimental results and the security of the algorithm. Finally, we conclude this paper in Section 5.

2 Improved Sequence Cross Transformation

2.1 Sequence Cross Transformation

Suppose there is a sequence, we can divide it into two equal parts from the middle, then rearrange this sequence using one by one cross method [15]. As a result, we can get a new scrambled sequence. For example, if the original sequence is $1, 2, 3, 4, 5, 6, 7, 8, 9, 10$, we can get another sequence $1, 6, 2, 7, 3, 8, 4, 9, 5, 10$ utilizing sequence cross transformation.

The flowchart of the sequence cross transformation process is illustrated in Figure 1.

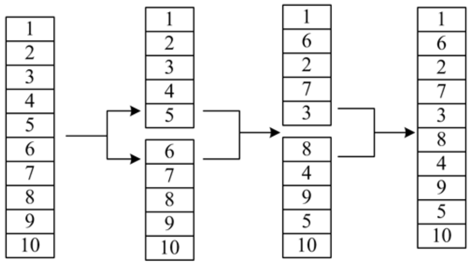


Figure 1: The flowchart of the sequence cross transformation process

2.2 Improved Sequence Cross Transformation

However, the sequential cross transformation method has periodicity [15]. In other words, the scrambled sequence will be restored to the original sequence after several transformations. To solve this problem, we propose an improved sequential cross transformation method in this paper. The improved sequence cross transformation method mainly includes the following three steps:

1) For an sequence, divide it into two equal parts from the middle.

2) Use circshift method to transform the two subsequences.

3) Rearrange this sequence using one by one cross method.

The main difference between the improved sequence cross transformation and sequence cross transformation is that the improved sequence cross transformation add a new step, *i.e.* Step (2), which can disrupt the transformation structure of sequence cross transformation and destroy its periodicity.

3 Image Encryption Algorithm

Denote the size of the plain image P as $M \times N$, where M and N represent the height and width of the image respectively. In this paper, the classical Lorenz chaotic system is used to generate chaotic sequences, which is described by [9]:

$$\begin{cases} \dot{x_1} = a(x_2 - x_1) \\ \dot{x_2} = bx_3 - x_1x_3 - x_2 \\ \dot{x_3} = -cx_3 + x_1x_2 \end{cases} \tag{1}$$

where x_1, x_2, x_3 are state variables and a, b, c are system parameters. We use the system parameters a, b, c and chaotic initial values x_0, y_0, z_0 to generate chaotic state variables x_1, y_1, z_1 of chaotic system (1). Furtherly, the chaotic sequences are transformed into new sequences suitable for image encryption.

3.1 The Encryption Method

The specific steps of the encryption algorithm can be described as follows:

Step (1): Calculate the initial values x_0, y_0, z_0 of the Lorenz chaotic system by:

$$\begin{cases} x_0 = \frac{1}{2}\frac{\sum_{ij} P_{ij}}{255MN} + 0.01 \\ y_0 = \frac{1}{3}\frac{\sum_{ij} P_{ij}}{255MN} + 0.02 \\ z_0 = \frac{1}{4}\frac{\sum_{ij} P_{ij}}{255MN} + 0.03 \end{cases} \tag{2}$$

Step (2): Choose the system control parameters a, b, c of the Lorenz chaotic system.

Step (3): Iterate the Lorenz chaotic system (1) for $N + 2000$ times with the initial values x_0, y_0, z_0, remove the former 2000 values and then we can obtain three chaotic sequences x_s, y_s, z_s of length L, where $L = M \times N$. Calculate three sequences S_1, S_2, S_3 with x_s, y_s, z_s by

$$\begin{cases} S_1 = |x_s| \times 10^{15} \mod \frac{MN}{2} \\ S_2 = |y_s| \times 10^{15} \mod \frac{MN}{2} \\ S_3 = |z_s| \times 10^{15} \mod 256 \end{cases} \tag{3}$$

Step (4): Transform the image matrix P into an one-dimensional pixel vector P_V.

Step (5): Set $t = 1$. Divide P_V into two equal parts from the middle. As a result, we can get two subsequences P_{V1} and P_{V2}.

Step (6): Utilize circshift transformation to transform the vector P_{V1} and P_{V2}, the specific formula is as follows:

$$\begin{cases} P_{V1} = circshift(P_{V1}, S_1(t)) \\ P_{V2} = circshift(P_{V2}, S_2(t)) \end{cases} \tag{4}$$

where $circshift(P, S(t))$ stands for shifting the elements in array P to the right by $S(t)$ positions.

Step (7): Rearrange P_{V1} and P_{V2} using one by one cross method and connect them to form a new vector P_1.

Set $t = t + 1$. Step (5)-Step (7) is repeated until $t > K$ so that the image is fully scrambled, where K is the number of scanning rounds.

Step (8): Perform the xor operation on P_1 using the random sequences S_3:

$$\begin{cases} C_V(1,1) = P_1(1,1) \oplus S_3(1,1) \\ C_V(1,i) = P_1(1,i) \oplus [C_V(1,i-1) \oplus S_3(1,i)] \end{cases} \tag{5}$$

where $i = 2, 3, \cdots, L$ and symbol " $\oplus$ " is the bitwise exclusive or operator and C_V is the ciphertext vector.

Step (9): Convert C_V into encrypted gray image C.

3.2 The Decryption Method

The decryption process is similar to the encryption process which mainly contains the following steps:

Step (1): Transform the cipher image C into one-dimensional pixel vector C_V.

Step (2): Calculate the scrambled image vector P_1 as follows:

$$\begin{cases} P_1(1,i) = C_V(1,i) \oplus [C_V(1,i-1) \oplus S_3(1,i)] \\ P_1(1,1) = C_V(1,1) \oplus S_3(1,1) \end{cases} \tag{6}$$

where $i = 2, 3, \cdots, L$. Set $t = K$.

Step (3): Divide sequence P_1 into two equal-length sequences P_{V1} and P_{V2}.

Step (4): Utilize circshift transformation to transform the vector P_{V1} and P_{V2}, the specific formula is as follows:

$$\begin{cases} P_{V1} = circshift(P_{V1}, -S_1(t)) \\ P_{V2} = circshift(P_{V2}, -S_2(t)) \end{cases} \tag{7}$$

where $circshift(P, -S(t))$ stands for shifting the elements in array P to the left by $S(t)$ positions. Then connect P_{V1} and P_{V2} to form a new vector P_0.

Step (5): Use inverse sequence cross transformation to transform P_0 then we can get a new vector P_1.

Set $t = t - 1$. Repeat Step(3)-Step(5) K rounds.

Step (6): Convert vector P_1 into plain image P.

4 Test and Analysis of the Proposed Scheme

The Matlab software is used as an experimental platform for experiments. Three images *i.e.* Lena, Hat and Plant (216×216) are taken for testing. The system parameters of the chaotic system are set as $a = 10, b = 8/3, c = 28$. In the following subsections, the experimental results and several different security analysis are given.

4.1 The Image Scrambling Effect

Sequence cross transformation is an image scrambling method, but it has periodicity. In this paper, an improved sequence cross transformation method is proposed. In order to compare the scrambling effects of them, we take the Lena image as a example. When different numbers K of scrambling rounds are taken, the scrambling images obtained by sequence cross transformation and the proposed method are shown in Figure 2.

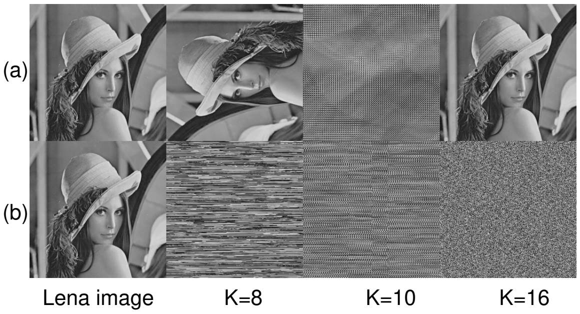


Figure 2: The Image scrambling performances. (a) Sequence cross transformation. (b) Improved sequence cross transformation.

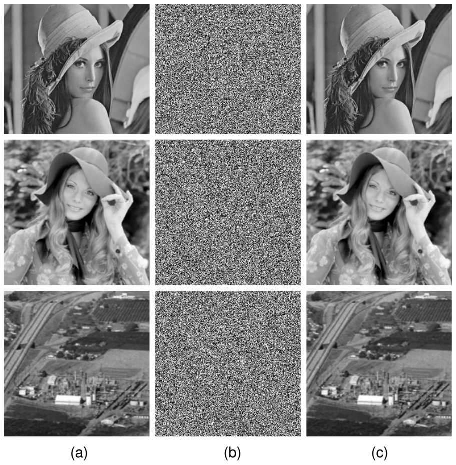


Figure 3: The experimental results of the encrypted image. (a) The plain image. (b) The ciphered image. (C) The decrypted image.

It can find that when $K = 16$, the scrambling image of sequence cross transformation changes to the original image which can also prove that the sequence cross transformation algorithm has periodicity. On the contrary, the proposed method has better and better scrambling effect with the increase of K, which shows that the proposed method can break the periodicity when used for image scanning.

4.2 The Encrypted Image

Set $K = 80$, the plain images and their corresponding encrypted images and decrypted images are illustrated in Figure 3. It is not easily to find any association between the encrypted images and the plain images, while the decrypted image looks the same as the plain image, which shows that the proposed method has good encryption and decryption effect.

4.3 Key Space Analysis

It is expected that the key space be large enough to resist violent attacks. In the proposed algorithm, the secret key is comprised of the chaotic system parameters a, b, c,

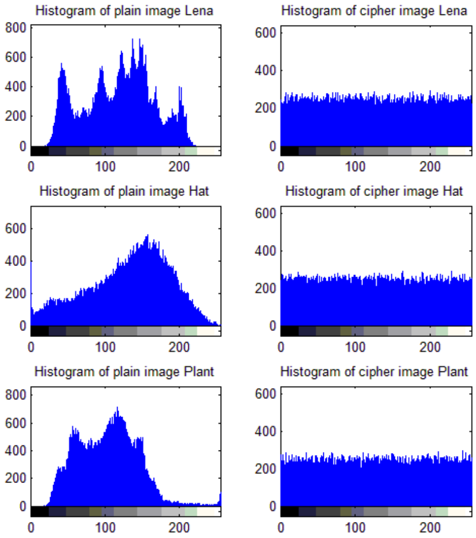


Figure 4: Histograms of plain images and cipher images

the chaotic system initial values x_0, y_0, z_0 and the number of scanning rounds K. Assuming that the precision of the system parameters is 10^{15}, the key space of the proposed algorithm is more than 10^{90}, which is large enough and is resistant to brute force attacks.

4.4 Histogram Analysis

The gray histogram is an important statistical feature which directly reflects the distribution characteristics of the image pixels. Figure 4 shows the histogram of the plain images and ciphered images. When the gray histogram is uniform and flat, it is difficult for the attacker to obtain the information of the original plain image from the cipher image. As can be seen from Figure 4, the pixel distribution of the original images are extremely uneven, while the distribution of the pixel values of the cipher images are very uniform and flat so that they have a good resistance to statistical analysis.

4.5 Correlation Analysis

There is a high correlation between adjacent pixels of a normal image. However, we hope that the adjacent pixel values are irregular and the correlation is low after the image is encrypted. We can calculate the correlation of adjacent pixels by [16]:

$$r_{xy} = \frac{\sum_{i=1}^{N}((x_i - E(x))(y_i - E(y)))}{\sqrt{(\sum_{i=1}^{N}(x_i - E(x))^2)(\sum_{i=1}^{N}(y_i - E(y))^2)}} \qquad (8)$$

where $E(x) = \sum_{i=1}^{N} x_i$, $E(y) = \sum_{i=1}^{N} y_i$, x_i and y_i are gray-level values of the selected adjacent pixels, and N is the number of sample pixels.

5000 pixels and its adjacent pixels in horizontal, vertical and diagonal directions in plaint images and cipher images are randomly select respectively for testing the correlation of adjacent pixels, and the results are shown in Table 1. In addition, we plotted the correlation distribution of r_{xy} of Lena image in Figure 5.

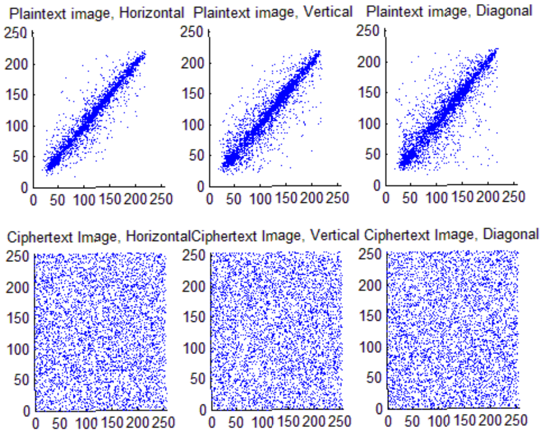


Figure 5: Correlation distributions of plain image Lena and cipher image Lena in each direction

Table 1: Correlation coefficients of the plain images and cipher images

Images		Horizontal	Vertical	Diagonal
Lena	Plain image	0.9407	0.9696	0.9086
	Cipher image	-0.0139	-0.0068	0.0157
Hat	Plain image	0.9770	0.9785	0.9569
	Cipher image	-0.0140	-0.0438	-0.0121
Plant	Plain image	0.9634	0.9413	0.8975
	Cipher image	0.0118	-0.0029	0.0159

From Table 1 and Figure 5 we could find that the adjacent pixels of the plain image have strong correlation, while the proposed image encryption method can eliminate the correlation of adjacent pixels.

4.6 Information Entropy Analysis

Information entropy can be used to measure the randomness of an image, and the theoretical maximum value of an image is 8 [22]. The formula of information entropy is as follows:

$$H(m) = -\sum_{i=0}^{255} P(m_i)\log_2 P(m_i) \qquad (9)$$

where m_i denotes the ith gray level for the digital image and $P(m_i)$ is the probability of m_i.

After calculation, the information entropy of plaint images are 7.4204, 7.7317 and 7.2990, while the information entropy of cipher images are all very close to the ideal value 8, which reflects that the encryption algorithm proposed in this paper has high security (See Table 2).

Table 2: The results of entropy analysis

Images		Information entropy
Lena	Plain image	7.4204
	Cipher image	7.9972
Hat	Plain image	7.7317
	Cipher image	7.9971
Plant	Plain image	7.2990
	Cipher image	7.9969

4.7 Analysis of Differential Attack Resistance

A secure cryptosystem should have a good anti-differential attack capability which depends on its sensitivity to plaintext image. The number of pixels change rate (NPCR) and the unified averaged changed intensity (UACI) [18] are usually applied to measure the ability to resist differential attack. Setting two plain images, there is only one-pixel difference between them. The formulas for calculating NPCR and UACI can be expressed as:

$$NPCR = \frac{\sum_{ij} D_{ij}}{W \times H} \times 100\% \tag{10}$$

$$UACI = \frac{1}{W \times H} \frac{\sum_{ij} (C_1(i,j) - C_2(i,j))}{255} \times 100\% \tag{11}$$

where C_1 and C_2 are the encrypted images for the plain images and D_{ij} is defined by

$$D_{ij} = \begin{cases} 0 & \text{if } C_1(i,j) = C_2(i,j) \\ 1 & \text{if } C_1(i,j) \neq C_2(i,j) \end{cases} \tag{12}$$

Table 3: NPCR and UACI

Images	NPCR(%)	UACI(%)
Lena	99.61	33.61
Hat	99.58	33.41
Plant	99.63	33.61

The calculation results of NPCR and UACI of the presented algorithm are also all very close to the ideal values as security required [18]. As a result, the presented algorithm can effectively resist differential attack (See Table 3).

4.8 Performance Comparison with Other Methods

In order to further illustrate the effectiveness of the proposed method, we compare the results of the Lena image of the proposed method with other image encryption methods proposed in [12] and [22]. The specific results are shown in Table 4.

From Table 4 we could find that the performance of the proposed method is close to that of [12] and [22]. Besides, the proposed method is simple and easy to be implemented so that it is suit for image encryption.

Table 4: Performance comparison with other methods

Index	Ref. [12]	Ref. [22]	Proposed
Correlation (Horizontal)	-0.0005	0.0150	-0.0139
Correlation (Vertical)	-0.0011	0.0044	-0.0068
Correlation (Diagonal)	-0.0015	0.0036	0.0157
NPCR(%)	99.56	99.66	99.61
UACI(%)	33.28	33.80	33.61
Information entropy	7.9971	7.9993	7.9972

5 Conclusions

In [15], the authors proposed the sequence cross transformation algorithm which is an image scrambling method in essence. The main drawback of the sequence cross transformation algorithm is that it has periodicity, therefore, the effect of image scrambling is not ideal. To solve this problem, we proposed an improved sequence cross transformation algorithm, which can overcome the shortcomings of sequence cross transformation and get better scrambling. Further, the improved sequence cross transformation algorithm is used for image scrambling and encryption. Experimentation is done on three classical images and the results of encryption and decryption test and security analysis confirm the effectiveness of the proposed image encryption method.

Acknowledgments

This work is partially supported by the National Natural Science Foundation of China (No.11871417). The authors gratefully acknowledge the anonymous reviewers for their valuable comments.

References

[1] S. Ahadpour, Y. Sadra, "A chaos-based image encryption scheme using chaotic coupled map lattices," *International Journal of Computer Applications*, vol. 49, no. 2, pp. 15-18, 2012.

[2] R. E. Boriga, A. C. Dascalescu, and A. V. Diaconu, "A new fast image encryption scheme based on 2D chaotic maps," *IAENG International Journal of Computer Science*, vol. 41, no. 4, pp. 249-258, 2014.

[3] L. F. Chen, D. M. Zhao, F. Ge, "Image encryption based on singular value decomposition and arnold transform in fractional domain," *Optics Communications*, vol. 291, pp. 98-103, 2013.

[4] M. Ghebleh, A. Kanso, D. Stevanovi, "A novel image encryption algorithm based on piecewise linear chaotic maps and least squares approximation," *Multimedia Tools and Applications*, vol. 77, no. 6, pp. 7305-7326, 2018.

[5] Y. Guo, L.P. Shao, L. Yang, "Bit-level image encryption algorithm based on Josephus and Henon chaotic map," *Application Research of Computers*, vol. 32, no. 4, pp. 1131-1137, 2015.

[6] P. Li, J. Xu, J. Mou, F. F. Yang, "Fractional-order 4D hyperchaotic memristive system and application in color image encryption," *EURASIP Journal on Image and Video Processing*, vol. 26, no. 10, pp. 11-23, 2017.

[7] H. Liu, C. Jin, "A color image encryption scheme based on arnold scrambling and quantum chaotic," *International Journal of Network Security*, vol. 19, no. 3, pp. 347-357, 2017.

[8] Y. B. Mao, G. R. Chen, S. G. Lian, "A novel fast image encryption scheme based on 3D chaotic Baker maps," *International Journal of Bifurcation & Chaos*, vol. 14, no. 10, pp. 3613-3624, 2004.

[9] B. Munmuangsaen, B. Srisuchinwong, "A hidden chaotic attractor in the classical Lorenz system," *Chaos, Solitons & Fractals*, vol. 107, no. 2, pp. 61-66, 2018.

[10] H. Natiq, N. M. G. Al-Saidi, M. R. M. Said, A. Kilicman, "A new hyperchaotic map and its application for image encryption," *The European Physical Journal Plus*, vol. 133, no. 6, pp. 5-18, 2018.

[11] P. Singh, A. K. Yadav, K. Singh, "Phase image encryption in the fractional Hartley domain using arnold transform and singular value decomposition," *Optics and Lasers in Engineering*, vol. 91, pp. 187-195, 2017.

[12] J. Tang, Z. Yu, L. Liu, "A delay coupling method to reduce the dynamical degradation of digital chaotic maps and its application for image encryption," *Multimedia Tools & Applications*, vol. 78, no. 17, pp. 24765-24788, 2019.

[13] X. Y. Wang, N. N. Guan, "A novel chaotic image encryption algorithm based on extended zigzag confusion and RNA operation," *Optics & Laser Technology*, vol. 131, no. 11, pp. 106366, 2020.

[14] X. Y. Wang, H. H. Sun, "A chaotic image encryption algorithm based on improved Joseph traversal and cyclic shift function," *Optics & Laser Technology*, vol. 122, no. 2, pp. 105854, 2020.

[15] J. Wang, X. Zhang, "A novel image scrambling and encryption algorithm based on sequence cross transformation," *Computer Applications and Software*, vol. 26, no. 12, pp. 111-114, 2009.

[16] X. Wu, H. Kan, and J. Kurths, "A new color image encryption scheme based on DNA sequences and multiple improved 1D chaotic maps," *Applied Soft Computing*, vol. 37, pp. 24-39, Dec. 2015.

[17] J. H. Wu, X. F. Liao, B. Yang, "Image encryption using 2D H non-Sine map and DNA approach," *Signal Processing*, vol. 153, no. 12, pp. 11-23, 2018.

[18] Y. Wu, J. P. Noonan, and S. Agaian, "NPCR and UACI randomness tests for image encryption," *Journal of Selected Areas in Telecommunications*, vol. 1, no. 2, pp. 31-38, 2011.

[19] X. Xu, J. Feng, "Research and implementation of image encryption algorithm based on zigzag transformation and inner product polarization vector," in *IEEE International Conference on Granular Computing*, pp. 556-561, 2010.

[20] R. Ye, "A novel chaos-based image encryption scheme with an efficient permutation-diffusion mechanism," *Optics Communications*, vol. 284, no. 22, pp. 5290-5298, 2011.

[21] S. L. Yin, J. Liu, L. Teng, "Improved elliptic curve cryptography with homomorphic encryption for medical image encryption," *International Journal of Network Security*, vol. 22, no. 1, pp. 155-160, 2020.

[22] C. Zhao, L. Yang, H. Zhu, *et al.*, "An image encryption scheme using generalized Arnold map and affine cipher," *Optik*, vol. 125, pp. 6672-6677, 2014.

Biography

Chunming Xu is an associate professor at the mathematics and statistical from Yancheng Teachers University, PR China. His main research interests include image processing and artificial intelligence.

Yong Zhang is an associate professor at the mathematics and statistical from Yancheng Teachers University, PR China. His main research interests include combinatorics and optimization.

Research on Multimedia Application on Information Hiding Forensics and Cybersecurity

Chin-Feng Lee[1], Chi-Yao Weng[2], Chih-Hung Wang[3],Goutam Chakraborty[4],
Kouichi Sakurai[5], Kuo-Yu Tsai[6]

(Corresponding author: Chin-Feng Lee)

Department of Information Management, Chaoyang University of Technology[1]
No. 168, Jifeng E Rd, Wufeng District, Taichung 413, Taiwan
Email: lcf@gm.cyut.edu.tw
Department of Computer Science, Ntional Pingtung University, Pingtung, Taiwan[2]
Department of Computer Science and Information Engineering, National Chiayi University, Chia-Yi, Taiwan[3]
Department of Department of Software and Information Science, Iwate Prefectural University, Japan[4]
Department of Mathematical Informatics, Kyushu University, Japan[5]
Department of Information Engineering and Computer Science, Feng Chia University,Taichung, Taiwan[6]

(Received May 6, 2021; Revised and Accepted Sept. 21, 2021; First Online Oct. 19, 2021)

Preface of Special Issue on Multimedia Application on Information Hiding Forensics and Cybersecurity

Abstract

With the rapid development of multimedia signal processing and digital communication technology, information transmitted on the network faces various security risks, such as (1) being intercepted by illegal users and leaking corporate secrets; (2) illegal tampering, resulting in data confusion and information errors; (3) Illegal users forge legal identities and send misinformation, bringing chaos to the regular order of production and operation and causing damage and loss. On the other hand, information hiding technology can protect information security and integrity, while digital forensics can be used to collect, test, and analyze digital evidence. In addition, with the rapid development of artificial intelligence (AI), many problems in multimedia security can be effectively solved. This special issue of "Multimedia Application on Information hiding Forensics and Cybersecurity" focuses on new information hiding methods and forensics in multimedia data based on intelligent technology. Following the IJNS journal's strict review procedures, each manuscript submitted to this special issue has undergone multiple rounds of technical reviews. In the end, this special issue selected a total of 3 papers which cover the topics: (1) steganography, (2) forensics and identity verification, and (3) social network artificial intelligence security.

Keywords: Anomaly Detection; Cybersecurity; Deep Learning; Digital Forensics; Information Hiding; Intrusion Detection

1 Introduction

Recently, advanced technology has been developed so user can use the advanced technology to transmit all the information via the media and network. Online social networks such as Facebook, Twitter, LinkedIn and Instagram are widely used in social, entertainment and other aspects, generating terabytes of data every day. If these data are improperly used for profit generation or abused by criminals, social network vulnerabilities will be created. With the popularity of social media, uploading and backing up digital images has become the norm. A huge amount of digital images is circulating on the Internet every day, and issues related to information security also follow. Information hiding technology can conceal, disguise or encrypt information without being discovered by the enemy or even intercepted or tampered with. In order to protect the security of the information, Information hiding is an indispensable technology. Moreover, with the rapid development of the fifth-generation mobile network and the Internet of Things, the low-latency and real-time transmission of the network may also bring huge challenges in the fields of intrusion detection, risk analysis, and threat prevention. Therefore, a large number of multimedia forensics and network security technologies have aroused extensive research attention in multimedia data integrity assessment.

Some novel attacks or new variants of malicious behaviors are not efficiently detected by most existing defensive tools because these attacks or behaviors evade detection by using network traffic obfuscation, or even employ adversarial machine learning techniques for further exploitations. Due to sophistication and scalability of the current network system, cybersecurity maintenance and hacker tracking conducted by traditional technologies may cause obvious problems of performance and accuracy. It is necessary to develop more robust detection, tracing and analysis models by leveraging machine learning or advanced analytics approaches to enhance security in network.

The goal of this special issue is to provide researchers engaged in all event information forensics with the primary platform to present their recent research results. It also provides an important opportunity for multidisciplinary studies connecting and creates new techniques to solve those multimedia forensics and cybersecurity problems with high performance.

2 Information Hiding

With the advent of artificial intelligence and the Internet of Things era, data applications have become extremely frequent and diverse. Data is the core asset of individuals, enterprises, or organizations. Therefore, it is important to protect the security and integrity of sensitive data and prevent information from being viewed, copied, or even tampered with. Information hiding technology, through concealment, disguise or encryption, is not detected by the enemy, and is further intercepted or tampered to protect the security of information.

Below we will discuss the origin of protecting information security, the concept and implementation of information hiding.

2.1 The Rise of Information Hiding

Information hiding can provide secure communication in the presence of malicious third parties, which can be traced back to ancient times. Whether in peacetime or during war, the ancient Egyptians, Greeks, Spartans, and Romans used various forms of information hiding in sports and military battles, and even used information hiding to manipulate people's beliefs and morals. The Egyptians used unordered hieroglyphics, the Greeks used steganography, the Spartans used Scytale, and the Romans used the Caesar code [22].

The two terms steganography and cryptography originated from ancient Greek words. The prefixes of the two terms are ”steganos” for protected (covered), and ”kryptos” for hidden (secret) [22]. The two meanings are obviously very close; both steganography and cryptography are methods of hiding information. But the two can be distinguished: the steganography method embeds information into a carrier by making it hard to notice, but the encryption method modifies the message by making it hard to identify.

In modern society, data protection involves many interdependent policy and technical issues including information confidentiality, anonymity, integrity, intellectual property, etc. Ensuring confidentiality and anonymity, is crucial for IoT applications. In fact, any failure would seriously threaten users' privacy. Thus, a wide deployment of IoT applications might be hindered to provide data confidentiality, data encryption such as: DES, RSA, SSL, and etc. uses complex encryption algorithms and ”keys” to convert ordinary information (or plaintext, Plaintext) into incomprehensible ciphertext. The Caesar Cipher is an example of cryptography. It is designed to ensure that the plaintext of the message is replaced by the ciphertext.

Cryptography and steganography are two different techniques that maintain data confidentiality and integrity. Cryptography refers to the art of converting plain text (messages) into an unreadable format, while on the other hand, steganography refers to hiding the existence of secret messages in some way (without traceability), In today's large-scale Internet and big data, this steganography that can accomplishes “invisibility” is really a very effective way of securing data transfer.

2.2 Data Hiding Technologies

In the natural ecology, many creatures, such as owls, stick insects, dead leaf butterflies or chameleons, use protective colors to avoid enemies, or to facilitate hunting. In fact, this is an art of the nature's steganography.

