IJNS

International Journal of Network Security

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

Vol. 26, No. 1 (January 2024)

INTERNATIONAL JOURNAL OF NETWORK SECURITY

PUBLISHING OFFICE

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

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

PUBLISHER: Candy C. H. Lin

23-75, P.O. Box, Taichung, Taiwan 40199, R.O.C.

Volume: 26, No: 1 (January 1, 2024)

International Journal of Network Security

Three-party Authentication Protocol Based on Riro for Mobile RFID System

Sheng-Hua Xu
(Corresponding author: Sheng-Hua Xu)

Network Information Center, Guangdong Polytechnic Normal University
Guangzhou 510006, China
Email:shengzaifenglin@163.com

(Received Sept. 17, 2022; Revised and Accepted Feb. 3, 2023; First Online Dec. 8, 2023)

Abstract

The existing protocols only apply to traditional RFID systems. This paper proposes a three-party authentication protocol for mobile RFID systems to address this problem. The protocol is based on a novel encryption operation: the operation of the remainder, rounding and reordering. The protocol skillfully combines the length variable and Hamming weight of the encryption parameter itself, performs the remainder rounding operation on them, and reorders them in different ways according to the operation results, which can break the sequence of the encryption parameter itself and reduce the introduction of other variables. At the electronic tag end, the counter mechanism is added to record the information amount sent to the electronic tag by the reader in the last round so that different operations can be performed, which can resist replay attacks and asynchronous attacks. Analyzing the protocol from multiple attack types shows that the proposed protocol can resist common types of attacks. By analyzing the performance indicators of the protocol based on the computation amount at the tag end and the overall traffic of the protocol, it is shown that the overall computation amount and communication overhead of the proposed protocol is superior to other comparison protocols.

Keywords: Internet of Things; Mobile Radio Frequency Identification System; Remainder Integer Reordering Operation; Three-party Authentication

1 Introduction

RFID is a technology that can read the data stored in a specific objects without contacting it [22], this technology is the most widely used in RFID system. A classic and traditional RFID system includes but is not limited to tags, servers and readers. Tags have many advantages, such as small size, easy to carry, simple installation, low cost, long service life [12, 13, 20], etc. Due to the above factors, RFID technology has been involved in all aspects, such as access control system [7], campus one card system [8], bus card system [2], etc.

In the traditional RFID system, the reader writer and server are generally placed in a fixed location, and the location will not change almost throughout the service life, so that they can interact with each other through a more secure wired way, such as data interaction through coaxial cable or optical fiber [10, 21]. However, with the continuous development of society and the increasing social needs of people, the traditional RFID system can't meet the complex needs of people, such as: people want to stay at home to do online shopping and online payment. This demand cannot be realized in terms of traditional RFID system. The emergence of mobile RFID system perfectly solves this problem.

In a mobile RFID system, the installation position of the reader/writer and the server may be changing at any time, and the distance between them may be far. If the coaxial cable or optical fiber and other wired methods are used for data interaction at this situation, the requirements cannot be met, so they also need to be transformed into wireless methods for data interaction. Due to the openness of wireless mode, the data transmitted in the wireless channel is easy to be eavesdropped by other irrelevant personnel, which may lead to the leakage of private data, thus causing irreparable losses to users [5, 11, 16]. Therefore, it is necessary to design a security protocol to ensure the data security of wireless transmission. Most of the existing protocols are only applicable to traditional RFID systems, that is, when designing these protocols, they only consider wireless communication between tags and readers, while wired communication between servers and readers. Therefore, if existing protocols are still used in mobile RFID systems, users' private information will be at risk of being eavesdropped and leaked, so new protocols which can be applicable to mobile RFID systems need to be designed.