The research community has already done lots of noteworthy research in image steganography. The terms “data hiding” and “steganography” have been used interchangeably. However, ”data hiding” and “watermarking” are clearly separated according to the techniques and the applications. Secret data are hidden within seemingly innocuous host media or cover media such as images, music, video, audio and text to produce the stego-media or watermarks depending on which steganography schemes are employed. The main application of the data hiding technique is the sharing of secret information via a covert channel [13, 27, 31]. Thus, the aim of data hiding is thus to achieve higher capacity and lower distortion. The aim of a digital watermarking technique is to protect the ownership of copyright, and as such, its robustness is a critical research topic. In the following subsections, we focus on exploring data hiding and digital watermarking on imagery.

Data hiding techniques can be categorized into three branches which are steganography methods based on spatial, transform, and compression domains. The data hiding methods in spatial domain are simpler, easier to implement and lower computation complexity. The steganography scheme based on the transform domain is to first transform the cover image using transform oriented methods such as Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT), Discrete Fourier Transform (DFT), Singular Value Decomposition (SVD), and then hide the secret data in the frequency coefficients to improve security and robustness. In order to save the image storage space and reduce the transmission time, the image is usually compressed before being transmitted to the receiver to speed up the transmission rate. Many data hiding methods are related to image compression such as vector quantization (VQ) [19, 24, 58, 78, 86, 131] or side-match VQ compression (SMVQ) [43, 44, 62, 123]. The SMVQ technique has been widely used in various image compression and hiding techniques nowaday. It effectively decreases the bit rates of image with a great compressed quality. However, the prediction is not precise while the variation between the encoded block and its neighboring upper and left blocks is large. Due to the imprecision will cause derailment problem, to solve this problem, an adaptive data hiding scheme to embed secret data into an SMVQ compressed image has been proposed [62]. A data hiding scheme [41] based on side match vector quantization (SMVQ) and search-order code (SOC) is proposed. In this scheme, the VQ image is compressed by SMVQ technology to reduce redundancy, and then the image is recompressed by SOC.

In literature, various types of image steganographic methods have been proposed and can be two tracks which are irreversible data hiding (IRDH) and the reversible data hiding (RDH). IRDH means that after the secret data have been extracted from the stego-image, the cover image is distorted permanently and cannot be restored correctly. RDH method embeds secret data into a host image, and the stego image can be recovered to original one after the secret data are extracted by receiver. It can also be called lossless data hiding. RDH is highly used in some special application such as military, medical or satellite images processing, to ensure the integrity of secret data and the host image. The following study, we focused on both irreversible and ieversible data hiding schemes for digital images in spatial domain.

2.2.1 Irreversible Data Hiding in Spatial Domain

The most common steganography technique, using mostly image and sound carrier files, is called Least Significant Bit (LSB) Substitution or Replacement [15]. As the name implies, LSB Substitution uses the least significant bits of cover pixels to embed secret data. Later, based on edges, texture, and intensity level of the cover images, the image steganographic methods have been developed to enhance the embedding capacity and visual quality by adaptive LSB substitution onto variation of LSB's pixel or bit-planes. Zhang and Wang [137] devised an efficient steganographic embedding scheme by exploiting modification direction (EMD). Use of the EMD scheme not only maintains a high visual quality, but also increases the embedding capacity. The EMD first transformed a binary sequence of secret data into a sequence of digits in a $(2n+1)$-ary notational system and developed a modulus function to embed each secret digit into a pixel-group that contains n cover pixels. The system parameter n is the secret key that is used for the embedding and extracting process. At most one cover pixel of the pixel-group is increased or decreased by 1 in order to embed the secret data; thus, the best embedding capacity is 1 bit per pixel in the $(2n+1)$-ary notation system. Improving embedding capacity of EMD is an attractive research topic in data hiding in recent years. Many works [57, 59, 60, 72, 120, 122, 125, 137] improved the exploiting modification direction method to enhance the embedding capacity as well as visual quality. The reduplicated exploiting modification direction (REMD) method [21] exploited the edge detection of cover images and pixel interpolation technique to enhance the embedding capacity. Some others modulus operations have been designed [48, 61, 71] to resolve the spatial redundancy problem in EMD.

In 2003, Wu and Tsai proposed Pixel Value Differencing [127], also called PVD to allow a greater modification when a larger difference between adjacent pixels thus creating higher embedding capacity. PVD provides a balance between embedding capacity and image quality. Moreover, PVD has the ability of resisting statistical analysis attack. Thus a great many pixel value differencing (PVD) based steganographic methods have been proposed [68, 83, 114, 128, 130, 132] to achieve high embedding capacity, good image quality while ensuring the security of the image.

The magic matrix-based data hiding method is a novel method proposed in recent years. A magic matrix can be the Sudoku matrix [17], a turtle shell matrix [18, 85] an octagonal matrix [74], and other patterns of matrices [63, 69, 70]. As far as the data embedding method using the magic matrix is concerned, the key element to improve the embedding ability and image quality is to construct a magic matrix and the conditions for traversing the area. With the magic matrix-based data hiding methods, decoding becomes easy. After the stego-image and magic matrix are input, the ciphertext can be extracted from the magic matrix based on the pixel pairs of the stego-image. The matrix in the above methods was mostly constructed in a 2D manner; i.e. the x-axis and the y-axis of the matrix were constructed by a pair of pixels but a 3D matrix-based data hiding method [64] can extend the traversal area to significantly increase the embedding capacity.

2.2.2 Reversible Data Hiding (RDH)

Many RDH methods have been proposed over the years. Among them, the two major techniques are difference expansion (DE) was proposed by Tian in 2003 [117] and histogram shifting (HS) proposed by Ni *et al.* in 2006 [92]. Since the DE and HS can reduce the embedding distortion and provide a sufficient embedding payload, various RDH techniques have been developed along these two lines to maintain a good payload-distortion performance. The difference expansion method (DE) makes use of the similarity between adjacent pixels in spatial domain, calculates the difference value between one pair of pixels and expands the pixel value to embed 1 bit of secret data. In order to discriminate between the two situations, the extract information is recorded using a location map. However, the location map will reduce the pure payload of camouflage images. Many research works [3, 5, 42, 54, 67, 101, 116, 124] have been proposed different methods of improvement.

Another main technique of RDH is histogram shifting method (HS) which uses statistics to obtain the number of times each pixel value occurs and then defines the highest frequency occurrence as the peak point and the lowest as the zero point. Shifting the pixel value between the peak and zero-value points results in a vacated position in which to embed the secret message. HS maintains high image quality since the pixel values are only adjusted by 1 unit at most. But it makes lower embedding capacity in result. Continuously, other scholars have also proposed relevant techniques based on HS methods. Those methods are mainly trying to build a sharp histogram of the Laplacian-like distributions. The prediction-error expansion (PEE) method by Thodi and Rodríguez [116] who introduced histogram-shifting (HS) which significantly reduced location map (LM) size. Since then, many variations of PEE based methods [40, 73, 87] were developed.

In 2013, Li *et al.* [77] proposed a high-fidelity reversible data hiding method called pixel value sorting (PVO). They combine the concept of PEE and HS methods to produce camouflage pixels of good image quality. To calculate the prediction error in block-by-block manner, the pixels in the image block are first sorted in ascending order. Then predict the smallest pixel with the second smallest pixel, and predict the largest pixel with the second largest pixel. Other improvement of PVO are proposed [66, 75, 93–95, 102].

2.3 Digital Watermarking

Image watermarking can be classified as either robust image watermarking for copyright protection or fragile image watermarking for integrity verification. Many researchers refer to the topic of robust image watermarking [4,7,106,108]. The robust image watermarking scheme usually embeds a watermark into a cover image. To protect the copyright, the embedded watermark should be able to be extracted and verified by the owner, even from a modified image. The modification may be malicious, intentional tampering or other common image attacks. Fragile watermarking schemes can be further divided into semi-fragile and complete fragile schemes. The major difference depends on the integrity criteria [84]. Semi-fragile watermarking [16, 20, 79] is also called soft authentication [84] and provides relatively relaxed integrity criteria. Some kinds of invisible modification are allowed, such as JPEG compression. This scheme is useful when the protected media need to be compressed at different rates in order to satisfy the transmission bandwidth. The complete fragile watermarking scheme, also called hard authentication [84], offers greater protection and integrity than soft authentication. That is, this scheme does not allow any kind of modification or tampering with the protected image. In addition to detecting whether a protected image has been modified, the hard authentication scheme must be capable of locating the tampered area.

Since most image tampering detection methods are based on image blocks and ignores characteristics of the image blocks, poor image quality results from hiding the watermark. Adaptive embedding rules in spatial domain for image tamper detection and recovery by applying various hiding, detection and recovery methods according to the block's smoothness. Kinds of effective self-embedding watermarking method for image tamper and recovery capability positioning have been conducted on the transform domain to generate authentication and recovery data from the image, and get better image quality than methods used in the past. Self-embedding fragile watermarking algorithms [53, 65, 99, 100, 107, 109, 133] can perform detection of the manipulated areas as well as recovery of these detected areas. Moreover, image authentication technologies based on block-wise and pixel-wise detection methods [65, 133] can effectively reduce the error rates and enhance the recovered image quality.

2.4 Combination of Steganography and Cryptography

The accelerated development of the Internet and multimedia technology has greatly promoted human communication. A lot of information is stored in digital form. Protecting content is a priority. data hiding involves hiding confidential data in other seemingly harmless host media or cover media, such as text, video, audio, images, and compression codes. The embedded confidential data can then be used as a verification code to protect intellectual property or as confidential data for shared information. If confidential communication is the goal, steganography is preferred; in addition, if copyright protection is the goal, watermarking should be considered. In view of the trend of covert communication between people via the Internet, the focus of this work is to provide secure communication through digital images and avoid being discovered by unauthorized users.

Steganography is the art of information hiding, used to secretly transmit useful information through communication channels. Some people even think that steganography is a better way to protect messages than cryptography, because cryptography only recodes or disrupts the content of the message, while steganography allows the original message to be hidden in the carrier without being noticed. Since the security of secret message can be protected using cryptography or steganography, it might be a good way to combine both methods to achieve a hybrid system for producing better protection of the message. In case, when the steganography fails and the message can be detected, it is still of no use as it is encrypted using cryptography techniques. The combination of many steganographic techniques with different cryptographic algorithms can be referred in [14, 133].

3 Digital Forensics

The growth of digital devices and advanced communication technologies have rapidly developed for enhancing our life; however, it will suffer the problem of digital crimes. Digital forensics aims to collect crime related evidence from various digital environment and analyze it [47] [26]. Computer forensics and investigation has become an important field and it has becoming a specialized and accepted investigative technique with legal tools that validate the discipline. The forensics concept is basically an investigating process dedicated to finding the "truth" with legal routes [98]. The forensics is not to assign crime or innocence but rather to find the facts in the form of digital evidence.

The history of digital forensics has been mentioned for three decades. The digital forensics is the science and developed for how to detection, extraction and analysis the facts from the digital environment, and is also the critical requirements in cyberspace [110]. One important issue of digital forensics is to prepare accepted reports [25] for the court. The report could include the victim, an attempt, and vulnerable times. Additionally, recent researches in digital forensics are mostly concentration to gather and analyze the evidences that will provide to the court. Digital forensics attempts can be divided into two groups, including evidence collection and evidence analysis, shown in Figure 1. In this figure, the issue of evidence collection includes collecting data from physical devices, integrity, and scalability. Tree important components in evidence analysis are hard disk, memory and network forensics recorded and analyzed the tracks from behaviors of cyber criminals.

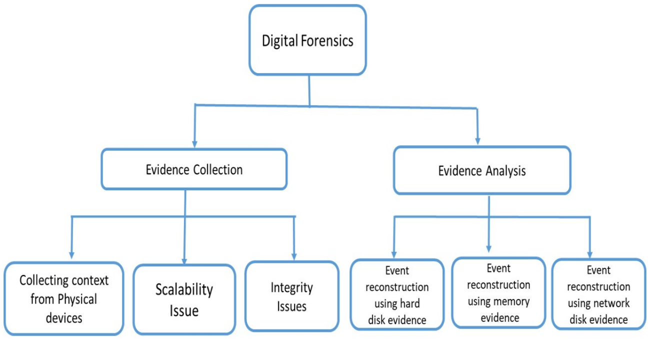

Digital Forensics
Evidence Collection
Evidence Analysis
Collecting context from Physical devices
Scalability Issue
Integrity Issues
Event reconstruction using hard disk evidence
Event reconstruction using memory evidence
Event reconstruction using network disk evidence

Figure 1: The classification of digital evidence collection and analysis

Several methods and forensic tools, e.g. Guidance Software's Encase [36] and AccessData's FTK [1], are applied to collect evidence from the system. Therefore, volume of information is extracted by the forensic tools. The low-level information extracted from the system is a challenge matters. Extracting millions of low-level information to present high-level information will take more time consuming and exhausting works. Some automatic methods are provided to extract high-level information from raw-data [38] [49] [52]. This work would review and introduce the research works on digital evidence collection and analysis and indicate some challenges.

3.1 Evidence Collection

The law executors (or called as digital investigator) must to collect pieces of evidence from the compromised systems for providing the accepted report to the court. The law executors have resulted in design and implementation of various digital forensics tools, including free tool and commercial tool. So, some tools have been designed for given information about various items, such as network connections, running processes, log files, registry, system services, and so on, to user.

Digital investigators have to collect the total context in the physical devices, such as hard disk, memory, and so on. Regarding this, some digital tools such as Encse [35] and FTK image [2] are introduced. Those tools are widely accepted to be a forensic methodology but it should turn off a computer and then acquire a bit-stream image from the hard disk. This is not a sufficient option because some case that some critical servers cannot be shut down during the extracting the evidences. Moreover, some forensically valuable information such as decoded passwords and un-obfuscated malware are found in the physical memory. Several digital forensics tools for collecting physical memory context are designed, such as FTK Imager, and Moonsols [90].

Images extracting from the hard disk or physical memory are valuable forensics information. Thus, some methods or tools try to extract the data from these images. For example, Stevenal and Casey's method [112] extract Windows command data from memory image. R. Beverly *et al.* [8] proposed a method to separate the packets from the network, and to associate data from the memory of operation system.

During the data collection using forensics tools, the integrity of the data collected data is an important issue. Several papers [111] [113] show the problem of by passing the integrity of the collected data. Ref. [111] states that many memory collection techniques are destroyed by simple anti-forensics techniques. This is that these skills require source code to be run on the compromised system. The memory collection technique is based on page table manipulation and hardware. Assume the memory collection should not rely on operation system facilities, it can achieve more robust environments to resistant anti-forensics tricks.

The other issue regarding data collection is the scalability. Extracting the information from digital media or physical devices can generate the volume and variety of digital evidences. Conventional annual methods examination of evidence may require more times to obtain the data. Serval papers [103] [89] [119] [10] are proposed to speed the examination times [119]. T. Vidas *et al.* [119]proposed OpenLV method to speed up the loading of forensics image. Open LV can reduce the examination time to review the digital forensics during investigation process. O. Brady *et al.* [10] develop a DESO (Digital Evidence Semantic Ontology) method to examine a repository and classifier of digital evidence.

3.2 Evidence Analysis

After extracting the evidence from different compromised system or others digital environment, digital investigations should analyze them and to generate accepted report. The extracted evidences are usually low-level information and are not in a same format. Therefore, the investigators encounter mass of low-level raw information. Several researches on digital forensics have addressed the issue of reconstructing consistence events. The extracted evidences from the compromised systems are various, and the researcher often focus on them. Many researchers have reconstructed events from extracting the information on hard disk or memory image.

The extracted evidence form the physical devices is more reliable and referenced in the court than that of extracted from memory. Physical devices event is reconstruction by finding pattern or correlation among low-level evidence. However, some papers use files system metadata to find the signatures and reconstruct events [51]. The correlation events in digital forensics analyzes the data that can be stored in databased or in ontologies. The ontologies method is good way to reconstruct events that uses a formal description of ontologies automates.

Additionally, the miscellaneous method is the other approach for event reconstruction. Ref. [32] presents a reconstructing events approach based on rigorous. This method applied the Finite State Machine (FSM) method to describe the compromised system. The FSM method can show all possible scenarios of the incident using backtracing transitions. The proposed FSM method cannot be used and stated all condistions in complex real systems because of thousands states and transitions grow exponentially. Ref. [12] present a model of cyber forensics ontology. The ontological model consists of a five hierarchical layer and the final layer generated the specified instances for certifying and specializing. The first layer consists of technology and profession. In the second layers, technology is divided into two parts of hardware and software; and profession is separated into the parts of law, academia, military and private. Each of these part is also divided into sublayers based on various concepts. This ontology model can be used for event reconstruction.

3.3 Challenges

Digital forensics has become an important issue to invest and identify the information of computer-based or computer-assisted crime. The challenges of digital forensic is divided into three categories: technical, legal, and resource. The technical challenges consist of differing formats, steganography, anti-forensics, and encryption. The legal challenges include jurisdictional issues and lack standardized legislation. The resource challenges are volume of data, time taken to obtain and analyze forensic data.

Network forensics is a sub-branch of the digital forensics. It refers to invest and analyze all traffic packages across a network. The attacker can use the data-stealing malware or analysis information to achieve the cybercrime. The packets in the network have various types, such VoIP, multimedia, video, sounds, images, and so on. The Network forensics for VoIP Packets is introduced. The paper "VoIP Packets Filtering for Mobile Instant Messaging Using N-gram Models" by Tung and Yen present an iterative algorithm for discovering attack patterns via a feedback mechanism, with the degrees of belief for attack instances propagated to the next iteration to further refine the search. This work observes the packet features of P2P connections established using various communication software from various fields other than the IP information of each packet. An automated method that can effectively filter out VoIP packets with an N-gram model is developed to improve the efficiency of criminal investigations.

While the challenges have been documented, the researcher should understand the practical reality and relevance of the challenges. Because a challenge exists, the investigators should face on these challenges. The future work of forensics must focus on developing effective approach to overcome the challenges and provide a robust environment to develop forensic solutions. Consideration to forensic capabilities, the forensics technologies should be advanced development and usage in order to invest and reconstruct useful information for the court.

4 Machine Learning/Deep Learning in Cybersecurity

Cybersecurity, including networked security and application security, can be enhanced by leveraging advanced analytical tools such as machine learning (ML) and deep learning (DL) technologies. As the fast development of artificial intelligence (AI), the security threats can be efficiently detected and fought by collecting and analyzing huge data of malicious and benign network traffic. The best practice of ML/DL on cybersecurity is establishing the classification models on the intrusion detection (or anomaly detection) mechanism. Machine learning has been a sophisticated technology applied on intrusion detection and there are some commercial platforms, such as Cisco Secure Network Analytics, Microsoft Azure Sentinel, etc. that integrate machine learning into their security detection systems. Many research papers [9, 138] have also discussed how to adopt the two famous categories, traditional machine learning model (or named shallow model) and deep learning model, for intrusion detection to increase the accuracy of prediction on a variety of attacking scenarios.

Table 1: Taxonomy of Machine Learning Methods

Supervised Learning	(Traditional ML) ANN, KNN, SVM, Decision Tree, Random Forest, Linear/Logistic Regression
	(DL) DNN, CNN, RNN, LSTM, GRU
Unsupervised Learning	(Traditional ML) K-means, Outlier Mining
	(DL) Autoencoder, GAN

4.1 Taxonomy of Machine Learning on Intrusion Detection

The machine learning is usually divided into supervised learning and unsupervised learning methods according to the source of data used in training being labeled or not [55,129]. Supervised learning means it learns valuable information from reliably labeled data, while the unsupervised learning means it focuses on the unlabeled data to find the outliers. Nevertheless, it is an expensive job to prepare suitable labeled data for training, especially on network packets containing malicious behaviors since collecting them needs some professional experiments. Although the unlabeled data are easier to be prepared, the unsupervised learning method must rely on the assumption that the amount of normal data is much larger than that of malicious data. Furthermore, the unsupervised learning method generally cannot reach the same high detection rate as the supervised learning method. Some popular machine learning and deep learning algorithms are described below and they often can be categorized into supervised learning and unsupervised learning [30, 82] as shown in Table 1.

Support Vector Machine (SVM). SVMs can perform classification by finding an optimization of separation to distinguish two classes on an N-dimensional hyperplane. The purpose of SVMs is to maximize the margin of the two classes, which can assist the classifier to obtain better performance. Although SVMs can have good effectiveness in the high dimensional data space, they may not reach desired results for training huge amount of data or being applied on multi-class classification model. Some research papers have been proposed to apply SVMs for the popular datasets for IDS [56, 97].

Decision Tree (DT) and Random Forest (RF). The decision tree provides supervised learning and can be regarded as a decision aid tool to infer some useful rules and create the model for prediction of the final result. Using decision tree in classification can obtain high performance and easily handle multi-class tasks. However, it has a disadvantage of easily overfitting since the training data may create an overly complex tree; consequently, the model keeps the noise but fails to catch the important information. Overfitting may cause that the prediction for the training data is accurate, but for the unseen data gets poor results. The studies such as [46, 104] shows the recent research results for intrusion detection.

The random forest (RF) developed by [11] is a concept of ensemble method by the integration of decision trees to create a forest through randomly distributing the training data, and make the model stronger and more accurate. The RF method can be stable in performing classification because it may not be seriously affected by the new data as the result is decided by all trees together. However, the computational complexity of RF is high and it also needs more storage space. Some previous papers [88, 126] further applied RF in the malicious URL detection.

Clustering Methods. This method supports unsupervised learning such that it does not need any specific knowledge or labeled data to perform classification. It can group the data to the same cluster according to the similarity defined by the distance value. In the clustering algorithms, K-means is a representative method, in which K denotes the number of clusters the model establishes. Based on the distance between two data points, K-means method can partition the data into K non-overlapping subgroups and put the data which are very similar to each other into the same cluster. Peng *et al.* [96] recently proposed an intrusion detection system based on mini batch K-means (MBKM) mechanism combined with principal component analysis to be suitable for the environment of big data.

The clustering analysis can be used in different research areas such as credit card fraud detection from a large amount of transaction records. Chen *et al.* proposed their paper ”Retrieving Potential Cybersecurity Information from Hacker Forums” to adopt the clustering method on a new application to efficiently discover the cybersecurity-related information or cyber threat intelligence (CTI) from hacker forums because of the complexity and diversity of their contents.

Autoencoder (AE). It is usually used in unsupervised learning to design two symmetrical procedures, an encoder and a decoder, to support outlier detection [76, 115]. The encoder performs the feature extraction on the input raw data and carries out the dimensionality reduction to extract the most important features from the input data. On the contrary, the decoder is designed like a decompressor to reconstruct the data from the inputted compressed representation of extracted features. According to the concept of AE, the large amount of similar traffic (assumed to be benign data) will obtain low reconstruction losses, while the outlier traffic (assumed to be anomaly) will get a very high reconstruction loss to make it easily be identified.

Convolutional Neural Network (CNN) and Long Short-term Memory (LSTM). The CNN and LSTM belong to the field of deep learning model that can support high accuracy classification of intrusion events. The LSTM is a kind of recurrent neural network (RNN) and designed to find the temporal context on large time scale for supporting learning of long-term dependencies.

There are some research papers combined CNN and LSTM to build their classification models with high accuracy for detecting intruders. In 2020, Zhang *et al.* [135] used three CNNs (Multiscale CNNs) to extract different spatial features and combined LSTM to handle the temporal features. Hassan *et al.* [39] used CNN of two convolution layers and one pooling layer combined with weight-drop LSTM (named CNN-WDLSTM) to extract the temporal features and avoid the overfitting.

4.2 Datasets and Imbalanced Data of Attacking Types

In the following, some popular datasets are described and the possible problem of imbalanced data is also explained.

4.2.1 Datasets

The datasets can be used to support the assessment of the IDS model for training and testing procedures. Establishing a dataset needs its specific network environment such that the collected data may contain some drawbacks for training in IDS model. We discuss several datasets usually used as the benchmark.

NSL-KDD Cup 99 [29]. it is an improved version of KDD CUP 99 [50] to provide the following refinements: (a) Remove duplicate and redundant data records from the original KDD CUP 99 to make the classifier avoid the biased situation and have better performance; (b) Provide appropriate and reasonable amount of training and testing data. That means the experiments conducted on it can use the entire set without performing sampling process. The evaluation results from different researchers can be easily compared to each other. The details of NSL-KDD Cup 99 include 40 class labels belonging to the following four groups: DoS (Denial of Service), U2R (User to Root), R2L (Remote to Local) and Probe.

UNSW-NB15 [91]. It was originally developed by the Australian Centre for Cyber Security (ACCS) to generate network packets by using IXIA PerfectStorm tool. The TCPdump tool is used to collect 100GB raw data and it is finally organized to be 2540,044 traffic records. Table 2 shows the number of records for each data type in UNSW-NB15. The imbalanced data problem that means the classes of Backdoor, Analysis, Shellcode and Worms having relatively small amounts of data, also can be found in Table 2.

Table 2: The number of records for each data type in UNSW-NB15

Data Type	Number of Records
Benign	2,218,761
Exploits	44,525
DoS	16,353
Backdoor	2,329
Analysis	2,677
Fuzzers	24,246
Genetic	215,481
Reconnaissance	13,987
Shellcode	1,511
Worms	174

CICIDS2017 [105]. This dataset was constructed by Canadian Institute of Cybersecurity (CIC). The realistic benign traffic was generated by profiling abstract human behaviors to build them for 25 users. The total number of data records in CICIDS2017 is 2,830,743 extracted by CICFlowMeter tool based on complete network traffic and configuration. It includes a benign type and 14 attack types of data, but also has imbalanced data problem for XSS, SQL Injection and Infiltration attack types.

4.2.2 Mitigating Problem of Imbalanced Data

In the following, the generative adversarial network (GAN), focal loss, and synthetic minority over-sampling technique (SMOTE) are described to explain how to mitigate the problem of imbalanced data for enhancing detecting rate of the attacking types which have fewer records in dataset.

Generative Adversarial Network (GAN). It was designed by Goodfellow *et al.* [33] in 2014 and can effectively generate new data statistically similar to the training data. Two neural networks, a generator and a discriminator, compete to each other. The generator is expected to produce new instances to make a fool of the discriminator. However, the discriminator tries to classify whether the input data is from original training data (called real data) or the generator (called fake or generated data).

However, the process of training GAN is very unstable and may have the problems of mode collapse and vanishing gradients. To solve these problems, Arjovsky *et al.* [6] proposed an improved method to use Earth-Mover (also named Wasserstein-1) distance. Gulrajani *et al.* [37] further proposed an alternative way (called WGAN-GP) to enforce Lipschitz constraint to make the training easier. In the literature, Lin *et al.* [81] leveraged Wasserstein GAN to generate malicious traffic records used for trying to deceive the IDS, which can make the IDS have the capability of resisting various real world attack scenarios. Apart from that, Fiore *et al.* [28] and Zheng *et al.* [136] applied GAN mechanism on the field of detecting frauds of financial data such as credit card records. Huang and Lei [45] proposed a novel model of Imbalanced Generative Adversarial Network (IGAN) to process the problem of imbalanced attacking types.

Focal Loss. It was proposed by Lin *et al.* [80] to originally solve extremely imbalanced problem between foreground and background classes for one-stage detector. A new loss function called focal loss was designed and evaluated by training a simple dense detector.

Synthetic Minority Over-sampling Technique (SMOTE). It was proposed by Chawla *et al.* in 2002 [21], which can solve the imbalanced data problem by increasing the proportion of minority (abnormal) class (i.e., over-sampling) and decreasing the proportion of the majority (normal) class (i.e., under-sampling) in the total training set [134].

4.3 Feature Selection

The feature selection is a technique to find the subset of relevant features in dataset so that the model can be improved to be cost-effective and performing classification efficiently. Due to the dimensionality reduction and irrelevant or redundant data removing, the feature selection can improve execution performance and classification accuracy of the model. The common used approaches of feature selection include statistical methods, genetic algorithms, entropy calculation, mutual information, information gain, etc. There are some previous research papers [34, 138, 139] to improve the model and get a better classification results by using fewer features in training.

5 Conclusions

The goal of this special issue "Multimedia Application on Information Forensics and Cybersecurity" is to provide proper and premier forum for researchers working on information hiding for all events to present their recent research results. It also provides an important opportunity for multidisciplinary studies connecting on industry and creating new techniques to solve those multimedia forensics and cybersecurity problems with improved performance.

There are 3 papers for publication which are related to the topics of this special issue.

The paper “VoIP Packets Filtering for Mobile Instant Messaging Using N-gram Models” by Tung and Yen who present an iterative algorithm for discovering attack patterns via a feedback mechanism, with the degrees of belief for attack instances propagated to the next iteration to further refine the search [118]. This work observes the packet features of P2P connections established using various communication software from various fields other than the IP information of each packet. An automated method that can effectively filter out VoIP packets with an N-gram model is developed to improve the efficiency of criminal investigations.

In “Data Hiding Based on The Chinese Remainder Theorem”, Wang *et al.* investigated the modulus of the Chinese Remainder Theorem (CRT) to the LSB Replacement method and GEMD method to break down the confidential data under the binary system, and embed the confidential data after breakdown into the designated pixel of the image in the carrier with the use of LSB Replacement method or the GEMD hiding method to form a concealed image [121]. Additionally, pixel groups selected by modulus value in this article were used to enhance security.

In “Retrieving Potential Cybersecurity Information from Hacker Forums”, Chen *et al.* introduced an improved data preprocessing method to reduce feature dimension and a hybrid method combining text tagging and clustering analysis techniques to discover cybersecurity information from unstructured hacker forums [23]. This research applies NLP, tagging, and clustering techniques to explore and capture cybersecurity information in hacker forums. The proposed CTI information retrieval method applies tagging and Word2Vec word embedding to extract key features and employs K-means and LDA two-stage clustering to discover CTI information from unstructured data.

The publication of this special issue was achieved based on joint cooperation by all involved. The Guest Editorial team would like to thank all the authors for the submission of their contributions and also appreciate the technical reviewers who contribute their knowledge and experiences to review the submitted manuscripts. We express our great gratitude and heartfelt thanks to the IJNS staffs, especially to Dr. Min-Shiang Hwang, IJNS Editor-in-Chief, who have provided significantly great support and guidance to the whole process. We do hope that the three papers published in this special issue could benefit the readers with wider range of information and better understanding in the fields.

Acknowledgments

We are grateful to Mr. Cheng-Yu Tu and Mr. Jing-Tong Wang with the department of Computer Science and Information Engineering of National Chiayi University in Taiwan who participated in this research work. This research was partially supported by the Ministry of Science and Technology, Taiwan, Republic of China under the Grant MOST 108-2221-E-153-004-MY2, MOST

110-2221-E-153-002-MY2, and MOST 110-2218-E-004-001-MBK.

References

[1] AccessData Group, *FTK - Forensic Toolkit*, May 21, 2021. (http://www.accessdata.com/products/digital-forensics/ftk)

[2] AccessData Group, *FTK Imager*, May 21, 2021. (http://accessdata.com/)

[3] A. M. Alattar, "Reversible watermark using the difference expansion of a generalized integer transform," *IEEE Transactions on Image Processing*, vol. 13, no. 8, p. 1147–1156, 2004.

[4] L. An, X. Gao, X. Li, D. Tao, D. Cheng, and J. Li, "Robust reversible watermarking via clustering and enhanced pixel-wise masking," *IEEE Transactions on Image Processing*, vol. 21, no. 8, p. 3598–3611, 2012.

[5] A. Arham, H. A. Nugroho, and T. B. Adji, "Multiple layer data hiding scheme based on difference expansion of quad," *Signal Processing*, vol. 137, no. 8, p. 52–62, 2017.

[6] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein generative adversarial networks," in *Proceedings of the 34th International Conference on Machine Learning*, pp. 214–223, 2017.

[7] A. Bajaj, "Robust and reversible digital image watermarking technique based on RDWT-DCT-SVD," in *2014 International Conference on Advances in Engineering and Technology Research*, pp. 1–5, 2014.

[8] R. Beverly, S. Garfinkel, and G. Cardwell, "Forensic carving of network packets and associated data structures," *Digital Investigation*, vol. 8, pp. 78–89, 2011.

[9] M. H. Bhuyan, D. K. Bhattacharyya, and J. K. Kalita, "Network anomaly detection: methods, systems and tools," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 1, pp. 303–336, 2014.

[10] O. Brady, R. Overill, and J. Keppens, "Deso: Addressing volume and variety in large-scale criminal cases," *Digital Investigation*, vol. 15, pp. 72–82, 2015.

[11] L. Breiman, "Random forests," *Machine Learning*, vol. 45, pp. 5–23, 2001.

[12] A. Brinson, A. Robinson, and M. Rogers, "A cyber forensics ontology: Creating a new approach to studying cyber forensics," *Digital Investigation*, vol. 3, pp. 37–43, 2006.

[13] L. Caviglione, "Trends and challenges in network covert channels countermeasures," *Applied Sciences*, vol. 11, p. 1641, 2021.

[14] K. Challita and H. Farhat, "Combining steganography and cryptography: New directions," *International Journal of New Computer Architectures and their Applications (IJNCAA)*, vol. 1, no. 1, pp. 199–208, 2011.

[15] C. K. Chan and L. M. Cheng, "Hiding data in images by simple LSB substitution," *Pattern Recognition*, vol. 37, pp. 469–474, 2004.

[16] C. S. Chan and C. C. Chang, "An efficient image authentication method based on hamming code," *Pattern Recognition*, vol. 40, no. 2, pp. 681–690, 2007.

[17] C. C. Chang, Y. C. Chou, and T. D. Kieu, "An information hiding scheme using sudoku," in *2008 3rd International Conference on Innovative Computing Information and Control (ICICIC 2008)*, 2008.

[18] C. C. Chang, Y. Liu, and T. S. Nguyen, "A novel turtle shell based scheme for data hiding," in *The 10th International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP 2014)*, 2014.

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

[20] E. C. Chang, M. S. Kankanhalli, X. Guan, Z. Y. Huang, and Y. H. Wu, "Robust image authentication using content based compression," *ACM Multimedia System Journal*, vol. 9, no. 2, pp. 121–130, 2003.

[21] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, no. 1, pp. 321–357, 2002.

[22] A. Cheddad, J. Condell, K. Curran, and P. M. Kevitt, "Enhancing steganography in digital images," in *2008 Canadian Conference on Computer and Robot Vision*, pp. 326–332, 2008.

[23] C. M. Chen, D. W. Wen, Y. H. Ou, W. C. Chao, Z. X. Cai, "Retrieving potential cybersecurity information from hacker forums," *International Journal of Network Security*, vol. 23, no. 6, pp. 1126-1138, 2021.

[24] W. J. Chen and W. T. Huang, "VQ indexes compression and information hiding using hybrid lossless index coding," *Digit Signal Process*, vol. 19, no. 3, p. 433–443, 2009.

[25] L. Daniel and L. Daniel, *Digital Forensics for Legal Professionals: Understanding Digital Evidence From The Warrant To The Courtroom.* Syngress, 2011.

[26] M. AI Fahdi, N. L. Clarke, and S. M. Furnell, "Challenges to digital forensics: A survey of researchers and practilitioneers attitdes and opinions," in *Proceedings of 2013 Information Security for South Africa*, pp. 1–8, 2013.

[27] T. S. A. Fatayer, *Secure Communication Using Cryptography and Covert Channel*, Springer International Publishing, 2020.

[28] U. Fiore, A. D. Santis, F. Perla, P. Zanetti, and F. Palmieri, "Using generative adversarial networks for improving classification effectiveness in

credit card fraud detection," *Information Sciences*, vol. 479, pp. 448–455, 2019.

[29] A. Frank, A. Asuncion, *UCI Machine Learning Repository*, Irvine, CA: University of California, School of Information and Computer Science, 2010. (http://archive.ics.uci.edu/ml)

[30] S. Gamage, J. Samarabandu, "Deep learning methods in network intrusion detection: A survey and an objective comparison," *Journal of Network and Computer Applications*, vol. 169, p. 102767, 2020.

[31] V. Gasimov, E. Mustafayeva, "Implementing covert channels to transfer hidden information over whatsapp on mobile phones," *American Journal of Engineering and Applied Sciences*, vol. 6, no. 2, pp. 32–35, 2020.

[32] P. Gladyshev, A. Patel, "Finite state machine approach to digital event reconstruction," *Digital Investigation*, vol. 1, pp. 130–149, 2004.

[33] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," *arXiv:1406.2661*, 2014.

[34] F. Gottwalt, E. Chang, and T. Dillon, "CorrCorr: A feature selection method for multivariate correlation network anomaly detection techniques," *Computer & Security*, vol. 83, pp. 234–245, 2019.

[35] Guidance Software, *Encase Forensic*, May 21, 2021. (ttps://www.guidancesoftware.com/)

[36] Guidance Software, *Opentext Encase Forensi*, May 21, 2021. (https://security.opentext.com/encase-forensic)

[37] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. C. Courville, "Improved training of Wasserstein GANs," in *Proceedings of Conference on Neural Information Processing Systems*, pp. 5767–5777, 2017.

[38] C. Hargreaves and J. Patterson, "An automated timeline reconstruction approach for digital forensic investigations," *Digital Investigation*, vol. 9, pp. S69–S79, 2012.

[39] M. M. Hassan, A. Gumaei, A. Alsanad, M. Alrubaian, and G. Fortin, "A hybrid deep learning model for efficient intrusion detection in big data environment," *Information Sciences*, vol. 513, pp. 386–396, 2020.

[40] W. Hong, T. S. Chen, Y. P. Chang, and C. W. Shiu, "A high capacity reversible data hiding scheme using orthogonal projection and prediction error modification," *Signal Processing*, vol. 90, pp. 2911–2922, 2010.

[41] C. H. Hsieh and J.C. Tsai, "Lossless compression of VQ index with search order coding," *IEEE Transaction on Image Processing*, vol. 5, no. 1, pp. 1579–1582, 1996.

[42] 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, p. 250–260, 2009.

[43] C. T. Huang, M. Y. Tsai, L. C. Lin, W. J. Wang, and S. J. Wang, "VQ-based data hiding in iot networks using two-level encoding with adaptive pixel replacements," *Journal of Supercomputing*, p. 1–20, 2016.

[44] C. T. Huang, W. J. Wang, C. H. Yang, and S. J. Wang, "A scheme of reversible information hiding based on SMVQ," *Imaging Science Journal*, vol. 61, no. 2, p. 195–203, 2013.

[45] S. Huang and K. Lei, "IGAN-IDS: An imbalanced generative adversarial network towards intrusion detection system in ad-hoc networks," *Ad Hoc Networks*, vol. 105, p. 102177, 2020.

[46] B. Ingre, A. Yadav, and A. K. Soni, "Decision tree based intrusion detection system for NSL-KDD dataset," in *2017 International Conference on Information and Communication Technology for Intelligent Systems (ICTIS 2017)*, vol. 2, pp. 207–218, 2017.

[47] A. Joseph and K. J. Singh, "A survey on latest trends and challenges in cyber forensics," in *Proceedings of International Conference on Mechanical, Civil, Electronics and Communications and Computer Science Engineering*, vol. 4, 2016.

[48] K. H. Jung and K. Y. Yoo, "Improved exploiting modification direction method by modulus operation," *International Journal of Signal Processing, Image Processing and Pattern*, vol. 2, p. 79–88, 2009.

[49] S. Kalber, A. Dewald, and F. C. Freiling, "Forensic application-fingerprinting based on file system metadata," in *Proceedings of 2013 Seventh International Conference on IT Security Incident Management and IT Forensics*, 2013.

[50] KDD, *KDD Cup 1999*, Oct. 2007. (http://kdd.ics.uci.edu/databases/kddcup99\/kddcup99.html)

[51] M. N. A. Khan, "Performance analysis of bayesian networks and neural networks in classification of file system activities," *Computers and Security*, vol. 31, pp. 391–401, 2013.

[52] M. N. A. Khan, C. R. Chatwin, and R. C. D. Young, "A framework for post-event timeline reconstruction using neural networks," *Digital Investigation*, vol. 4, pp. 146–157, 2007.

[53] C. Kim, D. Shin, and C. N. Yang, "Self-embedding fragile watermarking scheme to restoration of a tampered image using AMBTC," *Personal and Ubiquitous Computing*, vol. 22, pp. 11–22, 2018.

[54] H. J. Kim, V. Sachnev, Y. Q. Shi, J. Nam, and H. G. Choo, "A novel difference expansion transform for reversible data embedding," *IEEE Trans. Inf. Forensics Security*, vol. 3, p. 456–465, 2008.

[55] G. Kocher and G. Kumar, "Machine learning and deep learning methods for intrusion detection systems: recent developments and challenges," *Soft Computing*, vol. 25, pp. 9731–9763, 2021.

[56] M. V. Kotpalliwar and R. Wajgi, "Classification of attacks using support vector machine (SVM) on KDDCUP'99 IDS database," in *2015 Fifth International Conference on Communication Systems and Network Technologies*, pp. 987–990, 2015.

[57] W. C. Kuo and C. C. Wang, "Data hiding based on generalized exploiting modification direction method," *The Imaging Science Journal*, vol. 61, p. 484–490, 2013.

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

[59] C. F. Lee, C. C. Chang, P. Y. Pai, and C. M. Liu, "Adjustment hiding method based on exploiting modification direction," *International Journal of Network Security*, vol. 17, no. 5, pp. 607–618, 2015.

[60] C. F. Lee, C. C. Chang, and K. H. Wang, "An improvement of EMD embedding method for large payloads by pixel segmentation strategy," *Journal of Image and Vision Computing*, vol. 26, no. 12, pp. 1670–1676, 2008.

[61] C. F. Lee and H. L. Chen, "A novel data hiding scheme based on modulus function," *Journal of Systems and Software*, vol. 83, no. 5, pp. 832–843, 2010.

[62] C. F. Lee, H. L. Chen, and S. H. Lai, "An adaptive data hiding scheme with high embedding capacity and visual image quality based on SMVQ prediction through classification codebooks," *Image and Vision Computing*, vol. 28, no. 8, p. 1293–1302, 2010.

[63] C. F. Lee, Y. C. Li, S. C. Chu, and J. F. Roddick, "Data hiding scheme based on a flower-shaped reference matrix," *Journal of Network Intelligence (JNI)*, vol. 3, no. 2, pp. 138–151, 2018.

[64] C. F. Lee, J. J. Shen, S. Agrawal, Y. X. Wang, and Y. H. Lee, "Data hiding method based on 3D magic cube," *IEEE ACCESS*, vol. 8, pp. 39445– 39453, 2020.

[65] C. F. Lee, J. J. Shen, Z. R. Chen, and S. Agrawal, "Self-embedding authentication watermarking with effective tampered," *Sensors*, vol. 19, no. 10, pp. 2267–2284, 2019.

[66] C. F. Lee, J. J. Shen, Y. C. Kao, and S. Agrawal, "Overlapping pixel value ordering predictor for high-capacity reversible data hiding," *Real-Time Image Process*, pp. 1–21, 2019.

[67] C. F. Lee, J. J. Shen, and Y. H. Lai, "Data hiding using multi-pixel difference expansion," in *International Conference on Computer and Communication Systems*, p. 56–60, 2018.

[68] C. F. Lee, J. J. Shen, and K. T. Lin, *The Study of Steganographic Algorithms Based on Pixel Value Difference*, vol. 63, Springer International Publishing, 2017.

[69] C. F. Lee and Y. X. Wang, "Secure image hiding scheme based on magic signet," *Journal of Electronic Science and Technology (JEST)*, vol. 18, no. 1, pp. 93– 101, 2020.

[70] C. F. Lee, Y. X. Wang, and A. T. Shih, "Image steganographic method based on pencil-shaped pattern," *Lecture Notes in Electrical Engineering(LNEE)*, vol. 425, pp. 639–644, 2018.

[71] C. F. Lee, C. Y. Weng, and S. Aneesh, "Steganographic access control in data hiding using runlength encoding and modulo-operations," *Security and Communication Networks*, vol. 6, no. 3, pp. 277–284, 2011.

[72] C. F. Lee, C. Y. Weng, and K. C. Chen, "An efficient reversible data hiding with reduplicated exploiting modification direction using image interpolation and edge detection," *Multimedia Tools and Applications*, vol. 76, no. 7, p. 9993–10016, 2017.

[73] C. F. Lee, C. Y. Weng, and C. Y. Kao, "Reversible data hiding using lagrange interpolation for prediction-error expansion embedding," *Soft Computing*, vol. 23, no. 19, p. 9719–9731, 2019.

[74] H. S. Leng, *Data Hiding Scheme Based on Regular Octagon-Shaped Shells*, vol. 81, Springer International Publishing, 2018.

[75] J. J. Li, Y. H. Wu, C. F. Lee, and C. C. Chang, "Generalized PVO-K embedding technique for reversible data hiding," *International Journal of Network Security*, vol. 20, no. 1, pp. 65–77, 2018.

[76] X. Li, W. Chen, Q. Zhang, and L. Wu, "Building auto-encoder intrusion detection system based on random forest feature selection," *Computer & Security*, vol. 95, p. 101851, 2020.

[77] 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, p. 198–205, 2013.

[78] C. C. Lin, S. C. Chen, and N. L. Hsueh, "Adaptive embedding techniques for VQ-compressed images," *Information Sciences*, vol. 179, no. 1–2, p. 140–149, 2009.

[79] C. Y. Lin and S. F. Chang, "A robust image authentication method distinguishing JPEG compression from malicious manipulation," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, no. 2, pp. 153–168, 2001.

[80] T. Lin, P. Goyal, R. Girshick, K. He, and P. Dollar, "Focal loss for dense object detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, pp. 318–327, February 2020.

[81] Z. Lin, Y. Shi, and Z. Xue, "IDSGAN: Generative adversarial networks for attack generation against intrusion detection," *arXiv: 1809.02077v2*, September 2018.

[82] H. Liu and B. Lang, "Machine learning and deep learning methods for intrusion detection systems: A survey," *Applied Sciences*, vol. 9, no. 20, p. 4396, 2019.

[83] J. C. Liu and M. H. Shih, "Generalizations of pixel-value differencing steganography for data hiding in images," *Fundamenta Informaticae*, vol. 83, no. 3, p. 319–335, 2008.

[84] S. H. Liu, H. X. Yao, W. Gao, and Y. L. Liu, "An image fragile watermark scheme based on chaotic image pattern and pixel-pairs," *Applied Mathematics and Computation*, vol. 185, no. 2, pp. 869–882, 2007.

[85] Y. Liu, C. C. Chang, and T. S. Nguyen, "High capacity turtle shell-based data hiding," *IET Image Processing*, vol. 10, no. 2, pp. 130–137, 2016.

[86] Z. M. Lu, J. X. Wang, and B. B. Liu, "An improved lossless data hiding scheme based on image VQ-index residual value coding," *Journal of Systems and Software*, vol. 82, no. 6, pp. 1016–1024, 2009.

[87] L. Luo, Z. Chen, M. Chen, X. Zeng, and Z. Xiong, "Reversible image watermarking using interpolation technique," *IEEE Transactions on Information Forensics and Security*, vol. 5, pp. 187–193, 2010.

[88] M. S. I. Mamun, M. A. Rathore, A. H. Lashkari, N. Stakhanova, and A. A. Ghorbani, "Detecting malicious URLs using lexical analysis," in *2016 International Conference on Network and System Security (NSS 2016)*, pp. 467–482, 2016.

[89] F. Marturana and S. Tacconi, "A machine learning-based triage methodology for automated categorization of digital media," *Digital Investigation*, vol. 10, pp. 193–204, 2013.

[90] Moonsols, *Moonsols*, May 21, 2021. (`http://www.moonsols.com/`)

[91] N. Moustafa and J. Slay, "UNSW-NB15: A comprehensive data set for network intrusion detection systems (UNSW- NB15 network data set)," in *Military Communications and Information Systems Conference (MilCIS 2015)*, pp. 1–6, 2015.

[92] 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, p. 354–362, 2006.

[93] B. Ou, X. Li, and J. Wang, "Improved PVO-based reversible data hiding: A new implementation based on multiple histograms modification," *Journal of Visual Communication and Image Representation*, vol. 38, p. 328–339, 2016.

[94] 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, p. 760–772, 2014.

[95] F. Peng, X. Li, and B. Yang, "Improved PVO-based reversible data hiding," *Digital Signal Processing*, vol. 25, p. 255–265, 2014.

[96] K. Peng, V. C. M. Leung, and Q. Huang, "Clustering approach based on mini batch Kmeans for intrusion detection system over big data," *IEEE Access*, vol. 6, pp. 11897–11906, 2018.

[97] M. S. Pervez and D. M. Farid, "Feature selection and intrusion classification in NSL-KDD cup 99 dataset employing SVMs," in *The 8th International Conference on Software, Knowledge, Information Management and Applications (SKIMA 2014)*, pp. 1–6, 2014.

[98] E. Popovsky and Frincke, "Embedding forensic capabilities into networks: Addressing inefficiencies in digital forensics investigations," in *Proceedings of 2006 IEEE Information Assurance Workshop*, pp. 133–139, 2006.

[99] Z. Qian, G. Feng, X. Zhang, and S. Wang, "Image self-embedding with high-quality restoration capability," *Digital Signal Processing*, vol. 21, no. 2, pp. 278–286, 2011.

[100] C. Qin, P. Ji, X. Zhang, J. Dong, and J. Wang, "Fragile image watermarking with pixel-wise recovery based on overlapping embedding strategy," *Signal Processing*, vol. 138, pp. 280–293, 2017.

[101] Y. Qiu, Z. Qian, and L. Yu, "Adaptive reversible data hiding by extending the generalized integer transformation," *IEEE Signal Processing Letters*, vol. 23, no. 1, pp. 409– 413, 2016.

[102] X. Qu and H.J. Kim, "Pixel-based pixel value ordering predictor for high-fidelity reversible data hiding," *Signal Processing*, vol. 111, p. 249–260, 2015.

[103] V. Roussev and C. Quates, "Content triage with similarity digests: The m57 case study," *Digital Investigation*, vol. 9, pp. 60–68, 2012.

[104] S. Sahu and B. M. Mehtre, "Network intrusion detection system using J48 decision tree," in *2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI 2015)*, pp. 2023–2026, 2015.

[105] I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "Toward generating a new intrusion detection dataset and intrusion traffic characterization," in *Proceedings of the 4th International Conference on Information Systems Security and Privacy (ICISSP 2018)*, pp. 108–116, 2018.

[106] J. J. Shen and P. W. Hsu, "A robust associative watermarking technique based on similarity diagrams," *Pattern Recognition*, vol. 40, no. 4, pp. 1355–1367, 2007.

[107] J. J. Shen, C. F. Lee, F. W. Hsu, and S. Agrawal, "A self-embedding fragile image authentication based on singular value decomposition," *Multimedia Tools and Applications*, vol. 79, p. 25969–25988, 2020.

[108] J. J. Shen and J. M. Ren, "A robust associative watermarking technique based on vector quantization," *Digital Signal Processing*, vol. 20, no. 5, p. 1408–1423, 2010.

[109] D. Singh and S. K. Singh, "Effective self-embedding watermarking scheme for image tampered detection and localization with recovery capability," *Journal of Visual Communication and Image Representation*, vol. 38, pp. 775–789, 2016.

[110] S. Soltani and S. A. H. Seno in *Proceedings of 2017 7th International Conference on Computer and Knowledge Engineering*, pp. 247–253, 2017.

[111] N. Son, Y. Lee, D. Kim, I. James, S. Lee, and K. Lee, "A study of user data integrity during ac-

quisition of android devices," *Digital Investigation*, vol. 10, pp. 3–11, 2013.

[112] R. M. Stevens and E. Casey, "Extracting windows command line details from physical memory," *Digital Investigation*, vol. 7, pp. 57–63, 2010.

[113] J. Stüttgen and M. Cohen, "Anti-forensic resilient memory acquisition," *Digital Investigation*, vol. 10, pp. 105–114, 2013.

[114] G. Swain, "Adaptive pixel value differencing steganography using both vertical and horizontal edges," *Multimedia Tools and Applications*, vol. 75, p. 13541–13556, 2016.

[115] A. Telikani and A. H. Gandomi, "Cost-sensitive stacked auto-encoders for intrusion detection in the Internet of things," *Internet of Things*, vol. 14, p. 100122, 2021.

[116] M. Thodi and J. J. Rodríguez, "Expansion embedding techniques for reversible watermarking," *IEEE Transactions on Image Processing*, vol. 16, p. 721–730, 2007.

[117] J. Tian, "Reversible data embedding using a difference expansion," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 8, p. 890–896, 2003.

[118] C. T. Tung, C. P. Yen, "VoIP packets filtering for mobile instant messaging using N-gram models," *International Journal of Network Security*, vol. 23, no. 6, pp. 1108-1117, 2021.

[119] T. Vidas, B. Kaplan, and M. Geiger, "Openlv: Empowering investigators and first-responders in the digital forensics process," *Digital Investigation*, vol. 11, pp. 45–53, 2014.

[120] C. C. Wang, W. C. Kuo, Y. C. Huang, and L. C. Wuu, "A high capacity data hiding scheme based on readjusted GEMD," *Multimedia Tools and Applications*, vol. 8, p. 1–15, 2017.

[121] C. C. Wang, T. H. Lin, W. C. Kuo, "Data hiding based on the Chinese remainder theorem," *International Journal of Network Security*, vol. 23, no. 6, pp. 1118-1125, 2021.

[122] J. Wang, Y. Sun, H. Xu, K. Chen, H. J. Kim, and S. H. Joo, "An improved section-wise exploiting modification direction method," *Signal Processing*, vol. 90, p. 2954–2964, 2010.

[123] W. J. Wang, C. T. Huang, S. R. Tsuei, and S. J. Wang, "A greedy steganographic SMVQ approach of greedy-USBIRDS using secret bits for image-block repairing based on differences," *Multimedia Tools and Applications*, vol. 75, no. 22, p. 14895–14916, 2016.

[124] X. Wang, X. Li, B. Yang, and Z. M. Guo, "Efficient generalized integer transform for reversible watermarking," *IEEE Signal Processing Letters*, vol. 17, no. 6, pp. 567–570, 2010.

[125] X. T. Wang, C. C. Chang, C. C. Lin, and M. Li, "A novel multi-group exploiting direction method based on switch map," *Signal Processing*, vol. 92, p. 1525–1535, 2012.

[126] M. Weedon, D. Tsaptsinos, and J. Denholm-Price, "Random forest explorations for URL classification," in *2017 International Conference on Cyber Situational Awareness, Data Analytics and Assessment (CyberSA 2017)*, pp. 1–4, 2017.

[127] D. C. Wu and W. H. Tsai, "A steganographic method for images by pixel-value differencing," *Pattern Recognition Letters*, vol. 24, no. 9-10, p. 1613–1626, 2003.

[128] H. C. Wu, N. I. Wu, C. S. Tsai, and M. S. Huang, "Image steganographic scheme based on pixel-value differencing and LSB replacement methods," *IEE Proceedings Vision Image and Signal Processing*, vol. 152, no. 5, p. 611–615, 2005.

[129] Y. Xin, L. Kong, Z. Liu, Y. Chen, Y. Li, H. Zhu, M. Gao, H. Hou, and C. Wang, "Machine learning and deep learning methods for cybersecurity," *IEEE Access*, vol. 6, pp. 35365–35381, 2018.

[130] C. H. Yang, S. J. Wang, C. Y. Weng, and H. M. Sun, "Information hiding technique based on blocked PVD," *Journal of Information Management*, vol. 15, no. 3, p. 29–48, 2008.

[131] C. H. Yang, W. J. Wang, C. T. Huang, and S. J. Wang, "Reversible steganography based on side match and hit pattern for VQ-compressed images," *Information Sciences*, vol. 181, no. 11, p. 2218–2230, 2011.

[132] C. H. Yang, C. Y. Weng, H. K. Tso, and S. J. Wang, "A data hiding scheme using the varieties of pixel-value differencing in multimedia images," *Journal of Systems and Software*, vol. 84, no. 4, p. 669–678, 2011.

[133] B. B. Zaidan, A. A. Zaidan, A. K. Al-Frajat, and H. A. Jalab, "On the differences between hiding information and cryptography techniques: An overview," *Journal of Applied Sciences*, vol. 10, no. 15, pp. 1650–1655, 2010.

[134] H. Zhang, L. Huang, C. Q. Wu, and Z. Li, "An effective convolutional neural based on SMOTE and Gaussian mixture model for intrusion detection in imbalanced dataset," *Computer Networks*, vol. 177, p. 107315, 2020.

[135] J. Zhang, Y. Ling, X. Fu, X. Yang, G. Xiong, and R. Zhang, "Model of the intrusion detection system based on the integration of spatial-temporal features," *Computers & Security*, vol. 89, p. 101681, 2020.

[136] X. Zhang, Y. Han, W. Xu, and Q. Wang, "HOBA: A novel feature engineering methodology for credit card fraud detection with a deep learning architecture," *Information Sciences*, vol. 557, pp. 302–316, May 2021.

[137] X. Zhang and S. Wang, "Efficient steganographic embedding by exploiting modification direction," *IEEE Communications Letters*, vol. 10, no. 11, p. 781–783, 2006.

[138] Y. Zhou, G. Cheng, S. Jiang, and M. Dai, "Building an efficient intrusion detection system based on

feature selection and ensemble classifier," *Computer Networks*, vol. 174, p. 107247, 2020.