In order to solve the above difficulties, this paper designs a super lightweight protocol, which can realize the three-party authentication among tags, servers and readers, and can be used in mobile RFID systems [3,18]. The biggest innovation of the protocol is that an original encryption algorithm is designed, that is, bitwise synthesis operation. This encryption algorithm combines the Hamming weight value of the encryption parameter itself, compares the size, and carries out synthesis operations in different ways and in different orders according to the comparison results. Because there are three encrypted parameters each time, and the three parameters are different, the Hamming weight values corresponding to the three parameters involved each time are also different, which makes the order used in each synthesis different, increases the difficulty of cracking, and ensures that the storage space is not increased.

The remainder of this paper is organized as follows: Section 1 introduces the research background and problems in the application, and then leads to the research and focus of the paper. Section 2 introduces the current situation at home and abroad, advantages and disadvantages of this research; Section 3 introduces the specific implementation steps of the original bitwise composition operation. In Section 4, the security defects of Wu's protocol design are described in detail. In Section 5, the design steps of the protocol are given. In Section 6, the security of the protocol is analyzed from several specific attack types. Section 7 analyzes the performance indicators of different aspects of the protocol. The last section summarizes the whole paper.

2 Related Research Works

In [6], a protocol was designed by using the quadratic remainder theorem, which can satisfy the three-party authentication among tags, readers and servers. Reference [4] makes an in-depth analysis of the designed protocol in Reference [6], pointing out that the protocol can't resist replay attacks launched by attackers, and also gives an improved protocol. Although the improved protocol avoids replay attacks, there are still the following security risks: First, the shared key value of the protocol has not changed during the entire authentication process, which will enable attackers to obtain the non-updated shared key through some methods. Second, the information sent at the beginning of the protocol is in plaintext form, which allows attackers to obtain it through eavesdropping, so that attackers can launch cloning attacks, thus causing the disclosure of users' private information and data.

In [9], a lightweight authentication protocol based on PUF and hash function is given. Although it can resist cloning attacks by using PUF function encryption, the protocol still has other security problems. Such as, the protocol cannot supplement the consumed pseudonym information in a timely manner, which makes the protocol unable to resist the counterfeiting attack; the attacker can even make the tag end consume all pseudonym data through replay attack, and then force the tag to reregister.

In [1], a super lightweight protocol is proposed. The protocol uses super lightweight operations such as inversion, XOR and shift to encrypt important data. Although the author claims that the protocol has good security, it is found that the protocol cannot resist the tracking attack of attackers [15]. In the protocol design, some messages sent by the tag are not added to the random number during the encryption process, so that the message value of each round does not change. The attacker can obtain the unchanged message through long-term eavesdropping, and then analyze the label location, thus successfully implementing the tracking attack.

In [14], an authentication protocol is designed based on the classic DH key agreement algorithm, this protocol uses the classic elliptic ECC to encrypt private data. However, the protocol does not update the shared secret value used in each round, that is, the private key is not updated, making the protocol unable to resist replay attacks. Meanwhile, the protocol uses ECC to realize data encryption, which makes the tag end unable to undertake too much calculation, as a result, the promotion and application of the protocol is not optimistic.

In [19], a protocol is designed by using algorithms such as Chinese remainder theorem, hash function and physical non clonable function. The protocol designer claims that the protocol has strong security performance and can resist a variety of attacks. However, it is found that the protocol is not as secure, moreover, the correctness of the protocol in many aspects remains to be discussed and security issues will be analyzed in detail in the following sections.

Through the analysis of the existing classical protocols, it can be found that the protocols can't be used in mobile RFID systems, or can't provide strong security requirements, or can't be applied to low-cost RFID systems due to the large computation amount. In view of the above problems, this paper presents a super lightweight three-party authentication protocol that can be applied to mobile RFID systems. The main innovation of the protocol in this paper is as follows: using an original encryption operation to achieve data encryption transmission, namely bitwise composition operation. This operation makes full use of the Hamming weight values of the three encryption parameters, compares the three Hamming weight values, and performs the synthesis operations in different ways and in different orders according to the comparison results. Without adding parameters, this operation increases the difficulty of cracking and improves the security performance of the protocol.

3 Remainder Integer Reordering Operation

The following symbols can be used to Remainder Integer Reordering Operation ($Riro\left(X\right)$), which is specifically defined as follows:

1) Set the length of X Y is constant L.

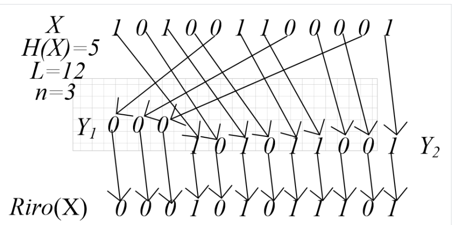


Figure 1: $Riro(X)(H(X) \geq n)$

2) Hamming weight of X is expressed by symbol $H(X)$.

3) The symbol n is used to represent the result of the remainder rounding operation of the length L of x and the hamming weight $H(X)$ of X, that is $n = \left[\frac{L}{H(x)}\right]$, rounding up.

4) When $H(X) \geq n$, select the $H(X)$ bit from the left to the right of X, and take a number every n bits in turn to form a subsequence Y_1; The remaining numbers can be arranged in order to form another subsequence Y_2. Finally, subsequence Y_1 and subsequence Y_2 can be reordered to get the result after encryption, namely $X = Y_1Y_2$.

5) When $H(X) < n$, select the n bit from the right to the left of X, and take a number every $H(X)$ bits in turn to form a subsequence Y_1; The remaining numbers can be arranged in order to form another subsequence Y_2. Finally, subsequence Y_1 and subsequence Y_2 can be reordered to get the result after encryption, namely $X = Y_1Y_2$.

The following examples can be used to deepen the understanding of the operations of remainder, rounding and reordering. When $X = 101001100001$, it can be obtained $L = 12, H(X) = 5$, according to the calculation rules of n, calculate $n = \left[\frac{L}{H(x)}\right] = \left[\frac{12}{5}\right] = 3$, when satisfy the situation $H(X) \geq n$. According to the description in (4), obtain the subsequence $Y_1 = 000$ and subsequence $Y_2 = 101011001$, and finally obtain $X = Y_1Y_2 = Riro(X) = 000101011001$. The above process diagram is shown in Figure 1.

4 Wu *et al.*'s Protocol Analysis

This section analyzes the security flaws of the protocol designed by Wu *et al.*

In view of the space limitations the specific steps of the protocol proposed by Wu *et al.* will not be listed here, and the specific implementation steps can be referred to [19]. Through in-depth analysis of the protocol, it can be found that the protocol has the following security defects and questionable contents.

First, in the section "related work", the protocol in [9] was proposed in 2018, the protocol in reference [12] was proposed in 2019, but the protocol in reference [17] was proposed in 2020. However, in the design process of the protocol in [9], some contents in reference [17] were cited. From the perspective of time, Wu *et al.*'s articles are disordered and their rationality is questionable.

Secondly, Wu *et al.* have repeatedly emphasized that the design protocol is lightweight, but analysis shows that the protocol is not lightweight. The main reasons are as follows: At the tag end, physical non-cloneable functions, the quadratic remainder theorem, the hashing function are used for calculation. As results, the final calculation amount of the protocol is not lightweight, so Wu *et al.* stressed that the correctness of the statement that the protocol is lightweight remains to be discussed.

Thirdly, the protocol proposed by Wu *et al.* can't resist replay attacks and asynchronous attacks. The specific analysis is as follows:

According to the communication steps of the protocol designed by Wu *et al.*, the following attacks can be carried out. The information VP in the ninth step can be obtained by eavesdropping, and then the attacker blocks the subsequent communication between the reader and the tag. Meanwhile, the attacker continuously sends the information obtained by eavesdropping to the tag. If the label verification information is passed, the label side will continuously update the relevant privacy information; When the number of update operations at one tag end exceeds three times, the privacy information shared among the tag, server and reader will lose consistency, which makes the subsequent three-party authentication impossible. So far, the replay attack launched by the attacker has succeeded, and has led to the loss of privacy information consistency between communication entities, that is, the asynchronous attack has also succeeded.