[139] Y. Zhu, J. Liang, J. Chen, and Z. Ming, "An improved NSGA-III algorithm for feature selection used in intrusion detection," *Knowledge-Based Systems*, vol. 116, pp. 74–85, 2017.

Biography

Chin-Feng Lee received her Ph.D. in Computer Science and Information Engineering from National Chung Cheng University, Taiwan in 1998. She is currently a professor of Information Management at Chaoyang University of Technology, Taiwan. Her research interests include image processing, watermarking, and machine learning.

Chi-Yao Weng received his Ph.D. degree in computer science from National Tsing Hua University, Hsinchu, Taiwan, in 2011. From 2011 to 2015, he was a Postdoctoral researcher at National Sun Yat-sen University and National Tsing Hua University, respectively. From August 2015 to January 2019, he was an assistant professor, and now an associate professor of the Department of Computer Science at National Pingtung University, Pingtung, Taiwan. His research interests include multimedia security, image privacy, information forensics, and mobile security.

Chih-Hung Wang was born in Kaohsiung, Taiwan in 1968. He received the BS degree in Information Science from Tunghai University and MS degree in Information Engineering from National Chung Cheng University, Taiwan in 1991 and 1993, respectively. He received the Ph.D. degree in Information Engineering from National Cheng Kung University, Taiwan in 1998. He is presently a full professor of Department of Computer Science and Information Engineering, National Chiayi University, Taiwan. His research interests include cryptography, network security, intrusion detection and data compression.

Goutam Chakraborty received his Ph.D. in Research Institute of Electrical Communication from Tohoku University, Sendai, Japan in 1993. He is a tenured full professor and head of Intelligent Informatics Lab. in the Department of Software & Information Science, Iwate Prefectural University, Japan. His research interests include evolutionary computations, wireless communication and networking, and machine learning.

Kouichi Sakurai (Member, IEEE) received the B.S. degree in mathematics from the Faculty of Science, Kyushu University, in 1986, the M.S. degree in applied science and the Ph.D. degree in engineering from the Faculty of Engineering, Kyushu University, in 1988 and 1993, respectively. From 1988 to 1994, he was engaged in research and development on cryptography and information security at the Computer and Information Systems Laboratory, Mitsubishi Electric Corporation. Since 1994, he has been working with the Department of Computer Science, Kyushu University, as an Associate Professor, where he became a Full Professor, in 2002. He is concurrently working with the Institute of Systems, Information Technologies and Nanotechnologies, as the Chief of the Information Security Laboratory, for promoting research co-operations among the industry, university, and government under the theme Enhancing IT-Security in Social Systems. From 2005 to 2006, he was successful in generating such co-operation between Japan, China, and Korea, for security technologies, as a Leader of the Cooperative International Research Project supported by the National Institute of Information and Communications Technology (NICT). Moreover, in March 2006, he established research co-operations under a Memorandum of Understanding in the field of information security with Prof. Bimal Kumar Roy, the first time Japan has partnered with The Cryptology Research Society of India (CRSI). He currently directs the Laboratory for Information Technology and Multimedia Security and is working with the CyberSecurity Center, Kyushu University. He is also with Department of Advanced Security, Advanced Telecommunications Research Institute International and involved in a NEDO-SIP-Project on supply chain security. He has published about 400 academic articles in cryptography and cybersecurity.

Kuo-Yu Tsai received the M.S. and Ph.D. degrees from the Department of Information Management, National Taiwan University of Science and Technology, in 2001 and 2009, respectively. He is currently an Assistant Professor with the Department of Information Engineering and Computer Science, Feng Chia University, Taiwan. He is a member of the Chinese Cryptology and Information Security Association (CCISA) and the E-Security Analysis and Management (E-SAM), and was elected as the Director of E-SAM, in 2018. His recent research interests include cryptography, network security, and cloud computing.

VoIP Packets Filtering for Mobile Instant Messaging Using N-gram Models

Cheng-Tan Tung and Chih-Ping Yen
(Corresponding author: Chih-Ping Yen)

Department of Information Management, Central Police University
Taoyuan 33304, Taiwan, R.O.C.
Email: peter@mail.cpu.edu.tw

(Received Nov. 6, 2020; Revised and Accepted June 21, 2021; First Online Oct. 19, 2021)
Special Issue on Multimedia Application on Information Hiding Forensics and Cybersecurity

Abstract

Since the vigorous development of mobile devices, such as smartphones, their communication functions have gradually replaced traditional traffic methods to become mainstream communication channels. However, given the increase in users emphasizing their right to privacy, communication software products are expected to encrypt the packets transmitted during communication to meet user requirements. Meanwhile, MIM (Mobile Instant Messaging) is increasingly being used by criminals as their medium of communication. The packets transmitted during such communication cannot be easily decrypted and understood by criminal investigators. Therefore, this research aims to maximize the use of the limited information provided by unknown packets and use the principle of VoIP (Voice over Internet Protocol) communication to establish a P2P (peer-to-peer) connection. Wireshark is used to monitor popular communication software, and it is tested in terms of whether a call process can be conducted at both ends. As a result, a P2P connection is established and integrated with the features of unknown packets. This work proposes that N-gram models other than the IP address of a packet can be used to find P2P connection packets. The experiment results show that the proposed model obtains an Accuracy of 95.77%, Precision of 98.36%, Recall of 93.4%, and F1-score of 95.81%. Thereby assisting criminal investigators in identifying criminals and improving the efficiency of criminal investigations are the purpose of this article.

Keywords: MIM; VoIP; P2P; N-gram

1 Introduction

A P2P network is an Internet system that has no central server and allows peers to exchange information. Its purpose is to (1) enable all peers to provide resources, including bandwidth, storage space and computing power. Therefore, when a new peer joins, the capacity of the entire system also increases. (2) Reduce the number of nodes in network transmission to avoid the risk of data loss [3]. The P2P applications include group software, file exchange, MIM, grid computing, Web Services, and business process exchange.

With the advancement of mobile phone functions and the development of network communications, MIM is gradually replacing the role of traditional local calls or mobile phones as a tool for modern communication. However, despite its convenience, MIM has become a platform for criminals, and the communication records in the form of packets that are involved in related crimes are increasingly challenging to identify and understand as they are transmitted across the Internet. MIM comes in various forms in the market. In Taiwan, the most common ones include LINE, Facebook Messenger, WeChat, and FaceTime [21]. Different MIM applications adopt specific communication format protocols and communication features. In the application of packet capture tools, the particular packet features are used to determine the MIM used and the user's usage status [4,17] Therefore, quickly and efficiently collecting information for the investigation of a large number of packets is a problem that criminal investigators must overcome. At present, for MIM investigations, criminal investigators use some digital forensics software tools, such as EnCase, FTK, Cellebrite UFED, XRY. Then use personal experience to manually observe the data packet header information (such as time, protocol, length) or the string characteristics of the data packet payload. Then, relevant instructions for using packet capture software are provided to assist investigators in reviewing voluminous packet data and efficiently filter VoIP packets with investigative value [17]. The observed features should be applied to automated filtering methods to shorten the packet analysis process further. We refer to the method proposed by Walnycky *et al.* [22] and use the principle of MITM under legal authorization for monitoring in a Wi-Fi environment. Taking the popular MIM platform in iOS App Store and Android Play

Store as the research object, we propose an automated method that can effectively filter VoIP packets using the N-gram model, thereby assisting criminal investigators in quickly identifying criminals and improving the efficiency of criminal investigation.

The rest of this paper is organized as follows. In the next section, we present related works. In Section 3, we introduce the proposed architecture. In Section 4, we present the experiment results. Finally, we summarize the conclusions.

2 Related Works

2.1 Feature and Filtering of VoIP Packets

For research on the subject of VoIP packets filtering, there are network attack, steganography & steganalysis, spamming, QoS (Quality of Service), fraud, and encrypted communications, etc. Wu [23] mentioned that some VoIP applications would use the technology of silence suppression, which reduces the sending rate of packets when one end of the call is not talking, thereby saving network bandwidth. The specific judgment method mentioned in the article is based on the threshold of 100 milliseconds. If there is no packet transmission for more than this time, the user may be in a listening state. Compared with other UDP protocol types of transmission, the packet flow of VoIP voice communication is more effective. According to experiments, there are not only more packets for voice-transmitted call traffic, but also a significant difference in average packet size. Stöber [18] found through experiments that most applications generate two types of packet traffic. About 70% of the traffic is initiated by the application itself, while only 30% produce when being used by users. Therefore, when the mobile device's packet traffic is intercepted, many packets that are not voluntarily generated by the user are often collected. At this time, you should try to filter to avoid affecting the observation result. Rezaei and Liu [15] also proposed for packet classification four methods, including (1) time series, (2) packet header, (3) payload data, and (4) statistical features. The above methods can be used together to increase the packet classification feature. Among them, the statistical feature method requires a large amount of traffic to calculate the relevant features of the packet length, so it is more suitable for offline classification. Mehta *et al.* [13], the network packet traffic on the port can be used slogans or protocol traffic into (1) Sensitive traffic: care to convey the time, such as VoIP, video conferencing, online gaming, through a develop traffic management in this way to ensure the service efficiency of traffic and increase the priority of such traffic. (2) Best-effort traffic: For service quality index (jitter, packet loss, delay), less sensitive traffic such as P2P, e-mail, traffic management is relatively fixed, it will usually be sent after sensitive traffic. (3) Undesired traffic (bad traffic): typically created by worms or other malicious botnet spam attacks. Tan [19] considers packet flow is described by its features, including source IP, source port, destination IP, destination port, arrival time, protocol, and packet length. More than 200 statistical attributes can be calculated using only packet header information, such as packet arrival time, average packet arrival time of traffic, number of downloaded packets, uploaded packets, packet probability density function, cumulative density function upstream, and downstream data packet density functions. And the cumulative density function, the number of one-way communication and its duration (download duration), etc.

2.2 Open System Interconnection Reference Model (OSI model)

The OSI model is a standard framework proposed by the International Organization for Standardization (ISO) that attempts to interconnect all kinds of computers as a network worldwide. The purpose is to solve the network interconnection problems of different architectures. The packet inspector used in this experiment of the research is Wireshark. With its graphical interface, it shows the relevant information of the data link layer, network layer, and transmission layer in the OSI model. The main function of the data link layer is to define how to package the transmitted data into the specifications of the data packet format. This layer contains the Media Access Control (MAC) address. Wireshark will display the mobile phone model and MAC address used for communication. The MAC address is a set of serial numbers. The MAC address of each network device is unique. This address information allows the host to identify the physical device. The network layer defines network routing and addressing functions, that is, the path selection of the packet transmission process and the address of each physical device in the network. The most important communication protocol in this layer is the Internet Protocol (IP), which mainly defines the format of IP data and the standard IP address format. During data transmitting, the IP address will be added to the transmitted data, and the data will be formed into a packet, and the router will analyze the IP information in the packet and select the path through which the packet should be transmitted. The transport layer is responsible for data transmission and control, records the real sender and destination (port) information to provide higher-level applications. This layer provides mostly two end-to-end protocols, including Transmission Control Protocol (TCP) and User Datagram Protocol (UDP). Because UDP [11] provides unreliable, connectionless data transmission, it is usually used in domain name system (DNS), dynamic host configuration protocol (DHCP), simple network management protocol (SNMP). In addition, for connections that require transmission speed rather than quality assurance, such as video, voice, and calls, UDP is a better choice than TCP when a small amount of packets is lost without affecting the overall use. Figure 1 shows a UDP packet format.

Bytes	0 ~ 1	2 ~ 3
0	Source port	Destination port
1	Length	Check sum
data		

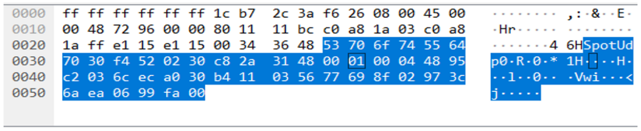
Figure 1: UDP packet format

2.3 VoIP

VoIP [5] mainly uses IP data packets to be exchanged between various network elements to complete the establishment and termination of a call. The principle is that the analog audio signal is compressed and coded into a digital signal through the caller device, and then encapsulated in an IP packet, which is transmitted over the network by TCP/IP and other related network protocols. Compared with traditional line telephones, VoIP is mainly based on the existing IP technology to develop upward. As long as any physical node has the ability to use the Internet, it can use the convenience brought by VoIP technology. Therefore, the current VoIP communication technology is gradually replacing traditional telephones, embedded in various communication software, and becoming the main communication channel for modern people. Since the principle of VoIP calls comes from the existing network communication technology, if you want to make a network call, you need at least two network terminal devices (such as mobile phones) IP addresses and port numbers. Therefore, a management server is required to manage and record the correspondence and conversion between the IP address and the actual phone. When users connect to the server for the first time, they need to register. At this moment, the server will record the user's phone number or account number and IP for subsequent connection calls. And this is the basic function of the server of a mobile phone application with a communication function [2].

2.4 Communication Protocol Tool for VoIP

With the rapid development of the Internet, the demand for network IP is increasing. In order to avoid the problem of insufficient IP addresses, the design of public network IP and private network IP is separated, and only public network IP can directly connect to the Internet. In contrast, private IP addresses are used by internal units such as enterprises or academics. But also because the private address cannot be transmitted through the Internet, it is necessary to correspond the private address with the public address, so that the private address can be connected to the Internet in the form of a public address. The technology in which addresses correspond to each other is called Network Address Translation (NAT). On the other hand, in the communication software widely used by users on the market, there will be P2P connections during the communication. The communication protocol can be roughly divided into two types: UDP and STUN (Session Traversal Utilities for NAT), each of which different purposes. The UDP protocol is responsible for converting the voice analog signal into a digital signal when the user talks, and transfers it into the IP packet for transmission. After deducting the first 42 bytes of the packet header, the remaining part is the packet payload (the blue element in Figure 2).



```
0000  ff ff ff ff ff ff 1c b7  2c 3a f6 26 08 00 45 00   ········ ,:·&··E·
0010  00 48 72 96 00 00 80 11  11 bc c0 a8 1a 03 c0 a8   ·Hr····· ········
0020  1a ff e1 15 e1 15 00 34  36 48 53 70 6f 74 55 64   ·······4 6HSpotUd
0030  70 30 f4 52 02 30 c8 2a  31 48 00 01 00 04 48 95   p0·R·0·* 1H····H·
0040  c2 03 6c ec a0 30 b4 11  03 56 77 69 8f 02 97 3c   ··l··0·· ·Vwi···<
0050  6a ea 06 99 fa 00                                  j·····
```

Figure 2: Packet payload

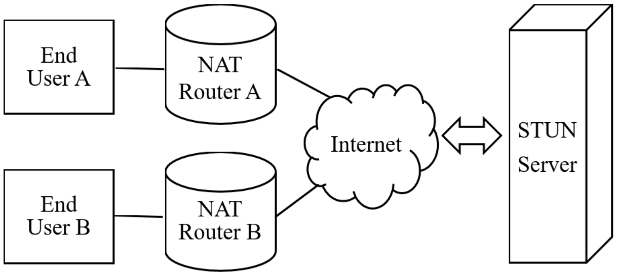


Figure 3: Structure of STUN protocol

The STUN protocol (session traversal utility for NAT) is developed on the basis of NAT. The STUN protocol allows users in a private location in the NAT intranet to know their external public address, port type, and NAT sort [20]. These messages can help two hosts behind different NAT routers establish UDP communication. The STUN protocol is mainly composed of three parts: STUN client, STUN server, and NAT router. The server is deployed on a server connected with multiple public IPs, as shown in Figure 3. The detailed connection process is defined by the RFC 5389 protocol [9].

2.5 Wireshark

Wireshark is a network packet analysis software. This tool is mainly used to capture the traffic passing through a specific network interface and display the detailed status of the packet. Wireshark has now developed into the most well-known open-source application security tool, and it can be run in various operation systems, including Windows, Mac OS, Linux, UNIX, and there are even APP versions for mobile devices on the Internet (mobile phone installation requires permission processing). To understand where the IP of the packet comes from, you can write down the IP address and use a specific website (such as WhosIP) to search in the browser. Still, this search method is time-consuming and inconvenient, so you can download GeoIP (Geographical IP) database (developed by MaxMind), which supports Wireshark and records three parts [6]: city, country, and autonomous system number (ASN) information. Comparing these three parts, we can know the source of the packet faster. ASN represents the routing settings under different network domains, different agencies or units will use this number to separate the network domains managed by each other. Therefore, the number can effectively identify the source of the packet compared to the IP address.

2.6 N-gram

N-gram is an algorithm based on a statistical language model. It is commonly used in natural language processing, data compression, automatic text generation, guesses or prompts entered by search engines, etc. Even further used in spam identification and spam comment identification [1], malware detection [10] The basic idea of N-gram is to operate a sliding window of size N according to the content of the language to form a sequence of byte fragments of length N. However, due to the considerable changes in the logic and sequence of the language, often a change in the order may result in different semantics, so the ability to predict language is currently realized through deep learning [8] N-gram is based on the assumption that the Nth word's appearance is only related to the previous N-1 words, and is not related to any other words. The probability of the entire sentence is the product of the probability of each word. These probabilities can be obtained directly from counting the number of simultaneous occurrences of N words in the corpus. Commonly used are the Uni-Gram (N=1), Bi-Gram (N=2), Tri-Gram (N=3), and Quad-Gram (N=4).

3 The Proposed Architecture

This research proposes that there should be more diversified screening mechanisms for specific ”VoIP packets on P2P connections” in the future. Initially, the filtering rules should be manually selected, such as packet header, load, timing, etc. If there are a large number of data packets, the features will be handed over to machine learning for calculation to find out the feature values of each attribute of the data packets, and finally establish a screening mechanism. In order to allow investigators to access important investigation information more quickly in the future, we propose an automated model architecture, as shown in Figure 4. The proposed model is divided into the training phase and testing phase. In the training phase, the features of the VoIP packet are selected first to filter out the training samples of VoIP, and then the N-gram is constructed to determine the threshold value of judgment. In the testing phase, the test sample is processed with N-gram and then evaluated for performance. The following subsections will detail the key steps in the proposed model.

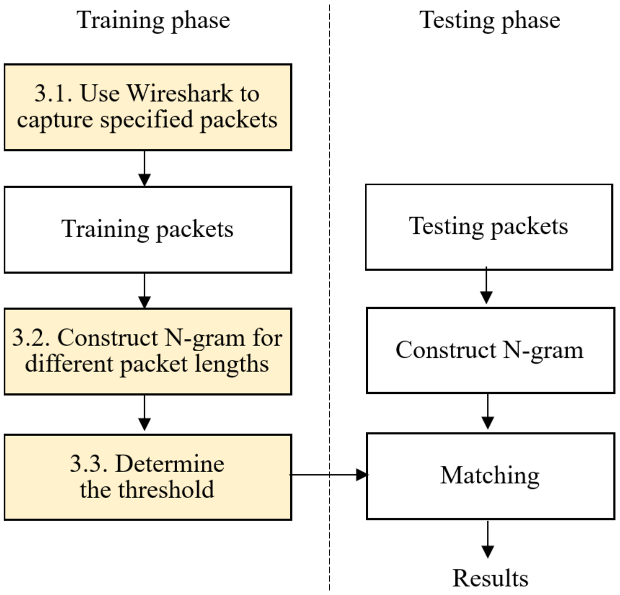
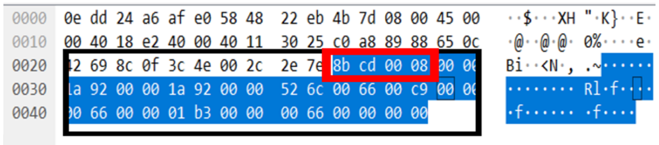

Training phase
Testing phase
3.1. Use Wireshark to capture specified packets
Training packets
3.2. Construct N-gram for different packet lengths
3.3. Determine the threshold
Testing packets
Construct N-gram
Matching
Results

Figure 4: The proposed model architecture

3.1 Use Wireshark to Capture Specified Packets

We collect VoIP packets of various MIM software, including LINE, WeChat, Facebook Messenger, Facetime, QQ, WhatsApp Messenger, Telegram, Jello, Skype, and label features to provide the purpose of capturing training packets.

First, we confirm whether the monitored mobile device has made a call. At this time, we enter ”UDP” in the filter field, which means that as long as Wireshark displays the packets transmitted by the UDP protocol. If there is a packet of this type, which means that this mobile device may have used audio-visual software or made voice/video calls. The second is to use the expanded field of GeoIP to find out the source and destination IP of the UDP packet, and use the IP address and ASN information to determine whether it is a communication service provider. Third, if the source and destination IP is a software distribution company that provides communication services, add ”UDP&&IP.GeoIP.asnum!=xxxx” or ”UDP&&IP.addr!=XXX.XXX.XXX” to the input command ”.XXX” refers to packets that belong to the UDP protocol but exclude specific ASN or IP addresses at the same time, that is, to exclude connection packets between mobile devices and communication software servers. Fourth, if the communication software attempts to establish a P2P connection, it will usually be filtered out. Then, verify whether the filtered packet is the other end of the communication, and we confirm it through the GeoIP field. If it is a P2P connection, the packet will be dynamically configured by the telecommunications company that provides the mobile device to connect to the Internet IP for a P2P connection. Take Taiwan Mobile Co., Ltd. as an example, the field provided by GeoIP



```
0000  0e dd 24 a6 af e0 58 48  22 eb 4b 7d 08 00 45 00   ··$···XH "·K}··E·
0010  00 40 18 e2 40 00 40 11  30 25 c0 a8 89 88 65 0c   ·@··@·@· 0%····e·
0020  42 69 8c 0f 3c 4e 00 2c  2e 7e 8b cd 00 08 00 00   Bi··<N·, .~······
0030  1a 92 00 00 1a 92 00 00  52 6c 00 66 00 c9 00 00   ········ Rl·f····
0040  00 66 00 00 01 b3 00 00  00 66 00 00 00 00         ·f······ ·f····
```

Figure 5: Feature of packet payload

displays the information "AS24158, Taiwan Mobile Co., Ltd.", which shows that its P2P packets are through the Internet of Taiwan Mobile Co., Ltd. Fifth, the statistics of the packet can be viewed from the "Statistics" in the function bar to obtain more information. If the mobile device has a call packet using communication software, we directly use the "Endpoints" in the "Statistics" function to view the relevant geographic records of each IP in the interception record. If the geographic location of the IP source is the target of detection, perform the packet Verify. Algorithm 1 demonstrates the entire procedure of filtering.

```
Algorithm 1 Capture_Specified_Packets
Input: PacketSet=set of all captured packets
Output: SpecifiedPacketSet=set of all the specified packets
function Capture_Specified_Packets(PacketSet)
    for packet in PacketSet do
        SpecifiedPacket=Use Wireshark to capture the packet containing the specified Strings
        insert SpecifiedPacket to SpecifiedPacketSet
    end
end
```

3.2 Construct N-gram for Different Packet Lengths

In the header column of the filtered packet, we use Sequence Number, Time, Destination IP, Source IP, Port, Protocol of packet to make statistics and summarize the features. Besides, the payload of VoIP packets is converted from an analog wave to a digital signal through a codec, and finally encrypted during transmission. Therefore, in the payload of the packet, a specific character string often appears in a specific position. According to the position and frequency of the word string, as shown in Figure 5, it can be compared with the packet header, and it can also be used as a source of feature extraction and then converted into a descriptive feature set.

In addition, different VoIP communication software has different packet length ranges [14]. For the above reasons, in order to extract the features of P2P communication, we first use N-gram to preprocess packets of different lengths. Because the N-gram method is simple and the calculation is fast, we take N consecutive bytes for packet input of different lengths, and N will be set from 1 byte to 8 consecutive bytes. Algorithm 2 demonstrates the entire procedure of constructing N-gram under a packet of specified length. According to experience, more than 8 bytes can no longer find features.

```
Algorithm 2 Construct_Ngram
Input: SpecifiedPacketSet=set of all the specified packets, N=number of consecutive bytes
Output: NgramSet=set of all the N-gram sequences
function Construct_Ngram(SpecifiedPacketSet)
    for packet in SpecifiedPacketSet do
        for i=1:N do
            Ngram=take out consecutive i bytes in packet
            insert Ngram to NgramSet
        end
    end
end
```

3.3 Determine the Threshold

TF-IDF (Term Frequency-Inverse Document Frequency) is a commonly used and effective weighting algorithm based on statistical methods in the fields of information retrieval, text categorization, and mining [12, 16]. TF is the number of times a term appears in a document, which is used to measure the importance of the term in a particular document. The larger the TF value, the more important the term. TF is a count standardized value to prevent bias towards longer documents. Here we will refer to the frequency at which an N-gram appears in the packet, as shown in Equation (1):

$$TF_{i,j} = \frac{n_{i,j}}{\sum_k n_{k,j}}, \tag{1}$$

where $n_{i,j}$ is the number of occurrences of N-gram i in packet j, and $\sum_k n_{k,j}$ is the sum of the number of occurrences of all N-grams in a packet.

IDF is used to measure the importance of a term in the entire document database. The smaller the value, the more the term appears in the majority of documents, which means the less important; the larger the value, the more important the term appears in a few documents. We also quote to express the importance of an N-gram in a specific packet length database, as shown in Equation (2):

$$IDF_i = log\frac{|D|}{1 + |\{j : i \in j | j \in D\}|}, \tag{2}$$

where $|D|$ represents the total number of packets in a particular packet length database, $|\{j : i \in j | j \in D\}|$ denotes the total number of packets containing N-gram i. Thus, the constant one in the denominator is to avoid Division by zero error.

Then, the aforementioned $TF_{i,j}$ and IDF_i are multiplied to produce a $TF\text{-}IDF_{i,j}$ weight value, which tends to retain the important N-gram and filter out the common N-gram. As shown in Equation (3):

$$TF\text{-}IDF_{i,j} = TF_{i,j} \times IDF_i. \tag{3}$$

That is, the importance of an N-gram increase in proportion to the number of times it occurs in the packet; at the same time, its importance decreases inversely with its frequency in the packet database.

Finally, for each VoIP packet of a specific length, we will use the average overlap score evaluation as the threshold to determine whether it is the target VoIP packet. When a detected packet has an average overlap score greater than or equal to this threshold, it is judged as a target VoIP packet; otherwise, it is judged as a general packet. Equation (4) is as follows:

$$Score_{Lj} = avg_{j\in Lj}(\sum_{k} TF\text{-}IDF_{k,j}). \qquad (4)$$

where L_j represents all the training data collections of the same length as packet j, $Score_{Lj}$ is the threshold of L_j to determine whether it is the target VoIP packet, and $avg(.)$ is the average function. Algorithm 3 is a procedure for determining the threshold value for a specific VoIP packet length.

Algorithm 3 Threshold_Determination

Input: M=the total number of specified VoIP packet lengths in the training set, $NgramSet$=set of all the $Ngram$ sequences in a specified VoIP packet length

Output: $Score$=threshold of VoIP packet

```
function Threshold_Determination(NgramSet)
    for $j=1:M$ do
        for Ngram of packet $j$ in NgramSet do
            TF[Ngram]=occurrence frequency array of each Ngram in the packet $j$
            IDF[Ngram]=importance array of each Ngram in the entire packet training set
            TF_IDF[j]=TF[Ngram]*IDF[Ngram]'
        end
        Total=Total+TF_IDF[j]
    end
    Score= Total/M
end
```

4 Experiment Results

4.1 Experimental Environment and Dataset

Figure 6 is a schematic diagram of a packet detection architecture. In the experiment, a laptop will be used to provide a wireless access point for the "packet monitoring mobile phone" to connect. At the same time, the Wireshark packet monitoring software is installed on the laptop. When the mobile phone uses the Wi-Fi provided by the laptop to make an Internet voice call, the laptop will record all network traffic, that is, packets. And the other end of the call is an I-brand mobile phone, which is connected by installing the SIM card mobile data, avoid using the same Wi-Fi environment to surf the Internet, which makes observation difficult. The Wireshark version used in this experiment is 3.0.5, the "packet monitoring phone" is Sony Xperia XA Ultra, which version is Android 7.0, and the other phone is iPhone which version is iOS 12.4.

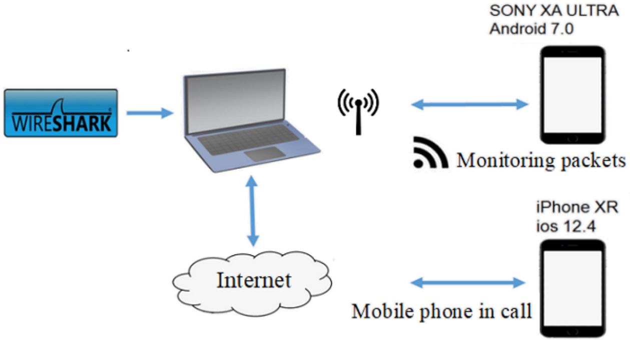


Figure 6: Schematic diagram of monitoring packet architecture

Furthermore, in order to provide performance validation of the proposed method, we employed Wireshark to collect 11,396 samples of VoIP packets and 10,608 samples of non-VoIP packets, forming a packet database of 22,004 in total.

4.2 Founded Features of Various MIM

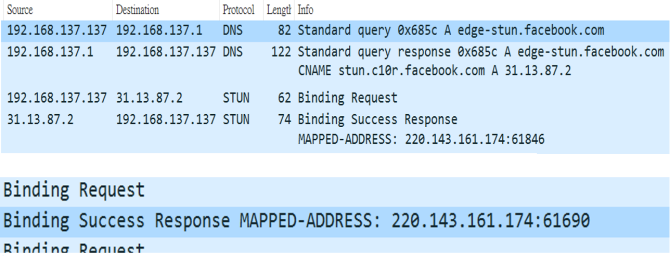
According to Subsection 3.1, in order to label the features of VoIP packets, we use various MIM software, including LINE, WeChat, Facebook Messenger, Facetime, QQ, WhatsAPP Messenger, Telegram, Jello, Skype. Then collect their data packets after establishing a P2P connection. After observation and comparison, the results and features are listed in Table 1. In addition, the following subsections also explain in detail the call features generated by the four more popular tools, such as LINE, Facebook Messenger, WeChat, and Skype.

4.2.1 LINE

During the ringing phase, LINE will send two Application Data packets encrypted with TLSV1.2, and then send a TCP packet with a fixed length of 66. At this time, the IP of the server responsible for connecting with the caller is 203.104.153.0/ 24. After that, the server of 147.92.130.0/24 will send a large number of UDP packets to the caller for the call. So far, the above process can be defined as a fixed feature of LINE to establish a voice call. When the connection is about 20 seconds later, the LINE server will try to establish a P2P connection, and the length of the first eight packets sent in a P2P format will be approximately the same as "290, 290, 335, 335, 170, 174, 174, 174" bytes. However, according to the research of Chen [7], before LINE is converted to P2P mode, it will send a UDP packet, and its payload will transmit the content of the other party's IP and PORT,

Table 1: Call results and features of various MIM

Communication software	Software type	Call server IP	Call protocol	Domain name	Autonomous system number	Whether to establish a P2P connection	Artificial observation of packet features
LINE (10.6.5)	Social software	147.92.130.0/24	UDP	linecorp.com	38631	After ten seconds, occasionally a P2P connection will not be established	Load
WeChat (7.0.12)	Communi-cation software	203.205.220.0/23	UDP	tencent.com	132203	A group of less than 5 packets appears in about 2 seconds during the call	Header
Facebook messenger (260.0)	Communi-cation software	31.13.87.0/24	STUN/ UDP	Facebook.com	32934	Start within a few seconds after the call	Header
Facetime (iOS12.4)	Communi-cation software	17.173.0.0/16	STUN/ UDP	apple.com	714	Only appear within a few seconds, and the P2P packet length is fixed at 58 bytes	Header
QQ (8.3.5)	Communi-cation software	119.28.6.0/23 203.205.236.0/23	UDP	tencent.com	132203	Start within a few seconds after the call, the initial P2P packet length is fixed at 114 bytes	Load
WhatsAPP Messenger (2.20.41)	Communi-cation software	31.13.87.0/24	STUN/ UDP	Facebook.com	32934	Start within a few seconds after the call	Header
Telegram (6.0.1)	Communi-cation software	91.108.16.0/22	UDP	telegram.org	62014 62041 59930	Occurs every 12 seconds during the call, one at a time	Header
Jello (2.18.2)	Communi-cation software	35.200.0.0/14	STUN/ UDP	google.com	15169	Start within a few seconds after the call	Header/ Load
Skype (8.59)	Communi-cation software	20.189.76.0/22 52.114.4.0/22 40.64.0.0/10	STUN/ UDP	microsoft.com	8075	Appears during ringing	Load



```
Source          Destination     Protocol Length Info
192.168.137.137 192.168.137.1   DNS       82 Standard query 0x685c A edge-stun.facebook.com
192.168.137.1   192.168.137.137 DNS      122 Standard query response 0x685c A edge-stun.facebook.com
                                             CNAME stun.c10r.facebook.com A 31.13.87.2

192.168.137.137 31.13.87.2      STUN      62 Binding Request
31.13.87.2      192.168.137.137 STUN      74 Binding Success Response
                                             MAPPED-ADDRESS: 220.143.161.174:61846

Binding Request
Binding Success Response MAPPED-ADDRESS: 220.143.161.174:61690
Binding Request
```

Figure 7: Facebook Messenger establishes a P2P connection

but in the experiment, the Wireshark filter syntax has not been found, It is speculated that LINE company should change the transmission rules, so it cannot be found in this experiment.

4.2.2 Facebook Messenger

Messenger uses the STUN protocol to establish packets for voice calls. At the beginning of the connection, it will use two packets to query Facebook DNS, the length of which is fixed at 82 and 122, and it will ask Facebook what the communication IP of the other end is. Because of the feature of the STUN protocol, P2P packet can often be found at the beginning of the call connection. The packet is the Binding Request and Binding Success Response sent by both end of the P2P connection, which is the IP of both parties that are required to establish a P2P connection as shown in Figure 7 Most communication software that uses the STUN protocol to send P2P packets will also send packets with arrays of similar information. In addition, when Messenger ends the call, it will send several ICMP packets and try to connect to the other port again. The appearance of these ICMP packets can determine that the call has ended. Besides, experiments have found that if the call is deliberately picked up late after the ringing (about 5 seconds later), it will affect the establishment of a P2P connection. Both of the communication will continue to send a set of Binding Request and Binding Success P2P packets at intervals of several seconds to dozens of seconds during the call, and the length is fixed at 162 and 106, but otherwise the packets during the call are still forwarded by the Messenger server.

4.2.3 WeChat

The connection server part of WeChat call, it is mainly responsible for the server with IP 203.205.220.0/23. The mechanism of WeChat establishment of the P2P connection is not to fix the call maintenance by both ends of the communication after a certain period of time, but perform in a specific mode: At the beginning of the connection in P2P mode, 7 to 8 packets with a fixed length of 162 will be sent, and the destination port of each packet is different; after that, it will continuously transmitted a group of 5 or less P2P packets during the call at a frequency of about two seconds, and the packet length is fixed at 142; also, the WeChat server is still responsible for most packet forwarding during the call. This is a significant feature of WeChat calls.

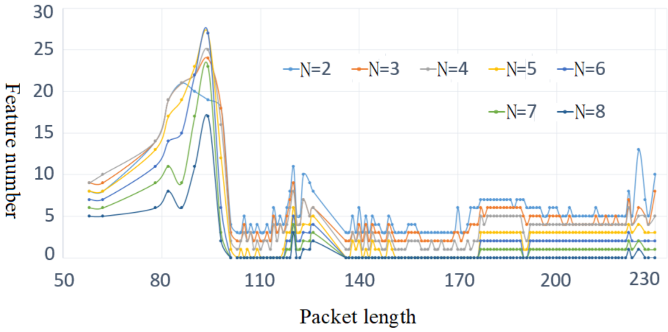


Figure 8: Relationship between packet length and feature number for different N value of N-gram

4.2.4 Skype

The calling protocol used by Skype is the STUN protocol. The packet for the P2P connection will be established during the ringing phase, and the packet length of the first P2P connection is usually 154; in addition, because Skype is operated by Microsoft, the experiment will see packets transmitted from Microsoft servers by nearby countries during a few seconds of ringing, such as 52.112.0.0/14 (Hong Kong), 40.64.0.0/10 (Seoul).

4.3 Relationship Between Packet Length and Feature Number

We know through experiments that the relationship between the VoIP packet length and the number of features for N-gram under different N values. As shown in Figure 8, the number of features between about 80 to 95 bytes in the packet length is the largest, and the average number of features is the largest. When N=2, the average number of features is the largest, but the trend of the number of features with a packet length close to 85 is different. We found that the MIM packet contains more consecutive 00s, which reduces the number of features. When N=8, most packets have no features. In addition, it is also found that in certain packet length intervals, the number of features is the same, such as 100 to 130, 140 to 170, and 180 to 210.

4.4 Performance Evaluation

4.4.1 Performance Metrics

In order to evaluate the performance of the proposed model, we used well-known metrics for classification, including Accuracy, Precision, Recall, and F1-score. In each packet evaluation, when the evaluation is VoIP, it is regarded as a positive class, otherwise it is regarded as a negative class. TP (True Positive) is a sample of VoIP packet, if the prediction is also a VoIP packet. TN (True Negative) means that the sample is not a VoIP packet and the prediction is not recognized as a VoIP packet. FP (False Positive) means that the sample is not a VoIP packet, but is predicted to be a VoIP packet. FN (False Negative) is a VoIP packet sample, but the prediction is not recognized as a VoIP packet. Therefore, Accuracy is the ratio of the number of correct classification to the total number of samples. As shown in Equation (5)

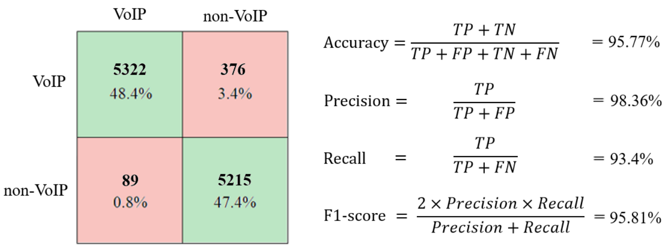
$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN}. \tag{5}$$

In addition, Precision refers to the ratio of the number of correctly classified instances to the total number of instances, as in Equation (6). Recall represents the percentage of positive examples correctly classified, as in Equation (7). F1-score is the harmonic average of Precision and Recall, as in Equation (8).

$$\begin{aligned} \text{Precision} &= \frac{TP}{TP+FP}, & (6)\\ \text{Recall} &= \frac{TP}{TP+FN}, & (7)\\ \text{F1-score} &= \frac{2 \times Precision \times Recall}{Precision + Recall}. & (8) \end{aligned}$$

4.4.2 Evaluation

The main purpose of an evaluation is to confirm the performance of the proposed model. The number of samples collected in this study is sufficient, so the experiment uses hold-out as a validation method, 50% of the data is selected randomly for training, and the remaining 50% is used for testing. That is, the testing data has a total of 11,002 packet samples, including 5,698 VoIP packet samples and 5,304 non-VoIP packet samples. As shown in Figure 9, it is the confusion matrix obtained by the test packet. The proposed model classifies 5,322 out of 5,698 VoIP packet samples as True (True Positive, TP), and 376 are classified as False (False Negative, FN). Among the 5,304 non-VoIP data packet samples, 5,215 are classified as True Negative (TN), and 89 are classified as False Positive (FP). We have obtained an Accuracy of 95.77%, Precision of 98.36%, Recall of 93.4%, and F1-score of 95.81%.



	VoIP	non-VoIP
VoIP	5322 48.4%	376 3.4%
non-VoIP	89 0.8%	5215 47.4%

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN} = 95.77\%$$

$$\text{Precision} = \frac{TP}{TP+FP} = 98.36\%$$

$$\text{Recall} = \frac{TP}{TP+FN} = 93.4\%$$

$$\text{F1-score} = \frac{2 \times Precision \times Recall}{Precision + Recall} = 95.81\%$$

Figure 9: Confusion matrix of the testing dataset, and performance of the proposed model

5 Conclusions

Through the packets captured during a call from each communication software, a stable call connection prompts the server to facilitate communication so as to establish a P2P connection. This feature reduces the server's burden. In the actual monitoring of a suspicious phone communication, the source or destination of the monitored phone's packets is a dynamic IP address provided by a telecommunications company and not by the server registered by the communication software's IP address. Moreover, a large number of UDP packets indicate that a given phone is currently in use. The mechanism for establishing a P2P connection varies with the IP address of the server. The process identifies which app is being used and how a P2P connection is established. Although the mechanism cannot obtain much information from a packet, it can identify connection features from the packet itself for observation. Nevertheless, the establishment of a P2P connection by communication software is beneficial for criminal investigations. Therefore, this research attempts to observe the packet features of P2P connections established using various communication softwares from various fields other than the IP information of each packet. An automated method that can effectively filter out VoIP packets with an N-gram model is also proposed to improve the efficiency of criminal investigations. The results of the experiment show that the proposed N-gram model obtains an Accuracy of 95.77%, Precision of 98.36%, Recall of 93.4%, and F1-score of 95.81%. Thereby assisting criminal investigators to quickly identify criminals and improve the efficiency of criminal investigations. In the future, more concepts of automatic filtering methods should be developed to solve the difficulties encountered by investigators. The efficiency of identifying suspicious callers on the basis of a large amount of packet content should also be improved.

Acknowledgments

This research was partially supported by the Executive Yuan of the Republic of China under the Grants Forward-looking Infrastructure Development Program (Digital Infrastructure-Information Security Project-110).

References

[1] S. Aiyar, N. P Shetty, "N-gram Assisted Youtube Spam Comment Detection," *Procedia Computer Science*, vol. 132, pp. 174-182, 2018.

[2] M. A. Azad, R. Morla, K. Salah, "Systems and methods for SPIT detection in VoIP: Survey and future directions," *Computers & Security*, vol. 77, pp. 1-20, Aug. 2018.

[3] J. Buford, H. Yu, E. K. Lua, *P2P Networking and Applications*, Morgan Kaufmann, 2008.

[4] Q. Y. Chen, "Wireless network signal monitoring based on LAN packet capture and protocol analysis on grid programming," *Computer Communications*, vol. 157, pp. 45-52, May 2020.

[5] T. Chakraborty, I. S. Misra, R. Prasad, *VoIP Technology: Applications and Challenges*, Springer, 2019.

[6] C. Chapman, *Chapter 7 - Using Wireshark and TCP Dump to Visualize Traffic*, Network Performance and Security, pp. 195-225, 2016.

[7] J. C. Chen, "Profile user activity through LINE encrypted traffic," *Communications of the CCISA*, vol. 23, no. 3, Jul. 2017.

[8] V. Habic, A. Semenov, E. L. Pasiliao, "Multitask deep learning for native language identification," *Knowledge-Based Systems*, vol. 209, Dec. 2020.

[9] IETF, *Session Traversal Utilities for NAT (STUN)*, Oct. 6, 2020. (https://datatracker.ietf.org/doc/rfc5389/)

[10] T. Islam, S. S. M. M. Rahman, Md. A. Hasan, A. S. Md. M. Rahaman, Md. I. Jabiullah, "Evaluation of N-gram based multi-layer approach to detect Malware in Android," *Procedia Computer Science*, vol. 171, pp. 1074-1082, 2020.

[11] Y. C. Lai, A. Ali, S. Hossain, Y. D. Lin, "Performance modeling and analysis of TCP and UDP flows over software defined networks," *Journal of Network and Computer Applications*, vol. 130, pp. 76-88, Mar. 2019.

[12] M. Lan, C. L. Tan, J. Su, Y. Lu, "Supervised and traditional term weighting methods for automatic text categorization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 4, pp. 721-735, 2009.

[13] P. Mehta, R. Shah, "A Survey of Network Based Traffic Classification Methods," *Database Systems Journal, Academy of Economic Studies, Bucharest, Romania*, vol. 7, no. 4, pp. 24-31, 2017.

[14] H. Oouch, T. Takenaga, H. Sugawara, M. Masugi, "Study on appropriate voice data length of IP packets for VoIP network adjustment," in *Global Telecommunications Conference (GLOBECOM'02)*, Nov. 2002.

[15] S. Rezaei, X. Liu, "Deep learning for encrypted traffic classification: An overview," *IEEE Communications Magazine*, vol. 57, pp.76-81, 2019.

[16] F. Sebastiani, "Machine learning in automated text categorization," *ACM Computing Surveys*, vol. 34, no. 1, pp. 1-47, 2002.

[17] L. F. Sikos, "Packet analysis for network forensics: A comprehensive survey," *Forensic Science International: Digital Investigation*, vol. 32, Mar. 2020.

[18] T. Stöber, M. Frank, J. Schmitt, I. Martinovic, "Who do you sync you are? Smartphone fingerprinting via application behavior," in *Proceedings of the Sixth ACM Conference on Security and Privacy in Wireless and Mobile Networks*, pp. 7-12, 2013.

[19] J. Tan, X. Chen, M. Du, "An internet traffic identification approach based on GA and PSO-SVM," *Journal of Computers*, vol. 7, no. 1, pp. 19-29, 2012.

[20] C. C. Tseng, C. L. Lin, L. H. Yen, J. Y. Liu, C. Y. Ho, "Can: A context-aware NAT traversal scheme," *Journal of Network and Computer Applications*, vol. 36, Issue 4, pp. 1164-1173, Jul. 2013.

[21] TWNIC, *Network Service Application Overview*, Oct. 10, 2020. (https://report.twnic.tw/2019/TrendAnalysis_globalCompetitiveness_en.html)

[22] D. Walnycky, I. Baggili, A. Marrington, F. Breitinger, J. Moore, "Network and device forensic analysis of Android Social-Messaging applications," *Digital Investigation*, vol. 14, p.77-84, 2015.

[23] C. C. Wu, K. T. Chen, Y. C. Chang, C. L. Lei, "Detecting VoIP traffic based on human conversation patterns, Principles, Systems and Applications of IP Telecommunications," in *Services and Security for Next Generation Networks*, pp. 280-295, 2008.

Biography

Cheng-Tan Tung Cheng-Tan Tung received his Master degree in computer science and information engineering from National Central University of Taiwan. He is currently an assistant professor of Department of Information Management, Central Police University, Taiwan. His research interests include Digital Forensics, Data Visualization, and Machine Learning.

Chih-Ping Yen is an Associate Professor, Department of Information Management, Central Police University. Received his Ph.D. degree from Department of Computer Science and Information Engineering, National Central University, Taiwan, in 2014. His research interest includes Digital Investigation, Artificial Intelligence & Pattern Recognition, Image Processing, and Management Information Systems.

Data Hiding Based on The Chinese Remainder Theorem

Chun-Cheng Wang[1], Tsung-Han Lin[2], and Wen-Chung Kuo[1]
(Corresponding author: Wen-Chung Kuo)

National Yunlin University of Science & Technology[1]
123 University Road, Section 3, Douliu City, Yunlin County, Taiwan (R.O.C.)
Email: simonkuo@yuntech.edu.tw
You-Shang Technical Corp[2]
Guangfu Rd., Fengshan Dist., Kaohsiung City, Taiwan (R.O.C.)

(Received Jan. 7, 2021; revised and accepted Sept. 21, 2021)
(Received Jan. 7, 2021; Revised and Accepted Sept. 21, 2021; First Online Oct. 19, 2021)
Special Issue on Multimedia Application on Information Hiding Forensics and Cybersecurity

Abstract

At the time of rapid change, the advancement of the Internet and the ceaseless development of technologies allowed for the frequent use of the Internet for communication and data transmission. Yet, these data can easily be replicated and transmitted, given their distinctive characteristics. Therefore, the secure transmission and receiving of data with peace of mind is an issue worth our second thoughts. The purpose of this article is to apply the modulus of the Chinese Remainder Theorem (CRT) to the LSB replacement method and the GEMD method to decompose the confidential data under the binary system and embed the decomposed confidential data into the designated pixels of the image. Then, use the LSB Replacement method or the GEMD hiding method in the carrier to form a hidden image. The receiver could then extract the designated pixel of the concealed image with the extraction of the LSB value or the GEMD data in the first place and compute the confidential data through the CRT with the same modulus value. The sorting of the modal value is the point of gravity and will affect the quantity for hiding and the PSNR value. Yet, this modulus is used to enhance the level of security.

Keywords: Chinese Remainder Theorem; Data Hiding; Generalized Exploiting Modification Direction (GEMD); Least Significant Bit (LSB) Replacement

1 Introduction

Data Hiding has long been used by people since the old days. This technique can be used to protect digital information, personal information, and account ID online or other documented hard copies of contracts from being stolen. This paper is focused on the security of information exchange via the Internet whereby confidential information can be embedded into media such as picture, voice, films, texts, and execution files, as shown in Figure 1.

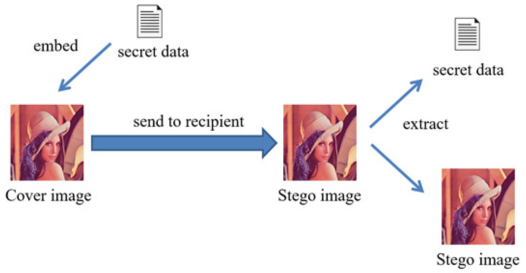


Figure 1: The process of using pictures to hide data

Data hiding aims at meeting the needs in 3 aspects:

Security: No one other than the legitimate participants can retrieve the confidential information hidden in the S image in disguise.

Imperceptibility: The quality of the information cannot vary significantly before and after hiding to avoid the awareness of a third party that the content has confidential information in hiding.

Payload: Increase the data payload as far as possible on condition that the quality of the image in disguise is unaffected.

In recent years, different data hiding methods have been proposed [1–10]. In this paper, the CRT is adopted due to its distinctive feature. Several combination prime numbers could be used to choose the modulus of a large numerical value, and this large number value will be referred to modulus calculation with several moduli individually to obtain the remainder, which could present a large numerical value with several moduli or to obtain a large value through reverse calculation. Sizable confidential information could be converted to the bit required for the binary system through the use of several products of moduli in order to split into several smaller numerical values, and apply the modulus to choose the appropriate quantity of pixel value, and hide the confidential information into the image with the LSB Replacement method or the GEMD method. Modulus value could be used to bring confusion to the confidential information to enhance its security, and to obtain the best combination of modulus value.

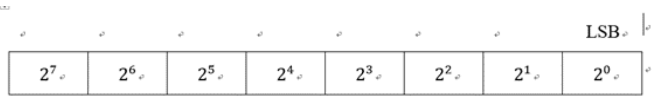
The primary framework of this paper is shown as follows: In Section II, the literature for reference of the data hiding technologies will be introduced. In Section III, data hiding under the CRT will be introduced. In Section IV, the empirical findings and analysis will be presented. In Section V, the conclusion of this study and the research direction in the future will be discussed.

2 Related Work

In this section, a review on the widely known LSB Replacement method and the GEMD technique proposed by scholars such as Kuo and others in 2009 on data hiding will be conducted. An important theorem: The Chinese Remainder Theorem will also be introduced.

2.1 LSB Replacement Method

LSB (Least Significant Bit) refers to the lowest bit of a binary number, which is shown in Figure 2.



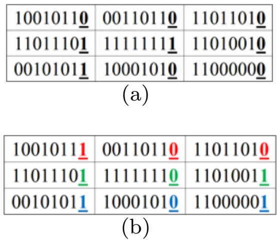
							LSB
2^7	2^6	2^5	2^4	2^3	2^2	2^1	2^0

Figure 2: The bit at the far right side is LSB

Under the LSB Replacement method [7], the bits of confidential information will be embedded into the image pixel of the carrier one-by-one. If the secret data is different from the LSB value of the image pixel of the carrier, use the confidential information to directly cover the LSB value of the pixel to get a disguised picture to achieve the objective of data hiding.

Example 1. *Assuming there is a 3×3 monochrome picture as a carrier with pixel value as shown in Figure 3(a), the confidential information $S = (100101101)_2$ is replaced bit by bit under the LSB Replacement method to the LSB of the image pixel of the carrier, and generate the Stego image as shown in Figure 3(b).*

The retrieving method is to take out the LSB value one-by-one from the pixel of the Stego image, and the LSB value is the secret data. The LSB Replacement method gives a very simple and quick method to hide confidential data into the carrier, and the image after hiding is not significantly different from the original image, which cannot be easily identify with the naked eye. Yet, the security level is low, and the confidential data could be accessed by retrieving the LSB value from the Stego image.



10010110	00110110	11011010
11011101	11111111	11010010
00101011	10001010	11000000

(a)

10010111	00110110	11011010
11011101	11111110	11010011
00101011	10001010	11000001

(b)

Figure 3: (a) Original pixel value; (b) Pixel value after embedding

2.2 Chinese Remainder Theorem (CRT)

The CRT provides the solution for the single variable of the prime numbers from prime number pairs as shown in Equation (1).

$$\left\{ \begin{array}{l} x \equiv a_1 \ (mod \ m_1) \\ x \equiv a_2 \ (mod \ m_2) \\ \quad \cdots \\ x \equiv a_n \ (mod \ m_n) \end{array} \right\} \tag{1}$$

Solution:

1) Find out $M = m_1 \times m_2 \times \cdots \times m_n$. This is the common modulus of all equations.

2) Find out $M_1 = M/m_1$, $M_2 = M/m_2, \cdots, M_n = M/m_n$.

3) Under the corresponding modulus $(m_1, m_2, \cdots, m_n)$, find out the multiplicative inverse elements of $M_1, M_2, \cdots, M_n$, which are $M_1^{-1}, M_2^{-1}, \cdots, M_n^{-1}$.

4) The common solution is shown as Equation (2).

$$x = (a_1 \times M_1 \times M_1^{-1} + \cdots + a_n \times M_n \times M_n^{-1}) \bmod M \tag{2}$$

Example 2. *Find x such that $x \equiv 2($ mod $3)$, $x \equiv 3($ mod $5)$ and $x \equiv 2(\text{mod}7)$.*
Solution:

1) Compute $M = 3 \times 5 \times 7 = 105$.

2) Caluculate $M_1 = 105/3 = 35$, $M_2 = 105/5 = 21$ and $M_3 = 105/7 = 15$.

3) Multiplicative inverse element is $M_1^{-1} = 2$, $M_2^{-1} = 1$ and $M_3^{-1} = 1$.

4) Use Equation (2), $x = (2 \times 35 \times 2 + 3 \times 21 \times 1 + 2 \times 15 \times 1) \bmod 105 = 23 \bmod 105$.

2.3 Generalized Exploiting Modification Direction

In 2013, Kuo and Wang proposed the Generalized Exploiting Modification Direction (GEMD) method [5] to enhance the embedding computing of EMD [10]. In the method, neighboring pixels from the original picture are taken as the cluster and $(n+1)$ bits of confidential data could be embedded into each cluster. The pixel value is revised on the basis of the difference between the value taken from the function and the confidential data. Each pixel value will be added, subtracted by 1 or keeping the same to embed the confidential data into the picture of the carrier to form a Stego image.
Extraction function:

$$f_b(g_1, g_2, \cdots, g_n) = \sum_{i=1}^{n}(g_i \times (2^i - 1)) \bmod 2^{n+1} \quad (3)$$

where g_i is the i-th pixel. The method of extracting confidential data is to equate the concealed pixel into the extraction function, and the confidential data can be obtained.

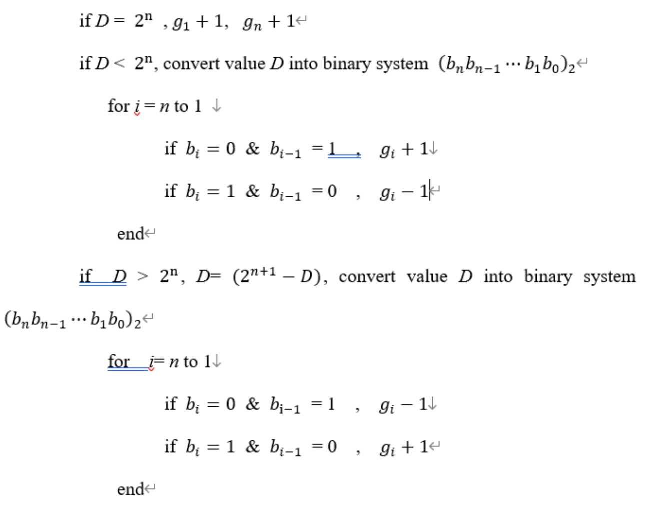
2.3.1 GEMD Hiding Algorithm

Step 1. Select n pixels and equate into the extraction function f_b to calculate the value.

Step 2. Select $(n+1)$ bits as confidential information S, and convert the secret data into decimal system.

Step 3. Find out the difference $D = (S - f_b) \bmod 2^{n+1}$.

Step 4. Adjust pixel on the basis of D as shown below:



```
if $D = 2^n$ , $g_1 + 1$, $g_n + 1$
if $D < 2^n$, convert value $D$ into binary system $(b_n b_{n-1} \cdots b_1 b_0)_2$
  for $i = n$ to 1
      if $b_i = 0$ & $b_{i-1} = 1$ ,  $g_i + 1$
      if $b_i = 1$ & $b_{i-1} = 0$ ,  $g_i - 1$
  end
if $D > 2^n$, $D = (2^{n+1} - D)$, convert value $D$ into binary system
$(b_n b_{n-1} \cdots b_1 b_0)_2$
  for $i = n$ to 1
      if $b_i = 0$ & $b_{i-1} = 1$ ,  $g_i - 1$
      if $b_i = 1$ & $b_{i-1} = 0$ ,  $g_i + 1$
  end
```

2.3.2 Example of Embedding with GEMD

If $n = 3$, and the value of the original pixel is $(179, 200, 191)$, while the secret data are $(1010)_2$, use the GEMD hiding algorithm to hide the data and get pixel value $(178, 200, 192)$. The steps are elaborated below:

Step 1. Calculate value f_b where $f_b(179, 200, 191) = (179 \times 1 + 200 \times 3 + 191 \times 7) \bmod 16 = 4$.

Step 2. Compute $(1010)_2 = 10$.

Step 3. Compute$D = (10 - 4) \bmod 16 = 6$.

Step 4. Because $D < 2^n$, modify $D = (b_3 b_2 b_1 b_0)_2 = (0110)_2$ and get stego pixels by using the following processing.

$b_3 = 0$, $b_2 = 1$, modify pixel value 191 into 192.

$b_2 = 1$, $b_1 = 1$, do not modify the value.

$b_1 = 1$, $b_0 = 0$, modify the pixel value of 179 into 178.

The stego pixels are (178,200,192).

2.3.3 Example of Extraction with GEMD

If the stego pixels are (178,200,192), then the confidential data extracted from the extraction function by using Equation (3), i.e., $f_b(178, 200, 192) = (178 \times 1 + 200 \times 3 + 192 \times 7) \bmod 16 = 10 = (1010)_2$.

3 Data Hiding Scheme Based on CRT

In this section, the CRT is adopted in combination with LSB Replacement metod and GEMD.

First of all, the sender chooses the value of several prime numbers as the modulus value at random, and uses the modulus value and the products of the modulus value to determine the quantity of secret data for hiding into a specific quantity of pixels. Both the LSB Replacement method and GEMD method are used to hide the data. The process is elaborated as below.

3.1 The Embedding Algorithm of CRT and LSB Replacement Method

Devide the confidential data in binary system into blocks of k bits in size. Each block of k bits is matched with each modulus value to find the remainder. We embed the remainder value into the $l_1 + l_2 + ... + l_n$ pixels of the cover image of the carrier under the LSB Replacement method.

The Embedding Algorithm

Input. Secret S (binary string), Cover image;

Output. Stego image, $m_1, m_2, \cdots, m_n$;

Step 1. Given pairwise coprime positive integers $m_1, m_2, \cdots, m_n$, and compute $M = m_1 \times m_2 \times \cdots \times m_n$.

Step 2. Find a maximum k such that $2^k \leq M$.

Step 3. Find the minimum l_i such that $2^{l_i} \geq m_i, (1 \leq i \leq n)$.

Step 4. Divide S into blocks where the embedding size of a block is k bits.

Step 5. Divide cover image into many blocks where the size of a block is $l_1 + l_2 + \cdots + l_n$ pixels.

Step 6. Convert each block of S, denoted as S_D, into a decimal number.

Step 7. If BitString = null, then do the following process.

```
for $i = 1$ to $n$ do
    $a_i \equiv S_D \bmod m_i$, where $a_i$ is represented as a binary number $BitString = BitString \,||\, a_i$
end for
```

Step 8. Using LSB replacement to embed BitString into $l_1 + l_2 + \cdots + l_n$ pixels of cover image.

Step 9. Repeat Step 7 until all blocks of S are embedded.

Here, we give an example to explain the proposed embedding algorithm, when $m_1 = 11$, $m_2 = 13$, and $m_3 = 17$ and the original pixels value shown as Figure 5.

Step 1. Calculate $M = 11 \times 13 \times 17 = 2431$.

Step 2. Find $k = 11$ such that $2^k \leq M$.

Step 3. Find $l_1 = 4$, $l_2 = 4$ and $l_3 = 5$ such that $2^{l_1} \geq m_1$, $2^{l_2} \geq m_2$ and $2^{l_3} \geq m_3$, respectively.

Step 4. Divide every 11 bits of confidential data into one block, i.e. $S = 10101110010 \mid 10111100101 \mid 10110000000 \cdots$;

Step 5. Divide every 13 pixels into one block, which is shown as Figure 4.

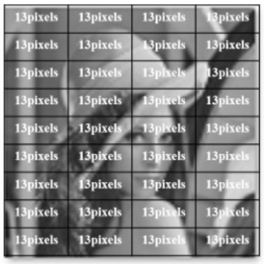


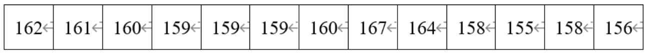
Figure 4: Divide the block results

Step 6. Compute $S_D = (10101110010)_2 = 1394$.

Step 7. Calculate $a_1 = 1394 \bmod 11 = 8 = (1000)_2$, $a_2 = 1394 \bmod 13 = 3 = (0011)_2$, $a_3 = 1394 \bmod 17 = 0 = (00000)_2$, and $BitString = 1000001100000$.

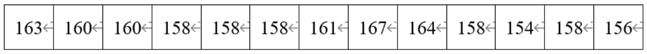
Step 8. There are 13 pixels, shown in Figure 5, the pixel value after the bit string was embedded by using the LSB Replacement method is shown in Figure 6.

Step 9. Repeat Step 7 until the confidential data of each block are embedded.



162	161	160	159	159	159	160	167	164	158	155	158	156

Figure 5: The original pixel value



163	160	160	158	158	158	161	167	164	158	154	158	156

Figure 6: The embedded pixel value

3.2 The Extraction Algorithm of CRT and LSB Replacement Method

Take the LSB value of l_i pixel from every $l_1 + l_2 + ... + l_n$ pixels of the stego image as the remainder value, and calculate with the use of CRT and we can get the confidential data by Extraction algorithm.

The Extraction Algorithm

Input: Stego image, $m_1, m_2, \cdots, m_n$;

Output: Secret S (binary string).

Step 1. Find minimum l_i such that $2^{l_i} \geq m_i$, $(1 \leq i \leq n)$.

Step 2. Divide stego image into blocks where the size of a block is $l_1 + l_2 + \cdots + l_n$ pixels.

Step 3. Extract the LSBs of l_i pixels, denoted as a_i, where is represented as a decimal number.

Step 4. Compute $S = (a_1 M_1 y_1 + a_2 M_2 y_2 + \cdots + a_n M_n y_n) \bmod M$ and then convert into binary.

Step 5. Repeat Step 3 until all S are extracted.

Here, we can extract the secret data from Figure 6 with $m_1 = 11$, $m_2 = 14$, and $m_3 = 17$.

Step 1. Find l_1, l_2, and l_3 such that $2^{l_1} \geq m_1$, $2^{l_2} \geq m_2$ and $2^{l_3} \geq m_3$, respectively.

Step 2. Calculate $l_1 + l_2 + l_3 = 13$, cut every 13 pixels of the stego image into one block.

Step 3. Extract l_i LSB value from the stego image, as shown in Figure 7. Therefore, $a_1 = (1000)_2 = 8$, $a_2 = (0011)_2 = 3$, and $a_3 = (00000)_2 = 0$.

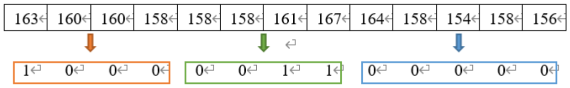


Figure 7: Extract LSB value

Step 4. Because $x \equiv 8(\mathrm{mod} 11)$, $x \equiv 3(\mathrm{mod} 13)$ and $x \equiv 0(\bmod 17)$, we can compute $M = 11 \times 13 \times 17 = 2431$ and then find $M_1 = 2431/11 = 221$, $M_2 = 2431/13 = 187$ and $M_3 = 2431/17 = 143$. Therefore, we can get $y_1 = 1$, $y_2 = 8$ and $y_3 = 5$. Finally, it can find the common $x = (8 \times 221 \times 1 + 3 \times 187 \times 8 + 0 \times 143 \times 5) \bmod 2431 = 1394 = (10101110010)_2$. In other words, $(10101110010)_2$ is the extracted confidential data.

3.3 The Hidding Algorithm of CRT and GEMD Method

Cut the confidential data in binary system into blocks of k bits in size, and calculate the remainder value by matching each k bits with each modulus value, and embed the remainder value into $l_1 + l_2 + ... + l_n - n$ pixels of the cover image by using the following algorithm.

The Embedding Algorithm of CRT and GEMD Method

Input: Secret S (binary string), Cover image;

Output: Stego image, $m_1, m_2, ..., m_n$.

Step 1. Given pairwise coprime positive integers $m_1, m_2, \cdots, m_n$, and $M = m_1 \times m_2 \times \cdots \times m_n$.

Step 2. Find a maximum k such that $2^k \leq M$.

Step 3. Find a minimum l_i such that $2^{l_i} \geq m_i$, $(1 \leq i \leq n)$.

Step 4. Divide S into blocks where the embedding size of one block is k bits.

Step 5. Divide cover image into blocks where the number of pixels in one block is $l_1 + l_2 + ... + l_n - n$ pixels.

Step 6. Convert each block of S, denoted as S_D, into a decimal number.

Step 7. Do the following process:

```
for  i = 1 to n do
    $a_i \equiv S_D \bmod m_i$, using GEMD to embed $a_i$ into
    $l_i - 1$ pixels of cover image
end for
```

Step 8. Repeat Step 7 until all blocks of S are processed.

Now, we give an example to explain our proposed method. It assumes that the cover pixels are shown as Figure 8, $m_1 = 11$, $m_2 = 13$, and $m_3 = 17$. By using the embedding algorithm of CRT and GEMD method.

Step 1. Calculate $M = 11 \times 13 \times 17 = 2431$.

Step 2. Find $k = 11$ such that $2^k \leq M$.

Step 3. Find $l_1 = 4$, $l_2 = 4$, and $l_3 = 5$ such that 2^{l_1}, $2^{l_2} \geq m_2$, and $2^{l_3} \geq m_3$.

Step 4. Cut every 11 bits of the confidential data into one block when $S = 10101110010 \mid 10111100101 \mid 10110000000....$

Step 5. Cut every $10(4 + 4 + 5 - 3)$ pixels in one block.

Step 6. Compute $S_D = (10101110010)_2 = 1394$.

Step 7. Compute there $a_1 = 1394 \bmod 11 = 8$ and find $l_1 = 4$ such that $2^{l_1} \geq 8$. Therefore, we take 3 pixels(162,161,160) from Figure 8 and embed with a_1 under GEMD method.

Setp 8. Calculate the difference $D = (8-5) \bmod 16 = 3$ and $f_b(162, 161, 160) = (162 \times 1 + 161 \times 3 + 160 \times 7) \bmod 16 = 5$.

$\because D < 2^3, d = (b_3 b_2 b_1 b_0)_2 = (0011)_2$.

$\because b_3 = 0,\ b_2 = 0$, do not modify the value.

$\because b_2 = 0,\ b_1 = 1$, modify pixel value of 161 into 162.

$\because b_1 = 1,\ b_0 = 1$, do not modify the value.

Therefore, we can get the stego pixels (162,162,160). Repeat using the embedding algorithm to get stego pixels (160,158,159,159,151,157,153) from (159,159,159,159,160,167,164). So, we can get the stego pixels shown as Figure 9 from the original pixels shown as Figure 8.

The Extraction Algorithm of CRT and GEMD Method

Input: stego image, $m_1, m_2, ..., m_n$;

Output: Secret S (binary string).

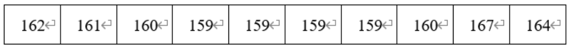
Step 1. Find a minimum l_i such that $2^{l_i} \geq m_i$, $(1 \leq i \leq n)$.

Step 2. Divide stego image into blocks where a block includes $l_1 + l_2 + \cdots + l_n - n$ pixels.

Step 3. Using GEMD Extraction function to compute of $l_i - 1$ pixels, denoted as a_i, where is represented as a decimal number.

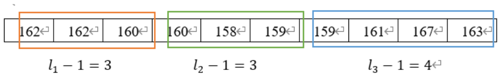
Step 4. Compute $S = (a_1 M_1 y_1 + a_2 M_2 y_2 + ... + a_n M_n y_n) \bmod M$ and then convert it into binary.

Step 5. Repeat Step 3 until all S are extracted.



162↵	161↵	160↵	159↵	159↵	159↵	159↵	160↵	167↵	164↵

Figure 8: The original pixels



162↵	162↵	160↵	160↵	158↵	159↵	159↵	161↵	167↵	163↵

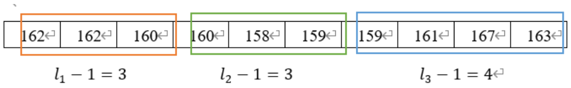
Figure 9: The stego pixels

Here, we also give to stego pixels,shown as Figure 9 when $m_1 = 11$, $m_2 = 13$, and $m_3 = 17$.

Step 1. Compute $l_1 = 4$, $l_2 = 4$ and $l_3 = 5$, such that $2^{l_1} \geq 11$, $2^{l_2} \geq 13$, $2^{l_3} \geq 17$.

Step 2. Compute $l_1 + l_2 + l_3 - n = 10$, split every 10 pixels of the stego image into one block.

Step 3. Extract $l_i - 1$ pixels from the stego image, as shown in Figure 10, and use GEMD extraction function to calculate and get a_i for $i = 1$, 2 and 3, such that $f_b(162, 162, 160) = (162 \times 1 + 162 \times 3 + 160 \times 7) \mod 16 = 8 = a_1$, $f_b(160, 158, 159) = (160 \times 1 + 158 \times 3 + 159 \times 7) \mod 16 = 3 = a_2$ and $f_b(159, 161, 167, 163) = (159 \times 1 + 161 \times 3 + 167 \times 7 + 163 \times 15) \mod 32 = 0 = a_3$.



162↵	162↵	160↵	160↵	158↵	159↵	159↵	161↵	167↵	163↵

Figure 10: Extract $l_i - 1$ pixel value

Step 4. Find out $x = 1394 = (10101110010)_2$, by using CRT method. In other words, $(10101110010)_2$ is the secret data.

4 Experimental Results and Analysis

Security, image quality, and hidden quantity should be taken into account for the data hiding technology. The focus of this paper is to propose the enhancement of security with image quality at a specific standard. Image quality will be measured by PSNR (Peak Signal to Noise Ratio). The higher the value, the higher the standard of image quality.

PSNR method:

$$PSNR = 10 \times log_{10}(\frac{255^2}{MSE})(dB) \qquad (4)$$

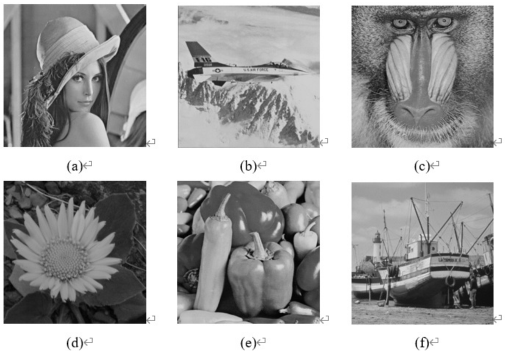
Where, MSE (Mean Square Error) will be calculated by the equation below:

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [S(i,j) - C(i,j)]^2 \qquad (5)$$

m, n stand for the size of the image, and $S(i,j)$ represents specific pixel value of the Stego image, and $C(i,j)$ represents specific pixel value of the image of the carrier.

4.1 Experimental Results

PC is adopted as the simulation environment as the method for this paper, matching with Intel (R) Core ™ i7-7700 processor and 16G memory under Windows 10 as the OS. Matlab is adopted as the program language. Six witty pictures of size 512 x 512 were used in the experiment, which is shown in Figure 11.



(a)↵ (b)↵ (c)↵ (d)↵ (e)↵ (f)↵

Figure 11: Pictures in the experiment

The following is the simulation result with the combined use of GEMD and CRT on the basis of the above condition in the experiment. There are 167 combinations of prime numbers falling between 3 and 1000 not being selected. Prime numbers marginally smaller than 2^k times are better combinations to select, such as: 29, 31, 61, 113, 127, 239, 241, 251, $\cdots$, and there are 34 combinations within 1000. These two combinations will be considered for random selection of 3, 4, 5, 6, 7, 8 and each will be run 100 times, and ranked in descending order by hiding quantity, as shown in Figure 12.

4.2 Analysis

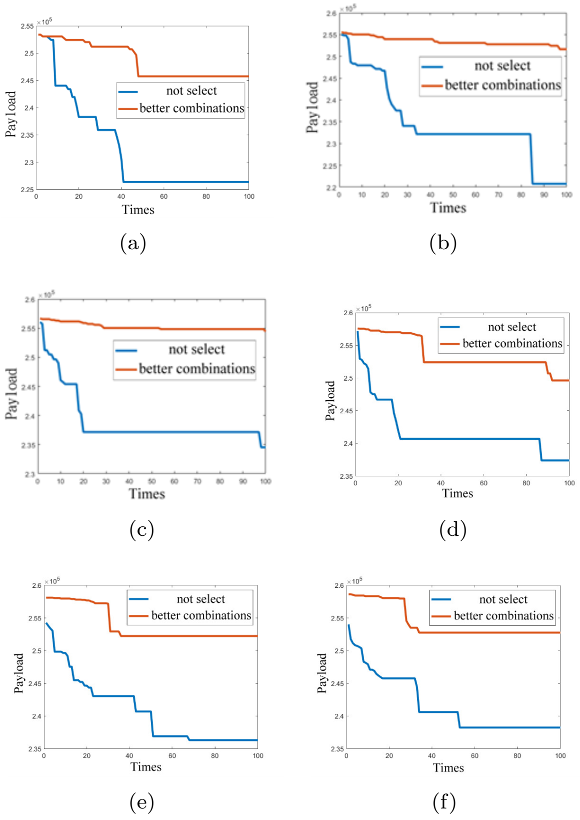
The method proposed by scholars Kuo and others is modified from EMD [10], which helps to improve the hiding quantity. The whole picture is used for splitting up and hiding with the same group of pixels. The method proposed in this paper is based on the modulus value to determine the split up of pixels. Accurate data cannot be obtained due to the different sequential orders of the modulus values. This paper is an attempt to use the method of using different image groups after splitting up in the whole picture to enhance security. Table 1 is the comparison of the method in this paper.

5 Conclusions

The findings from the experiment indicated that the combined use of the CRT and LSB Replacement method or the GEMD could keep very good image quality from the

Table 1: Comparison table

	LSB Replacement	Combining LSB Replacement Method and CRT	GEMD	Cobing GEMD and CRT
Pixel Selected	Constant	Modulus value and its sequence	Put every n into 1 group	Modulus value and its sequence
bpp	1	By value M	$\frac{n+1}{n}$	By value M
Security	No	Yes	No	Yes



(a) (b) (c) (d) (e) (f)

Figure 12: (a)n=3 (b)n=4 (c)n=5 (d)n=6 (e)n=7 (f)n=8 The simulation result of running 100 times with different modulus value by hiding quantity in sequential order

perspective of PSNR. As for the hiding quantity, selection will be made on the basis of the modulus value, which will give different results. The pixel for embedding under the LSB Replacement method and GEMD is constant and is not secure and could easily be guessed. In this paper, pixel groups were selected by modulus value. Different sequencing ways of modulus values give different results, which help to enhance security.

References

[1] K. Ahmadi, and E. Salari, "An image hiding algorithm using Chinese remainder theorem with reduced distortion," in *IEEE International Conference on Electro/Information Technology*, pp. 240–245, June 2014.

[2] D. Bhat, V. Krithi, K. N. Manjunath, S. Prabhu, and A. Renuka, "Information hiding through dynamic text steganography and cryptography: Computing and Informatics," in *International Conference on Advances in Computing, Communications and Informatics (ICACCI'17)*, pp. 1826–1831, Sep. 2017.

[3] C. C. Chang, N. T. Huynh, and H. D. Le, "Lossless and unlimited multi-image sharing based on Chinese remainder theorem and Lagrange interpolation," *Signal Processing*, vol. 99, pp. 159–170, Jun. 2014.

[4] C. C. Chang, T. C. Lu, G. Horng, Y. H. Huang, and Y. M. Hsu, "A high payload data embedding scheme using dual Stego-images with reversibility," in *9th International Conference on Information, Communications & Signal Processing* pp. 1–5, Dec. 2013.

[5] W. C. Kuo, and C. C. Wang, "Data hiding based on generalized exploiting modification direction method," *Imaging Science Journal*, vol. 61, no. 6, pp. 484-490, 2013.

[6] W. C. Kuo, R. J. Xiao, C. C. Wang, Y. C. Huang, "Study on security enhancing of generalized exploiting modification directions in data hiding," in *International Conference on Security with Intelligent Computing and Big-data Services (SICBS'18)*, pp. 103-115, 2018.

[7] Y. Liu, and J. Zhang, "Large–capacity LSB information hiding scheme based on two–dimensional code," in *7th IEEE International Conference on Electronics Information and Emergency Communication (ICEIEC'17)*, pp. 528–532, July 2017.

[8] K. Parasuraman, and D. G, "Reversible image watermarking using interpolation technique," in *International Conference on Electronics, Communication and Computational Engineering (ICECCE'14)*, pp. 200–205, 2014.

[9] F. Wang, Y. Guo, Z. Yin, X. Zhou, and Q. Zhao, "Data hiding method based on reference matrix," *Procedia Computer Science*, vol. 131, pp. 800–809, Jan. 2018.

[10] X. Zhang, and S. Wang, "Efficient steganographic embedding by exploiting modification direction," *IEEE Communications Letters*, vol. 10, no. 11, pp. 781–783, Nov. 2006.

Biography

Chun-Cheng Wang biography. Received the Ph.D. degree from National Yunlin University of Science and Technology, Republic of China, in 2017. He is currently a postdoctoral researcher in National Yunlin University of Science and Technology, Republic of China. His research interests are cryptography, image processing and information hiding.

Tsung-Han Lin biography. Received the M.S. degree from National Yunlin University of Science and Technology, Republic of China, in 2020. He is currently an engineer in You-Shang Technical Corp., Republic of China. His current job content is software development and maintenance.

Wen-Chung Kuo biography. Received the B.S. degree in Electrical Engineering from National Cheng Kung University and M.S. degree in Electrical Engineering from National Sun Yat-Sen University in 1990 and 1992, respectively. Then, He received the Ph.D. degree from National Cheng Kung University in 1996. Now, he is a professor in the Department of Computer Science and Information Engineering at National Yunlin University of Science & Technology. His research interests include steganography, cryptography, network security and signal processing.

Retrieving Potential Cybersecurity Information from Hacker Forums

Chia-Mei Chen[1], Dan-Wei Wen[2], Ya-Hui Ou[3], Wei-Chih Chao[1], and Zheng-Xun Cai[1]
(Corresponding author: Chia-Mei Chen)

Department of Information Management, National Sun Yat-sen University[1]
No. 70, Lianhai Rd, Gushan District, Kaohsiung 804, Taiwan
(Email: cchen@mail.nsysu.edu.tw)
Department of Management Sciences, Tamkang University, Taipei, Taiwan[2]
National Penghu University of Science and Technology, Taiwan[3]

(Received Mar. 6, 2021; Revised and Accepted Sept. 21, 2021; First Online Oct. 19, 2021)
Special Issue on Multimedia Application on Information Hiding Forensics and Cybersecurity

Abstract

To adapt to the rapidly evolving cyberattacks, cyber threat knowledge is essential for organizations to gain visibility into the fast-evolving threat landscape and timely identify early signs of an attack and the adversary's strategies, tactics, and techniques. In addition, to gaining insight into potential cyber threats, hacker forums are a valuable source. However, the complexity and diversity of the content in hacker forums make it challenging to retrieve useful cybersecurity information. This research proposes an improved data preprocessing method to reduce feature dimension and a hybrid method combining text tagging and clustering analysis techniques to discover cybersecurity information from unstructured hacker forums. The experimental results illustrate that the proposed solution could extract cybersecurity information efficiently.

Keywords: Cyber Threat Intelligence, Hacker Forum, Latent Dirichlet allocation, Natural Language Processing

1 Introduction

Organizations and businesses apply modern information technologies to expand services and improve customer satisfaction, while in the meantime they are facing potential cyberattacks. Cyberattacks have increased in frequency and sophistication, presenting significant challenges for organizations that must defend their data and systems from capable threat attackers. They utilize a variety of tactics, techniques, and procedures (TTPs) to compromise systems, disrupt services, commit financial fraud, and expose or steal intellectual property and other sensitive information. Given the risks these threats present, organizations seek solutions to improve information security and reduce cyberattack risks.

According to a guide to cyber threat information sharing published by the National Institute of Standards and Technology (NIST) [16], cyber threat information or cyber threat intelligence (CTI) is any information that can help an organization identify, assess, monitor, and respond to cyber threats. Cyber threat information includes indicators of compromise (IoC); tactics, techniques, and procedures used by threat actors; suggested actions to detect, contain, or prevent attacks; and the findings from the analyses of incidents. Organizations can improve their security postures in case such cybersecurity information is acquired.

Collecting such cybersecurity information is an important investment for organizations as it provides a proactive measure to prevent security breaches and saves financial losses. To obtain CTI, security teams gather unstructured data from multiple sources and analyze it to retrieve useful CTI about adversaries and attack signatures to make security decisions for organizations. The purpose of such CTI collection and discovery is to keep organizations informed of the potential threats and exploits.

Hacker forums are a popular internet community for hackers sharing hacking knowledge such as security breaches, hacking tools, malware, evasion techniques, and data leakage. For example, hackers discussed attack plans in the forums [44]; 7.5 million customer personal information was leaked from an online financial service company and sold in hacker forums [7]; a data breach broker sold databases of user records from 14 companies [37]; some forums offer hackers hiring, penetration test, and remote access services [29].

Hacker forums are a valuable source of cybersecurity intelligence [15]. Due to the massive volume of forum posts, extracting cybersecurity-related information from hacker forums is important to discover potential threats and security trends. Therefore, this study extracts information from hacker forums to discover vital cyber threat information to facilitate prompt response to cyberattacks.

Classification is a supervised learning approach that learns to figure out what class a new object should fit in by learning from training data with the class labels; clustering is an unsupervised learning approach that groups similar objects without knowing what their labels are. Classification uses predefined classes in which objects are assigned, while clustering identifies similarities between objects, which it groups according to those features in common and which differentiate objects from other groups. Therefore, classification could be used to detect patterns such as IoC (Indicator of Compromise) patterns, malicious URLs, domain names, etc.; clustering could be used to explore forum content and discover new information discussed in the forums. To identify threat intelligence, most literature applied either classification or cluster models [1, 2, 9]. This study combines the two approaches, text tagging and clustering, to explore the content of hacker forums and to discover the CTI information.

One key challenge of clustering is how to determine the number of clusters, as it depends on the level of granularity and analysis goals. This study compares different clustering models with various clustering evaluation measures including the elbow method, Silhouette Coefficient, Calinski-Harabaz Index, and Davies-Bouldin Index to find a valid approach to determine the number of clusters and discover CTI in hacker forums.

Hacker forums are supposed to discuss and share hacking-related subjects, while users may post freestyle or random information. Such posts would make analysis and extraction complicated. To propose an effective CTI extraction method for hacker forums, this study improves the traditional data cleaning method and reduces the feature dimension greatly. The posts in hacker forums contain diverse technical as well as non-technical related information. Therefore, the study proposes a novel analysis method that adopts two-stage clustering to identify new threat information, where the first stage clustering groups the content by theme topics and the second stage focuses on dividing into security-related event clusters.

2 Research Gaps And Questions

Several research gaps were identified from the literature review. First, current CTI efforts rely on the use of autofeeds from security vendors to generate threat intelligence. This means current security measures are often handled reactively based on existing attack cases. Second, hacker forums contain diverse non-security related information and free-style writing forms, which require effective data cleaning and clustering to extract security-relevant information. Finally, previous work focused on identifying security information by classification with patterns and rarely explored forum content to discover potential threat intelligence by clustering. With these research gaps, the following research questions have been proposed to guide the study:

- How to pre-process forum posts effectively to extract meaningful content?
- How to validate the effectiveness of the clustering results?
- How to explore hacker forums and extract proactive CTI efficiently by clustering?

The primary contribution of this study is to discover potential cybersecurity information by exploring hacker forums as a source of cyber threat intelligence and by applying a hybrid method of text tagging and clustering. This is achieved by using an automated process that consists of the following main phases: (1) data collection, (2) data cleaning and tagging, and (3) two-stage clustering of discovering topics pertaining to cybersecurity.

3 Literature Review

From the perspective of data collection, data can be divided into two categories: indicator-based and document-based. The first is indicator-based data feeds (Indicator Feeds). Indicator-based data feeds mainly share indicators of compromise (IoC) to achieve attack prevention in a short time, including the blacklist IP address, malicious domains, and malware hashes. The document-based data may contain rich and comprehensive threat information than the former one, which requires to apply NLP techniques and analysis models to retrieve them.

Tagging is efficient in extracting indicator-based CTI information as well as semantic information from unstructured corpus. Wollschlaeger *et al.* [43] proposed a semantic annotation framework based on tagging, where the tags address several independent aspects of semantics, increasing the expressiveness of information semantics. Wang and Chow [44] performed semantic extraction by tagging unstructured CTI data, and the experiment results show that the extracted entities and relationships by tagging provide valuable CTI information. Chen *et al.* [5] utilized tagging for capturing the semantics of web services in order to improve clustering performance.

The term frequency-inverse document frequency (TF-IDF) is a numerical statistic that reflects the importance of a word to a document in a collection of documents or corpus, where TF refers to the total number of times a given word appears in a document against the total number of all words in the document and IDF measures how common or rare a given word is across all documents. The TF-IDF can be expressed in the following equation.

$$tfidf(t,d) = tf(t,d) \times idf(t) \tag{1}$$

where t is a token or a given word and d is the document. The TF-IDF value increases in proportion to the number of times a given word appears in the document but is offset by the frequency of the word in the corpus to adjust the factor of words that frequently appeared.

Niakanlahiji *et al.* [4] employed a context-free grammar (CFG) model to extract candidate threat actions and applied TF-IDF to extract threat actions. Their results imply that TF-IDF is suitable for representing the importance of a candidate threat action among a list of tokens, so this study adopts it for extracting relevant short phrases from candidate threat actions.

Distributed representations of words in a vector space help learning algorithms to achieve better performance in NLP tasks by grouping similar words. Word2Vec (W2V) [27] is a family of word embedding (word vector) models of representing distributed representations of words in a corpus, where Continuous Bag-of-Words Model (CBOW) and Continuous Skip-gram Model are commonly used. It is a two-layer neural network and produces a vector space, where each unique word in a corpus is assigned a corresponding vector in the space.

A study [40] concluded that Word2Vec outperforms the traditional feature selection models including CHI, IG, and DF. As words may have different meanings (i.e., senses) depending on the context, identifying words in the correct meaning is important for extracting relevant information. Two previous studies [14, 31] concluded that Word2Vec can capture syntactic word similarities effectively and outperforms LSA (Latent semantic analysis) used commonly in word sense disambiguation.

Word2Vec models lose the ordering of the words. An unsupervised algorithm Doc2Vec (D2V) [22] represents each document by a dense vector, which overcomes the weaknesses of Word2Vec. Kadoguchi *et al.* [17] applied Doc2Vec and ML technology to classify information security data from dark web forums, and the results indicate that Doc2Vec is effective on feature selection and a multi-layer classifier can achieve 79% accuracy. Another study [34] applied Doc2Vec on classifying court cases and yields 80% accuracy. A performance study [34] demonstrated that Word2Vec and Doc2Vec perform better than N-gram on text classification and semantic similarity.

The above word embeddings are pre-trained models from co-occurrence statistics, while pre-trained contextual language models, BERT (Bidirectional Encoder Representations from Transformers) [10], generate word embeddings by jointly conditioning on left and right context. BERT-based models have been applied for search queries and classifications. Some studies [6, 30, 32] applied BERT for ranking query and document pairs and constructing a search query model, and some [12, 26, 41, 45] utilized BERT-based transformers to detect fake news.

Zhan *et al.* [46] conducted a performance analysis of BERT model and found out that BERT dumps redundant attention weights on tokens with high document frequency, such as periods, and that may lead to a potential threat to the model robustness. BERT extracts representations for query and document in the beginning and relies heavily on the interactions to predict relevance. The authors suggested some improvement may transform it into a more efficient ranking model. Khattab and Zaharia [18] developed an improved BERT-based ranking model that independently encodes the query and the document by delaying interactions. According to the literature review, it might not be suitable for exploring cyber threat information from unlabeled corpus like hacker forums.

Liao *et al.* [25] presented an automatic IoC extraction method based on the observation that the IoCs are described in a predictable way: being connected to a set of terms like “download”. It generated 900K IoC items with a precision of 95% and a coverage of over 90%. Kurogome *et al.* [21] proposed an automatic malware signature generation system from given malware samples, and the evaluation demonstrated that the produced IOCs are as interpretable as manually-generated ones.

Samtani *et al.* [36] applied classification and topic modeling techniques to extract source code from manually categorized data, where LDA (Latent Dirichlet allocation) finds the topics of the source code postings and classification categorized the programming language type. Benjamin and Chen [1] utilized recurrent neural network language models (RNNLMs) coupled with methodology from lexical semantics for learning hacker language. They demonstrated that RNNLMs can be used to develop the capability for understanding hacker language and different embedding models may impact the performance of the machine learning model.

Underground forums allow criminals to interact, exchange knowledge, and trade in products and services. Pastrana *et al.* [33] developed a web crawler to capture data from underground forums. Biswas *et al.* [2] applied a logistic regression model and sentiment analysis to achieve role-based hacker classification and examine hacker behaviors in dark forums. The overall classification accuracy is 80.57 %, and the keywords used in message posts are greatly linked to hacker expertise. Gautam *et al.* [11] employed machine learning approaches to classify underground hacker forum data into predefined categories, and the experimental results show that RNN GRU outperforms LSTM and yields the classification results of 99.025% accuracy and 96.56% precision.

Deliu *et al.* [9] explored the potential of Machine Learning (ML) methods to retrieve relevant threat information from hacker forums and compared the text classification performance of a Convolutional Neural Network (CNN) model against a traditional ML approach (Support Vector Machines). They concluded that SVM performs equally well as CNN.

Li *et al.* [23] combined Word2Vec and LDA to cluster academic abstracts and concluded that the combined model clusters the abstracts efficiently. Another study [38] also combined Word2Vec and LDA for web service clustering and demonstrated that the combined model outperforms a plain LDA.

The previous work demonstrated that hacker forums contain valuable CTI and mostly focused on applying classification models for extracting CTI from hacker forums. Traditional ML models can yield high levels of performance that are on par with modern ML models.

4 Methodology

This study developed a CTI discovery method as plotted in Figure 1 to answer the proposed research questions, and the notations used in this study are summarized in Table 1. The proposed method consists of the following components: data collection, data cleaning and tagging, word embedding, and CTI analysis and extraction. This study applies text tagging and word embedding to extract semantic information and develops a two-stage clustering method to retrieve security-related information. According to the literature review, word embedding models could represent semantic information [34], and the studies [5,44] demonstrated tagging could extract useful semantic information and improve clustering performance.

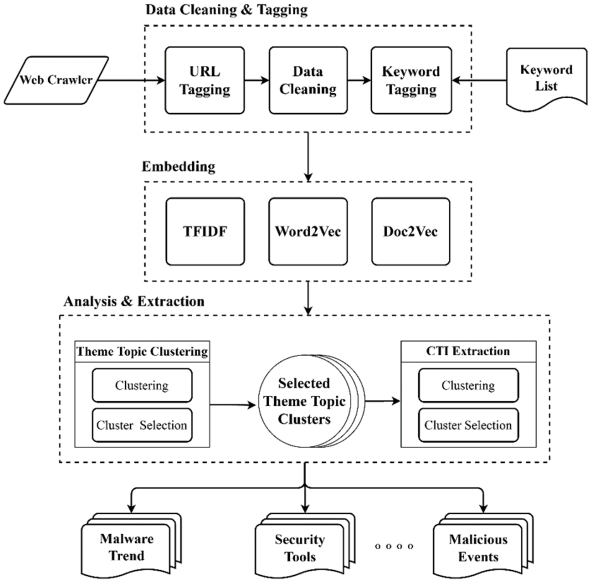


Figure 1: Research design

Table 1: Notations used in this study

Notation	Meaning
$\|A\|$	The number of elements in a set A
Corpus	The set of the documents in a corpus
C_{att}	The set of all the theme topic clusters
C_{stt}	The set of the selected theme topic clusters
C_{unfit}	The set of the theme topic clusters in extreme sizes
E_i	The set of the event clusters in the selected theme topic cluster i
TC_i	The theme topic cluster i
W_j	The j-th keyword of a cluster
$W_{TC\_i\_j}$	The j-th keyword in the theme topic cluster i
$S(W)$	The TFIDF score of a keyword W
R_{max}	The maximum ratio of a theme topic cluster to the corpus
R_{min}	The minimum ratio of a theme topic cluster to the corpus
$D(W_j)$	$(S(W_j) - S(W_{j+1}))/S(W_{j+1})$; the discrepancy of the j-th keyword to $j+1$-th's
H_{dis}	The discrepancy threshold of two consecutive keywords

4.1 Data Collection

Data collection can be achieved by developing a web crawler to gather posts from hacker forums. Some hacker forums employ anti-crawling techniques to hinder automated content extraction, which complicates collection automation. Forum posts may contain various data forms such as text, image, attachment, and threads of responses.

4.2 Data Cleaning and Tagging

The process of data cleaning and tagging reduces the volume of the corpus as well as the dimension of token vectors. This process consists of the following submodules: URL labeling, data cleaning, and keyword tagging, where data cleaning includes tokenization, stop word removal, token pruning, and tagging.

Common data preprocessing in text mining removes URL labels directly before proceeding with the rest of the data preprocessing steps. Li's study [24] observed that sellers might express the privacy information to be sold in a URL-like text format to catch the reader's attention. To retain such information, the proposed method performs URL labeling/tagging before data cleaning, as the text preprocessing steps might remove or disrupt it.

Users have different writing styles so that the documents often contain different terms with similar meanings. In text mining, a large keyword list (feature set) complicates the analysis and induces bias. Therefore, this study applies text tagging to reduce the feature dimension and to improve the information retrieval performance, while retaining the semantic information. Text tagging is achieved by keyword and regular expression matching in this study. The keyword tagging could achieve the purposes of token pruning and feature dimension reduction. The selected keywords are based on the previous studies [13, 19, 20] and categorized into two types: security and non-security relevant.

The tagged documents contain hashtag tokens in the format of #keyword#, where a matched term or regular expression is replaced by the associated hashtag. Based on our preliminary study on observing posts in hacker forums, tthis study defines 18 hashtags: 7 non-security hashtags (NH) and 11 security hashtags (SH). The non-security hashtags include #HIDDEN#, #IMAGE#, #ATTACHMENT#, #URL#, #QUOTE#, #MODERATOR#, and #PORN#; the security hashtags include #ICQ#, #ACC_PASS#, #E-MAIL#, #WEBSITE#, #EXPLOIT#, #ATTACK#, #MALWARE#, #PROXY#, #PAYMENT#, #TUTORIAL#, #AN-

TIVIRUS#.

The proposed data cleaning process consists of lemmatization and tokenization, stop word removal, irrelevant terms removal by rules. Lemmatization and tokenization divides text information into individual words, where this study deploys word tokenization from Python NLTK as an analysis [28] on open source tools showed that it gives the best output. After the tokenization, noisy text removal steps: punctuation removal, non-ASCII character removal, and stop word removal. A collected English text corpus may contain characters of other languages, such as Chinese, Japanese, or Russian, and such non-English terms are removed to improve the clustering accuracy.

Forum posts normally are not as formal as news articles or technical reports, so they may contain internet slang words, text faces (emoji in the text form), or typos which are non-security related terms for this study. By using the common English words as the base of the stop word list, the proposed data cleaning method acquires more stop words including common internet slang terms [29] to make token pruning more effective. To improve token pruning, it further removes nonsense or non-security terms by regular expression rules such as too long words or with many repeated letters.

4.3 Word Embedding

Word embeddings are a type of word representation that allows words with similar meanings to have a similar representation. As the embedding model may affect the proposed clustering performance, this study employs TFIDF to compute term importance and compares two embedding models, Word2Vec and Doc2Vec, in order to find an efficient embedding model.

4.4 Analysis and Extraction

The analysis and extraction module outlined in Figure 2 utilizes a two-stage clustering, where the first clustering (theme topic clustering) determines the theme topics of a corpus and the second clustering (CTI event detection) extracts cyber threat information of each selected topic. As some topic clusters produced from the first clustering may contain non-security related topics or general information without security focus, a set of selection criteria is developed to obtain security-focused topic clusters.

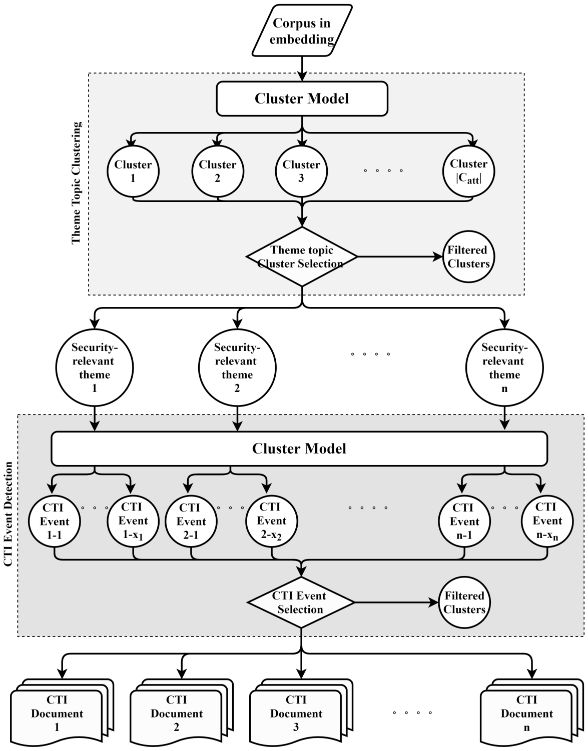


Figure 2: Analysis and extraction process

4.4.1 Theme Topic Cluster Selection Criteria

A key issue of cluster analysis is to identify clusters of the subject matter. This study develops a set of selection/filtering rules to extract security-relevant theme topic clusters, where Table 2 outlines the selection criteria. The first two rules exclude the clusters of extreme size, where an extreme size is smaller than the minimum portion or larger than the maximum portion of the corpus and expressed as below.

$$\begin{cases} |TC| < R_{\min} \times \mid \text{Corpus} \mid \\ |TC| > R_{\max} \times \mid \text{Corpus} \mid \end{cases} \tag{2}$$

Table 2: The proposed theme topic cluster selection rules

Rule ID (Action)	Description
TR1 (Removed)	A too-small cluster is removed.
TR2 (Removed)	A loo-large cluster is removed.
TR3 (Selected)	A cluster whose top k keywords are all security hashtags is selected.
TR4 (Removed)	A cluster whose top k keywords could not contribute the most term weighting is removed.
TR5 (Removed)	A cluster whose top m keywords contain non-security hashtags or keywords is removed.

Based on our preliminary study, a large cluster covers a broad range of documents and might not be able to distinguish a specific interest theme, while a small cluster contains little information to form a meaningful theme topic. Based on our preliminary study, a cluster i is considered to be too small, if the number of documents in the cluster is less than 1/50 of the corpus size, i.e.,$|TC_i| < 0.02 \times |Corpus|$;it is too large, if its size is larger than a quarter of the corpus, i.e., $|TC_i| > 0.25 \times |Corpus|$. That is, $R_min = 0.02$ and R_max = 0.25. The third rule selects clusters containing security hashtags, which implies that such clusters discuss mostly security-related information.

Based on our preliminary study by manually examining clustering results, a topic cluster with few keywords of high weighting often contains documents of a specific focus; on the contrary, that with many keywords of similar weighting likely contains diversified documents. Therefore, to identify a focused cluster, the fourth rule checks if there is a large discrepancy drop between two consecutive keywords, It computes the discrepancies of the top k keywords, where the discrepancy of the j-th keyword, $D\left(W_j\right) = \left(S\left(W_j\right) - S\left(W_{j+1}\right)\right)/S\left(W_{j+1}\right)$

S(W) is the TFIDF score of a keyword W in the cluster, and j $\in \{1, 2, \ldots, k\}$. If the first k discrepancies are not significant, which implies that this cluster contains no significant focused keywords and is not selected into the list. In this study, k=3, m=10, and a cluster is removed if the discrepancy $D\left(W_j\right) < 1.2$. For the fifth rule, a topic cluster containing non-security hashtags or keywords, such as #HIDDEN#, #QUOTE#, thankman, or job, implies that this cluster does not focus on security and is removed from the list.

4.4.2 Determining the Cluster Size

A fundamental step for unsupervised algorithms is to determine the number of clusters into which the data may be clustered. Exploring and retrieving meaningful information efficiently relies heavily on the cluster size. A good clustering produces clusters that are relatively homogeneous within themselves and heterogeneous between each other. Based on this idea, clustering metrics have been proposed to evaluate the quality of clustering results from different aspects. This study selects the number of clusters by considering the following common metrics: elbow method [39], Silhouette Coefficient [35], Calinski-Harabaz Index [3], and Davies-Bouldin Index [8].

4.4.3 CTI Event Detection

After applying the selection rules on the first stage clustering, the proposed system produces a set of security-focused topic clusters. The documents in a single topic cluster contain narrow-domain information as they contain similar keywords. The literature review [23] indicates that clustering narrow-domain texts could be challenging, as narrow-domain leads to keyword overlappings and makes it hard to distinguish sub-domains. As the past research suggests that LDA yields good clustering results, this study employs LDA to perform the second stage clustering. Like the first stage clustering, it may contain non-security focused event clusters, so the following filtering rules are applied.

ER1: A too-small cluster is removed, where a cluster of the size less than 3 is too small.

ER2: A cluster whose top m keywords contain non-security hashtags or keywords is removed.

5 System Evaluation

This study designs the following evaluation to address the proposed research questions as explained below.

- For the first research question, how to preprocess forum posts effectively to extract meaningful content, Experiment I compares the proposed data cleaning method with the traditional approach.
- For the second research question: how to validate the effectiveness of the clustering results, the study defines a clustering effectiveness measure, Embedding Cluster Score (EC_Score), to validate the results of the topic clustering. Experiment II evaluates the efficiency of the proposed method on the topic clustering with different embedding and clustering models.
- For the third research question, how to explore hacker forums and extract proactive CTI efficiently by clustering, this study proposed a hybrid solution that combines text tagging and clustering models to extract CTI information. Experiment III examines the performance of the CTI information extraction.

The study chooses a hacker forum dataset CrackingArena provided by AZSecure to evaluate the proposed solution, which was one of the largest hacker forums existing in 2018 with 11,977 active users. It contains a total of 44,927 posts dated from April 2013 to February 2018.

5.1 Experiment I: Evaluating the Effectiveness of Data Cleaning

Experiment I compares the performance of the proposed data cleaning and tagging method with the traditional data cleaning method that removes common stop words. The resulted corpora after the two data cleaning methods have been validated through human inspection. Table 3 lists the number of posts of each hashtag, and Table 4 lists the number of tokens (word terms) before and after data cleaning and tagging. The results illustrate that the proposed data cleaning and tagging method is effective in reducing the token/feature dimension. The total number of posts is 44,927 and is reduced to 1,543 after the proposed data cleaning process. This experiment also finds out that the forum posts contain quite a lot of nonsense terms such as long words or words with repeated letters.

5.1.1 Performance Measure

To identify an optimal cluster number of a given cluster model, this study considers the following commonly-used clustering metrics: elbow method, Silhouette Coefficient, Calinski-Harabaz Index (CHI), and Davies-Bouldin Index (DBI) as explained in the above section. To compare the performance of the different cluster models, this study defines a performance measure, Embedding Cluster Score (EC_Score), that considers two factors: (1) examining if

Table 3: The number of posts of each hashtag

Hashtag	posts
#HIDDEN#	315
#IMAGE#	791
#ATTACHMENT#	24
#URL#	774
#QUOTE#	171
#MODERATOR#	3
#ICQ#	79
#ACC_PASS#	12
#E-MAIL#	104
#WEBSITE#	113
#EXPLOIT#	30
#ATTACK#	30
#MALWARE#	20
#PROXY#	160
#PAYMENT#	118
#PORN#	99
#TUTORIAL#	55
#ANTIVIRUS#	27

Table 4: The efficiency comparison of token prune

Original token volume	Traditional	This study	
		Without removing nonsense terms	With nonsense terms
50,310	48,909	22,688	20,222

the cluster model can produce security-focused clusters effectively; (2) examining if the cluster model can produce a clustering result of similar-sized clusters.

For the first factor, the effectiveness is examined by the number of the selected theme topic clusters over the total number of the clusters. The selected clusters are security-related, so the more selected clusters imply the cluster model could generate security-focused clusters more effectively.

According to the selection rules listed in Table 2, the extreme-sized clusters are unfitted. For the second factor, a too-large cluster with dense data points implies that the applied word embedding model or the cluster model is not suitable to generate good clustering, while a too-small cluster results from overfitting. Both situations have a negative impact on information retrieval, so the score penalizes them. A good cluster model yields efficient clustering results with security-focused clusters and no or few unfitted clusters. Therefore, the EC_Score is expressed below.

$$EC_score = \frac{|C_{stt}|}{|C_{att}|} \times \left(1 - \frac{|C_{unfit}|}{|C_{att}| - |C_{stt}|}\right) \tag{3}$$

5.2 Experiment II: Evaluating the Performance of Theme Topic Cluster Model

The efficiency of a cluster-based extraction method might depend on with or without word embedding and the applied clustering model. Two embedding models, Word2Vec (W2V) and Doc2Vec (D2V), and their variations are evaluated; three clustering models, K-means, hierarchical cluster (HC), and LDA, are examined. One of the most common approaches, Exp II-1: TFIDF+K-means (without word embedding) is chosen to be the baseline comparison, and a summary of the Exp II results is outlined in Table 5. According to the summarized performance results described in Table 5, Exp II-3: W2V (Skip-Gram)+K-means yields the best theme topic clustering, as it has the highest EC_Score and produces the most security-relevant clusters efficiently without extreme sizes, and Exp II-9 proves to be the worst cluster model. Due to the paper limit, only the clustering results of the baseline, best, and worse clustering models are elaborated in detail, namely Exp II-1 (Baseline): TFIDF + K-means, II-3: W2V (Skip-Gram) + K-means, and II-9: D2V (PV-DM)+ HC.

Table 5: The performance results of Experiment II

EXP ID	$\lvert C_{att}\rvert$	$\lvert C_{stt}\rvert$	$\lvert C_{unfit}\rvert$	EC_Score
Exp II-1(Baseline): TFIDF + K-means	13	5	1	33.7%
Exp II-2: W2V (CBOW) + K-means	19	7	3	27.6%
Exp II-3: W2V (Skip-Gram) + K-means	15	7	0	**46.7%**
Exp II-4: D2V (PV-DM) + K-means	16	3	5	11.5%
Exp II-5: D2V (DBOW) + K-means	17	4	3	18.1%
Exp II-6: TFIDF + HC	16	6	5	18.8%
Exp II-7: W2V (CBOW) + HC	16	5	4	19.9%
Exp II-8: W2V (Skip-Gram) + HC	16	6	2	30%
Exp II-9: D2V (PV-DM)+ HC	26	2	14	3.1%
Exp II-10: D2V (DBOW)+ HC	13	4	4	17.1%
Exp II-11: LDA	11	4	6	5.2%

5.2.1 Exp II-1(Baseline): TFIDF+K-means

Figure 3 shows how to determine the optimal number of clusters by observing the curve changes of the cluster indexes described in the above section, where the navy blue vertical line indicates an optimal cluster number (13 clusters) and is identified when there are large slope changes appeared in the considered four cluster indexes. Table 6 lists the detailed clustering results and the selected theme topic clusters. The results show that the baseline (TFIDF + K-means) produces a quite good quality of clustering results with only 1 over-sized, unfitted, cluster.

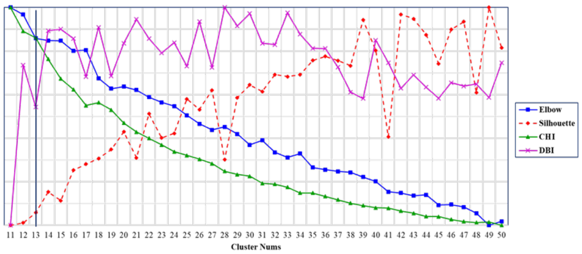


Figure 3: The cluster metrics of Exp II-1: TFIDF+K-means

Table 6: The clustering results of Exp II-1: TFIDF+K-means

ID	Top 3 terms	Rule	Posts
0	#PROXY#, proxy, #URL#'	Selected*	65
1	#IMAGE#, #HIDDEN#, #URL#	TR5	203
2	USER, ACTION, RedURL	Selected	32
3	shell, c99.txt, r57	Selected	23
4	Watchdog, community, stay	TR4	21
5	Proxy, #PROXY#, View	Selected	27
6	#URL#, #PAYMENT#, #IMAGE#	TR2	417
7	#PORN#, Site, #URL#	TR5	61
8	#IMAGE#, #URL#, #QUOTE#	TR5	81
9	account, #IMAGE#, post	TR5	183
10	#URL#, slot, machine	Selected	129
11	site, crack, config	TR5	178
12	FULLZ, Number, GOOD	TR4	36

5.2.2 Exp II-3: Word2Vec(Skip-Gram)+K-means

Figure 4 shows how to determine the optimal number of clusters by observing the curve changes of the cluster indexes described in the above section, where the vertical line indicates an optimal cluster number (15 clusters) and is identified when there are large slope changes appeared in the considered four cluster indexes. Table 7 lists the detailed clustering results and the selected theme topic clusters. The results demonstrate that the combination (W2V(Skip-Gram)+K-means) produces the best quality of clustering among all the cluster and embedding models and no unfitted cluster.

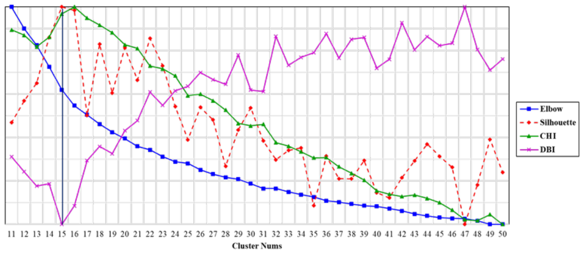


Figure 4: The cluster metrics of Exp II-3: W2V(Skip-Gram)+K-means

Table 7: The clustering results of Exp II-3: W2V(Skip-Gram)+K-means

ID	Top 3 terms	Rule	Posts
0	#URL#, fdfc119f0fb1ddbe54 5829f1777db354	Selected*	50
1	#PROXY#, proxy, list	Selected	58
2	FULLZ, Number, GOOD	TR4	36
3	#IMAGE#, #URL#, site	TR5	294
4	#IMAGE#, account, post	TR5	291
5	#IMAGE#, #URL#, #HIDDEN#	TR5	233
6	USER, ACTION, RedURL	Selected	33
7	shell, c99.txt, r57	Selected	23
8	#IMAGE#, #URL#, #HIDDEN#	TR5	91
9	slot, #URL#, machine	Selected	78
10	Proxy, #PROXY#, View	Selected	28
11	#PAYMENT#, CC, dump	Selected	35
12	stay, community, Watchdog	TR4	25
13	Site, #PORN#, Access	TR5	36
14	#URL#, #IMAGE#, Windows	TR5	145

5.2.3 Exp II-9: Doc2vec (PV-DM) + Hierarchical Cluster

Figure 5 illustrates how to determine the optimal cluster size by observing the curve changes of the cluster indexes described in the above section, where the vertical line indicates an optimal cluster size (26 clusters) is suggested by the indexes. Table 8 lists the detailed clustering results and the selected theme topic clusters. The results show that the combination (D2V(PV-DM)+HC) produces the worst and uneven clustering and could not identify security-focused clusters efficiently, where more than half (14 clusters) are unfitted (TR1 and TR2), about one third (8 clusters) contain non-security related topics (TR5), and only two security-related clusters are selected.

In summary, the results of Exp II indicate that both word embedding and cluster models impact the clustering performance. The worst cluster model fails to distinguish domain-relevant information so that it could not produce efficient clustering results. Furthermore, by comparing the clustering results of the best and worst models (Tables 7 and 8), the number of unfitted clusters affects the clustering efficiency as well, as extreme-sized clusters could not distinguish domain information well.

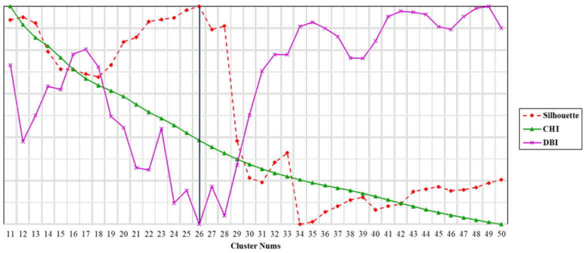


Figure 5: Analysis and extraction process

5.3 Experiment III: Evaluating the Performance of CTI Event Detection Model

If the first stage clustering fails to identify security-focused clusters, the second stage clustering for CTI information extraction might be affected. Therefore, Exp III employs the clustering results from the best cluster model obtained from Exp II (namely, Exp II-3) and adopts LDA to identify CTI events, where Table 9 summarizes the selected clusters from the best cluster model, Table 10 plots the LDA clustering results with coherence validation, and Table 11 outlines the resulted CTI event detection. In Table 10, high coherence indicates the clustering is efficient and could divide the data into a set of meaningful CTI events.

In the theme topic cluster ID: 0, URL Lists, the LDA-based CTI event detection model identifies 2 event clusters: account information and blog lists, where account information includes media platforms like Netflix and RapidGator.Net. The cluster ID: 1, Proxy 1, is further grouped into several types of proxy tools. The cluster ID: 6, System Configuration, contains various system configuration issues including rarefile.net, Sentry, UFC.TV, movies4you.tv, etc., so it is further grouped into 7 clusters. The cluster ID 7, Malicious Script, contains mostly malicious php script files shared by the same writer who posted the same script at various times, so it is grouped into one cluster. Likewise, the cluster of Gambling exhibits the same situation and results. The cluster of Proxy 2 is further grouped into two event clusters: proxy code and grabber tools by the LDA cluster model, as both belong to different types of proxy information. The cluster of Dump contains all about credit card information leakage and is further divided into 6 event clusters, where each event cluster contains data leakage from one data breach broker.

Table 8: The clustering results of Exp II-9: D2V(PV-DM)+HC

ID	**Top 3 terms**	**Rule**	**Posts**
0	USER, ACTION, GifStart=2	Selected*	29
1	#URL#, #IMAGE#, slot	TR5	182
2	#IMAGE#, #URL#, post	TR5	32
3	#IMAGE#, #URL#, #PORN#	TR5	226
4	CC, Classic, #E-MAIL#	TR1	19
5	#IMAGE#, #URL#, #HIDDEN#	TR5	81
6	#URL#, #IMAGE#, #HIDDEN#	TR5	229
7	#URL#, #IMAGE#, slot	TR5	319
8	#URL#, slot, #IMAGE#	TR5	76
9	der, yang, dan	TR1	5
10	shell, #URL#, c99	TR5	38
11	#URL#, NETFLIX, Site	TR1	9
12	#ACC_PASS#, dump, gold/plat/bus/corp/sign	TR1	5
13	import_module, process_report, process_report_data	TR1	1
14	#URL#, shell, c99.txt	TR4	66
15	#PROXY#, proxy, service	Selected	43
16	FULLZ, Site, GOOD	TR4	68
17	ACTION, recaptcha_response_field= manual_challenge, USER	TR1	3
18	IDM, Internet, download	TR1	2
19	#ACC_PASS#, #ANTIVIRUS#, #URL#	TR1	13
20	#WEBSITE#, DropBox.com, BitShare.com	TR1	1
21	#URL#, /etc/, Apache	TR1	1
22	href, div, /div	TR1	2
23	href, class, /li	TR1	1
24	x15, x78, x75	TR1	1
25	track1/2, -Dumps, pin	TR1	4

By manually examining the LDA clustering results as

Table 9: The selected clusters from the best first stage cluster model (Exp II-3)

ID	Theme topic	Keywords	Posts
0	URL Lists	#URL#, fdfc119f0fb1ddbe545829f1777db354, #E-MAIL#, NETFLIX, #PORN#, MoneyMakingDiscussion.Net, Visit, amateur, March, Bonus	50
1	Proxy 1	#PROXY#, proxy, list, #IMAGE#, combo, Proxy, test, Support, ban, VPN	58
6	System Configuration	USER, ACTION, RedURL, #URL#, blnDigits=1, blnMultiChar=0, Range=0, URLMode=0, Brightness=0, GifOffset=2	33
7	Malicious Script	shell, c99.txt, r57, c99, script, tool, r57.txt, inurl:c100.txt, inurl:c100.php, inurl:locus.txt	23
9	Gambling	slot, #URL#, machine, free, game, casino, Free, play, online, Slot	78
10	Proxy 2	Proxy, #PROXY#, View, Click, Code, #URL#, Text, Attention, directly, Sign	28
11	Dump	#PAYMENT#, CC, dump, #ICQ#, Classic, Dumps, #E-MAIL#, sell, Gold, Canada	35

Table 10: The LDA event clustering results

ID	Theme topic	Event topics	Alpha	Beta	Coherence
0	URL lists	2	0.71	0.11	0.6307
1	Proxy 1	16	0.11	0.21	0.6083
6	System configuration	7	0.61	0.91	0.6076
7	Malicious script	1	0.01	0.01	0.5944
9	Gambling	1	0.61	0.81	0.5923
10	Proxy 2	2	0.21	0.21	0.5831
11	Dump	6	0.81	0.01	0.5819

Table 11: The extracted CTI information

ID	Theme topic	Event cluster	Posts
0	URL lists	Account/password information of media platforms	8
		Russia blog lists	42
1	Proxy 1	Sockshub/rsocks	7
		Fast Proxy Tester/ Checker	11
		ProxyFire	5
6	System configuration	Various system config info	33
7	Malicious script	Sharing php-based malware scripts	23
9	gambling	Tupantitty online gambling	78
10	Proxy 2	Proxy Code	5
		Proxy Grabber	7
11	Dump	Selling privacy data in 6 types	36

described above, the proposed two-stage clustering approach discovers CTI information efficiently. In summary, based on the above three experiments, the evaluation concludes that the proposed CTI information retrieval method can explore hacker forums well and extract cybersecurity information efficiently.

6 Conclusion

Acquiring cyber threat knowledge is essential for organizations to gain visibility into the fast-evolving threat landscape. Hacker forums play an important role in disseminating threat information and correlate significantly with the number of cyber-attacks observed in the real world [42]. Most past research focused on identifying threat intelligence with patterns by classification models. Clustering and preprocessing the content of hacker forums is challenging as the number of clusters is hard to determine and forum writers tend to write freestyle and diversified article posts.

This study applies NLP, tagging, and clustering techniques to explore and capture cybersecurity information in hacker forums. The proposed CTI information retrieval method applies tagging and Word2Vec word embedding to extract key features and employs K-means and LDA two-stage clustering to discover CTI information from unstructured data. Based on Exp I, the proposed data cleaning and tagging method reduces the feature dimension significantly by more than two times better than the traditional data cleaning method, from the size of 48,909 to 20,222. Exp II and III demonstrate that the proposed theme topic cluster selection criteria trim off non-security relevant clusters effectively and the two-stage clustering method can capture cybersecurity-related article posts efficiently.

For determining the clustering size, this study finds out that considering multiple cluster evaluation metrics is effective in finding good clustering parameters. The proposed performance metric, EC_Score, is proved to be helpful for determining the best combination of word embedding and clustering models. This study has demonstrated that applying both text classification and clustering models can achieve great performance in exploring and extracting CTI information efficiently.

Future work can extend this research to explore online hacker forums in multiple languages or increase understanding of other hacker online community platforms. In addition to increasing the variety of platforms or languages, future work can look at social relationships among hackers and hacker groups or identifying the members creating and disseminating CTI by using social network analysis techniques. This work can also be expanded by introducing a temporal component to track the prevalence of a specific CTI topic over time, which is useful for identifying emerging CTI technologies.

References

[1] V. Benjamin and H. Chen, "Developing understanding of hacker language through the use of lexical semantics," in *IEEE International Conference on Intelligence and Security Informatics (ISI'15)*, pp. 79–84, 2015.

[2] B. Biswas, A. Mukhopadhyay, and G. Gupta, "Leadership in action: How top hackers behave a big-data approach with text-mining and sentiment analysis," in *Proceedings of the 51st Hawaii International Conference on System Sciences*, 2018.

[3] T. Caliński and J. Harabasz, "A dendrite method for cluster analysis," *Communications in Statistics-theory and Methods*, vol. 3, no. 1, pp. 1–27, 1974.

[4] S. Chandel, J. Wei, and B. T. Chu, "A natural language processing based trend analysis of advanced persistent threat techniques," in *IEEE International Conference on Big Data (Big Data'18)*, pp. 2995–3000, 2018.

[5] L. Chen, L. Hu, Z. Zheng, J. Wu, J. Yin, Y. Li, and S. Deng, "Wtcluster: Utilizing tags for web services clustering," in *International Conference on Service-Oriented Computing*, pp. 204–218, 2011.

[6] Z. Chen, M. Trabelsi, J. Heflin, Y. Xu, and B. D. Davison, "Table search using a deep contextualized language model," in *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 589–598, 2020.

[7] CloudSEK Threat Intelligence Team, *Dave Suffers Breach, 7.5m Users' Data Leaked, Meow Attack Deletes 4,000 Unsecured Databases, and More*, Sept. 13, 2020. (https://cloudsek.com/threatintel/dave-suffers-breach-7-5m-users-data-leaked-meow-attack-deletes-4000-unsecured-databases-and-more/)

[8] D. L. Davies and D. W. Bouldin, "A cluster separation measure," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-1, no. 2, pp. 224–227, 1979.

[9] I. Deliu, C. Leichter, and K. Franke, "Extracting cyber threat intelligence from hacker forums: Support vector machines versus convolutional neural networks," in *IEEE International Conference on Big Data (Big Data'17)*, pp. 3648–3656, 2017.

[10] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," *arXiv preprint arXiv:1810.04805*, 2018.

[11] A. S. Gautam, Y. Gahlot, and P. Kamat, "Hacker forum exploit and classification for proactive cyber threat intelligence," in *International Conference on Inventive Computation Technologies*, pp. 279–285, 2019.

[12] M. Guderlei, M. Aßenmacher, "Evaluating unsupervised representation learning for detecting stances of fake news," in *Proceedings of the 28th International Conference on Computational Linguistics*, pp. 6339–6349, 2020.

[13] A. Gupta and A. Anand, "Ethical hacking and hacking attacks," *International Journal of Engineering and Computer Science*, vol. 6, no. 4, 2017.

[14] A. Handler, *An Empirical Study of Semantic Similarity in WordNet and Word2Vec*, Theses and Dissertations, University of New Orleans, 2014.

[15] InfoSecurity, *Hackers Forums Provide Sense of Community, Information Security Intelligence*, Sept. 13, 2020. (https://www.infosecurity-magazine.com/news/hackers-forums-provide-sense-of-community/)

[16] C. Johnson, M. Badger, D. Waltermire, J. Snyder, and C. Skorupka, *Guide to Cyber Threat Information Sharing*, Report, National Institute of Standards and Technology, 2016.

[17] M. Kadoguchi, S. Hayashi, M. Hashimoto, and A. Otsuka, "Exploring the dark web for cyber threat intelligence using machine leaning," in *IEEE International Conference on Intelligence and Security Informatics (ISI'19)*, pp. 200–202, 2019.

[18] O. Khattab and M. Zaharia, "Colbert: Efficient and effective passage search via contextualized late interaction over bert," in *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 39–48, 2016.

[19] Y. Kim, "Convolutional neural networks for sentence classification," *arXiv preprint arXiv: 1408.5882*, 2014.

[20] S. Kumar and D. Agarwal, "Hacking attacks, methods, techniques and their protection measures," *International Journal of Advance Research in Computer Science and Management*, vol. 4, no. 4, pp. 2253–2257, 2018.

[21] Y. Kurogome, Y. Otsuki, Y. Kawakoya, M. Iwamura, S. Hayashi, T. Mori, and K. Sen, "Eiger: Automated ioc generation for accurate and interpretable endpoint malware detection," in *Proceedings of the 35th Annual Computer Security Applications Conference*, pp. 687–701, 2019.

[22] Q. Le and T. Mikolov, "Distributed representations of sentences and documents," in *International conference on machine learning*, pp. 1188–1196, 2019.

[23] C. Li, Y. Lu, J. Wu, Y. Zhang, Z. Xia, T. Wang, D. Yu, X. Chen, P. Liu, and J. Guo, "Lda meets Word2Vec: A novel model for academic abstract clustering," in *Companion Proceedings of the the Web Conference 2018*, pp. 1699–1706, 2018.

[24] W. Li, H. Chen, and J. F. Nunamaker Jr, "Identifying and profiling key sellers in cyber carding community: Azsecure text mining system," *Journal of Management Information Systems*, vol. 33, no. 4, pp. 1059–1086, 2016.

[25] X. Liao, K. Yuan, X. F. Wang, Z. Li, L. Xing, and R. Beyah, "Acing the ioc game: Toward automatic discovery and analysis of open-source cyber threat intelligence," in *Proceedings of the ACM SIGSAC Conference on Computer and Communications Security*, pp. 755–766, 2019.

[26] C. Liu, X. Wu, M. Yu, G. Li, J. Jiang, W. Huang, and X. Lu, "A two-stage model based on bert for short fake news detection," in *International Conference on Knowledge Science, Engineering and Management*, pp. 172–183, 2019.

[27] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," in *Advances in neural information processing systems*, pp. 3111–3119, 2013.

[28] V. Mohan, "Text mining: Open source tokenization tools: An analysis," vol. 3, pp. 37–47, 2016.

[29] E. Montalbano, *Experts on Seller Floods Hacker Forum with Data Stolen from 14 Companies*, Sept. 13, 2020. (`https://threatpost.com/threat-actors-introduce-unique-newbie-hacker-forum/157489/`)

[30] R. Nogueira, W. Yang, K. Cho, and J. Lin, "Multistage document ranking with bert," *arXiv preprint arXiv: 1910.14424*, 2019.

[31] K. Orkphol and W. Yang, "Word sense disambiguation using cosine similarity collaborates with Word2Vec and WordNet," *Future Internet*, vol. 11, no. 5, p. 114, 2019.

[32] R. Padaki, Z. Dai, and J. Callan, "Rethinking query expansion for bert reranking," in *European Conference on Information Retrieval*, pp. 297–304, 2020.

[33] S. Pastrana, D. R. Thomas, A. Hutchings, and R. Clayton, "Crimebb: Enabling cybercrime research on underground forums at scale," in *Proceedings of the World Wide Web Conference*, pp. 1845–1854, 2018.

[34] L. T. B. Ranera, G. A. Solano, and N. Oco, "Retrieval of semantically similar philippine supreme court case decisions using doc2vec," in *International Symposium on Multimedia and Communication Technology (ISMAC'19)*, pp. 1–6, 2019.

[35] P. J. Rousseeuw, "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis," *Journal of Computational and Applied Mathematics*, vol. 20, pp. 53–65, 1987.

[36] S. Samtani, R. Chinn, and H. Chen, "Exploring hacker assets in underground forums," in *IEEE international conference on intelligence and security informatics (ISI'15)*, pp. 31–36, 2015.

[37] Security Experts, *Experts on Seller Floods Hacker Forum with Data Stolen from 14 Companies*, Sept. 13, 2020. (`https://www.informationsecuritybuzz.com/expert-comments/experts-on-seller-floods-hacker-forum-with-data stolen-from-14-companies/`)

[38] M. Shi, J. Liu, D. Zhou, M. Tang, and B. Q. Cao, "WE-LDA: A word embeddings augmented LDA model for web services clustering," in *IEEE International Conference on Web Services (ICWS'17)*, 2017.

[39] R. L. Thorndike, "Who belongs in the family?," *Psychometrika*, vol. 18, no. 4, pp. 267–276, 1953.

[40] W. Tian, J. Li, and H. Li, "A method of feature selection based on word2vec in text categorization," in *37th Chinese Control Conference (CCC'18)*, pp. 9452–9455, 2018.

[41] R. Vijjali, P. Potluri, S. Kumar, and S. Teki, "Two stage transformer model for covid-19 fake news detection and fact checking," *arXiv preprint arXiv: 2011.13253*, 2020.

[42] Q. H. Wang, W. T. Yue, and K. L. Hui, "Do hacker forums contribute to security attacks?," in *E-Life: Web-Enabled Convergence of Commerce, Work, and Social Life*, pp. 143–152, 2012.

[43] B. Wollschlaeger, E. Eichenberg, and K. Kabitzsch, "Explain yourself: A semantic annotation framework to facilitate tagging of semantic information in health smart homes," in *HEALTHINF*, pp. 133–144, 2020.

[44] W. T. Yue, Q. H. Wang, and K. L. Hui, "See no evil, hear no evil? dissecting the impact of online hacker forums," *MIS Quarterly*, vol. 43, no. 1, p. 73, 2019.

[45] R. Zellers, A. Holtzman, H. Rashkin, Y. Bisk, A. Farhadi, F. Roesner, and Y. Choi, "Defending against neural fake news," *arXiv preprint arXiv: 1905.12616*, 2019.

[46] J. Zhan, J. Mao, Y. Liu, M. Zhang, and S. Ma, "An analysis of bert in document ranking," in *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 1941–1944, 2020.

Biography

Chia-Mei Chen has joined in the Department of Information Management, National Sun Yat-Sen University since 1996. She was the Section Chef of Network Division and Deputy Director, Office of Library and Information Services in 2009-2011. She had served as a coordinator of TWCERT/CC (Taiwan Computer Emergency Response Team/Coordination Center) during 1998 to 2013 and then as a consultant until 2018. Based on her CSIRT experience, she established TACERT (Taiwan Academic Network Computer Emergency Response Team) in 2009. She was a Deputy Chair of TWISC@NCKU, a branch of Taiwan Information Security Center during 2017 to 2020. She continues working for the network security society. Her current research interests include anomaly detection, network security, machine learning, text mining, and big data analysis.

Dan-Wei Wen is an assistant professor at the Department of Information Management, Tamkang University. She received her Ph.D. from the Department of Business Administration, National Cheng-Kung University. Her research interests include industry dynamics, catching-up strategy, and data mining.

Ya-Hui Ou received her Ph.D. degree from the Department of Information Management, National Sun Yat-sen University in 2017. She is an assistant professor in the Common Education Teaching Center, National Penghu University of Science and Technology, Penghu, Taiwan. Her research interests include network security and statistical analysis.

Wei-Chih Chao has received his Master's degree from the Department of Information Management, National Sun Yat-sen University. Currently he is a software engineer in an information security institute.

Zheng-Xun Cai received his Master's degree from the National Sun Yat-sen University in 2017 and continues pursuing the PhD degree at the same school. His research focuses on digital forensics, network analysis, and intrusion detection.

Reviewers (Volume 23, 2021)

Dariush Abbasinezhad
Tarek Abbes
Ahmed Abd El-Rahiem Abd El-Latif
Slim Abdelhedi
Mohd Faizal Abdollah
Ahmed Mohammed Abdullah
Subrata Acharya
Sodeif Ahadpou
Tohari Ahmad
Muhammad Najmi Ahmad-Zabidi
Mohammad Reza Ahmadi
Asimi Ahmed
Ganesh V. Aithal
Mehrnaz Akbari Roumani
Abdul-Gabbar Tarish Al-Tamimi
Aws N. Al-Zarqawee
Monjur M Alam
Shahid Alam
Tanweer Alam
Dilip S Aldar
Sara Ali
Ali Mohamed Allam
Khalid Abdulrazzaq Alminshid
Seth Alornyo
Ali Mohammed Alsahlany
Richard Amankwah
Ruhul Amin
Rengarajan Amirtharajan
R. Anand
Karl Andersson
Benjamin Arazi
K. S. Arvind
Muhammad Asad
Travis Atkison
Hany Fathy Atlam
Cossi Blaise Avoussoukpo
Anant M. Bagade
Amandeep Bagga
Nazrulazhar Bahaman
Nischay Bahl
Anuj Kumar Baitha
Saad Haj Bakry
R. R. Balakrishnan
Kavitha Balu
Maram Y Bani Younes
Tamer Mohamed Barakat
Utpal Barman
Pijush Barthakur
Eihab Bashier Mohammed Bashier
Adil Bashir
Sunny Behal
Rydhm Beri
Taran Singh Bharati
Akashdeep Bhardwaj
Lathies T. Bhasker
Sugandh Bhatia
Sajal Bhatia
Dharmendra Bhatti
Krishna Bhowal
Li Bin
Sumitra Binu
Zhengjun Cao
Liling Cao
Chi-Shiang Chan
Eric Chan-Tin
Mohan Kumar Chandol
Yogesh Chandra
Arup Kumar Chattopadhyay
Nirbhay K. Chaubey
Ali M Chehab
Chi-Hua Chen
Chin-Ling Chen
Jan Min Chen
Tzung-Her Chen
Xi Chen
Yang Chen
Yi-Hui Chen
Yushuang Chen
Zhixiong Chen
Qingfeng Cheng
Kaouthar Chetioui
Mao-Lun Chiang
Shu-Fen Chiou
Tae-Young Choe
Kim-Kwang Raymond Choo
Christopher P. Collins
Joshua C. Dagadu
Ashok Kumar Das

Prodipto Das
Sanjoy Das
Debasis Das
Ranjan Kumar Dash
Subhrajyoti Deb
Abdelrahman Desoky Desoky
Mooramreddy Sree Devi
Sankhanil Dey
Subhasish Dhal
Jintai Ding
Jingnan Dong
Xiaoli Dong
Nishant Doshi
Ahmed Drissi
Crystal Wilson Dsouza
Qi Duan
Ashraf Diaa Elbayoumy
Abd Allah Adel Elhabshy
Ahmed A. Elngar
Edwin Engin Yaz
Aoxiong Fan
Arizona Firdonsyah
Xingbing Fu
Vladimir Sergeevich Galyaev
Rakesh C Gangwar
Juntao Gao
Tiegang Gao
Xinwei Gao
N. B. Gayathri
G. Geetha
Mohammad GhasemiGol
Madhumala Ghosh
Ramesh Gopalan
Poornima Ediga Goud
Krishan Kumar Goyal
Ke Gu
Avinash k Gulve
Sumalatha Gunnala
Shuai Guo
C. P. Gupta
Jatin Gupta
Pynbianglut Hadem
Charifa Hanin
Ali Hassan
Wien Hong
Tsung-Chih Hsiao
Chengyu Hu
Defa Hu
Xiong Hu
Yen-Hung Hu
Huajun Huang
Chin-Tser Huang
Jianmeng Huang
Munawar Hussain
Bala Venkateswarlu Isunuri
Grasha Jacob
Amit Jain
Yogendra Kumar Jain
Swati Jaiswal
Teena Jaiswal
Bappaditya Jana
V. S. Janani
N Jeyanthi
lin zhi jiang
Shaoquan Jiang
Rong Jiang
Rui Jiang
Zhengping Jin
Ashish Joshi
Li Su Juan
Omprakash Kaiwartya
Yoshito Kanamori
Nirmalya Kar
Gagandeep Kaur
Wongyos Keardsri
Omar Khadir
Vaishali D. Khairnar
Asif Uddin Khan
Md. Al-Amin Khandaker
Malik Sikander Hayat Khiyal
Dong Seong Kim
Kingsford Kissi Mireku
Vikas K Kolekar
P. Dhandapani Raman D.
Kothandaraman
Anjan Krishnamurthy
Fengfei Kuang
Sajja Ratan Kumar
Manish Kumar
Naresh N Kumar
Saru Kumari
Yesem Kurt Peker
Owusu-Agyemang Kwabena
Albert Kofi Kwansah Ansah
Manmohan Lakhera
Then Lee
Cheng Li
Chun-Ta Li
Yanping Li
Zhaozheng Li
H. M. Lian
Changlu Lin

Chia-Chen Lin
Chih-Yang Lin
Iuon-Chang Lin
Yang-Bin Lin
Jiang Hong Ling
Desheng Liu
Li Liu
Shuang Gen Liu
Ting Liu
Ximeng Liu
Yanjun Liu
Yining Liu
K. Shantha Kumari Luke
Jayakumar
Zhiyong Luo
Ming Luo
Sagar Bhaskar Mahajan
Zahid Mahmood
Tanmoy Maitra
Doaa Mohsin Majeed
Arun Malik
Mahalinga V. Mandi
T. Manesh
Palvinder Singh Mann
Ali Mansouri
A. M. Meddeb-Makhlouf
Kamran Ali Memon
Bo Meng
Weizhi Meng
Yang Ming
Suhail Qadir Mir
Amit Mishra
Anuranjan Misra
Syed Shahul Hameed

Mohamed Ismail
Sirwan Ahmed Mohammed
Madihah Mohd Saudi
Guillermo Morales-Luna
Belmekki Mostafa
Alaa Moualla
Hamdy M. Mousa
Muhammad M. Muhammad
Kuntal Mukherjee
C. H. Mukundha
Bhagavathi Priya M
Muthumanikam
Ambika Nagaraj
Preeti Nagrath
K. Nandhini
Syed Naqvi
Kanagaraj Narayanasamy
Lakshmi Kannan Narnayanan
Prabir Kr Naskar
Sarmistha Neogy
Krishnamur G Ningappa
Sohail Noman
Chokri Nouar
Abdul Abiodun Orunsolu
Arezou Ostad Sharif
Nasrollah Pakniat
Dhiraj Pandey
S. K. Pandey
B. D. Parameshachari
Subhash S. Parimalla
Chintan J. Patel
Kailas Ravsaheb Patil
Suresh Kumar Peddoju
Hongmei Pei

Kanthakumar Pongaliur
A. Prakash
Krishna K. Prakash
Munivara Prasad
Hongquan Pu
Yudha Purwanto
Septafiansyah Dwi Putra
Murad Abdo Rassam Qasm
Qais Saif Qassim
Chuan Qin
Jiaohua Qin
Narasimhan Renga Raajan
Hashum Mohamed Rafiq
Abdul Hamid M. Ragab
V. Sampangi Raghav
Uma R. Rani
Ganga Rama Koteswara Rao
Golagani A.V.R.C Rao
Mohammad Maher Rasheed
V. Rathinasabapathy
Dhivya Ravi
Ramesh S Rawat
Siva Ranjani Reddi
Khaled Riad
Mohd Foad Rohani
Ou Ruan
Sanjay Kumar Sahay
Ashish Saini
Debabrata Samanta
Sabyasachi Samanta
Manju Sanghi
Arif Sari
Balamurugan K. S. Sathiah
Rajat Saxena

Michael Scott
Chandra Vorugunti Sekhar
Irwan Sembiring
Elena Sendroiu
Divyashikha Sethia
Vrutik M. Shah
Vrushank Shah
Kareemulla Shaik
Tarun Narayan Shankar
Udhayakumar Shanmugam
Rohith Shivashankar
Abhishek Shukla
Varun Shukla
Anuj Kumar Singh
Debabrata Singh
Jitendra Singh
Mahendra Pratap Singh
Mukesh Singh
Bala Srinivasan
Siva Shankar Subramanian
Karthikeyan Subramanian
T. SudalaiMuthu
K. S. Suganya
Guodong Su
Haiyan Sun
Fei Tang
Maryam Tanha
Ariel Soares Teles
Pratik Teli
Xiuxia Tian
Geetam Singh Tomar
Yuan-Yu Tsai
Pushpendra Kumar Verma
Ravi Verma
Vandani Verma
Vibhor Kumar Vishnoi
Phu Vo Ngoc
Tao Wan
Putra Wanda
Ding Wang
Fangwei Wang
Feng Wang
Guoqing Wang
Li Wang
Libin Wang
Linfan Wang
Qingping Wang
Xiaogang Wang
Xingbo Wang
Xu Wang
Ying Wang
C. H. Wei
Jianghong Wei
Zhe Wei
Axin Wu
Na-I Wu
Chengbo Xu
Degang Xu
Lei Xu
Chengbo Xu
Yashveer Yadav
Wei Yajuan
Jun Yan
Changsong Yang
Li Yang
Rui Yang
Wenjie Yang
Yifei Yao
Jun Ye
Pinghao Ye
Fangfang Yin
Huang Yiwang
Lin You
Huifang Yu
Lei Yu
Hang Yue
Taskeen Zaidi
Noor Zaman Zaman
Jianjun Zhang
Sherali Zeadally
Jianping Zeng
Fangguo Zhang
Futai Zhang
Jianhong Zhang
Jie Xiu Zhang
Qiu-Yu Zhang
Shanshan Zhang
Yanshuo Zhang
Yinghui Zhang
Zonghua Zhang
Hongzhuan Zhao
Mingju Zhao
Yuntao Zhao
Zhiping Zhou
Ye Zhu
Yingwu Zhu
Frank Zhu
Aaron Zimba

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.