Based on the above, the rationality and correctness of the protocol designed by Wu *et al.* are open to discussion in many aspects, and there are also security vulnerabilities. Based on this protocol framework, an improved three-party authentication protocol is proposed.

5 Three Party Authentication Protocol Design

The specific design idea of the proposed protocol will be elaborated in this section.

1) Relevant symbols of three-party authentication Protocol
The meanings and expressions of symbols related to the protocol will be given below.

DB represents back-end server;

T represents tag;

R represents reader;

K represents the shared key among DB, T, R;

K^{new} represents the shared secret keys among DB, T, R in the last round;

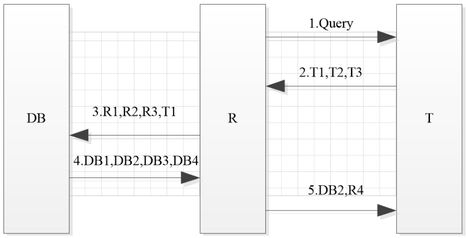


Figure 2: Schematic Diagram of Three Party Authentication Agreement

K^{old} represents the shared secret keys among DB, T, R in the current round;

TID represents the unique identifier representing of the tag;

RID represents the unique identifier representing of the reader/writer;

DBx indicates the random number generated by the backend server;

Tx indicates the random number generated by the tag;

Rx indicates the random number generated by the reader;

$DB1, DB2, DB3, DB4$ indicates the message calculated by the backend server;

$T1, T2, T3$ indicates the message calculated by the tag;

$R1, R2, R3$ indicates the message calculated by the reader;

$\oplus$ represents an OR operation;

& representation and operation;

$Riro(X)$ represent the remainder, rounding and reordering operations;

$Query$ indicates a request instruction;

$Count$ represents a counter at the tag end.

2) Three party authentication protocol steps

The schematic diagram of the three party authentication protocol designed in this paper is shown in Figure 2.

The specific implementation steps of the three-party authentication protocol designed in this paper can be described as follows.

Step 1: The reader sends a request instruction to the tag.

Step 2: The tag generates a random number Tx, calculates the message $T1, T2, T3$ in turn, and sends the message $T1, T2, T3$ to the reader.

$$\begin{aligned} T1 &= K \oplus Tx. \\ T2 &= Riro(K\&Tx). \\ T3 &= Riro((K \oplus TID)\&Tx). \end{aligned}$$

Step 3: The reader/writer verifies the authenticity of the tag through $T1, T2$. The specific verification method is as follows.

Deforms the message $T1 = K \oplus Tx$ and gets $Rx = RID \oplus R1$, then, according to the same operation rules to get $T2$. Compare the calculated result $T2$ with the received one $T2$, if the two are not equal, and the tag fails to pass the verification; if the two are equal, the tag passes the verification, the subsequent operation continues.

Then the reader generates random numbers Rx, calculates the messages $R1, R2, R3$ in turn, and finally sends the messages $R1, R2, R3$ to the server.

$$\begin{aligned} T2 &= Riro(K\&Tx) = Riro(K\&(K \oplus T1)); \\ R1 &= RID \oplus Rx; \\ R2 &= Riro(Rx\&(RID \oplus K)); \\ R3 &= Riro(T3\&(Rx \oplus Tx)). \end{aligned}$$

Step 4: The server verifies the reader/writer first, and then verifies the reader/writer and label at the same time.

The server verifies the reader/writer as follows.

Deforms the message $R1 = RID \oplus Rx$ and gets $Rx = RID \oplus R1$, then, according to the same operation rules to calculate $T2$. Combine the calculated $T3$ and use the same operation rules to calculate $R3$.

Compare the calculated result $R3$ with the received one $R3$, if the two are not equal, the reader/writer and the tag fail to pass the verification; if the two are equal, the reader/writer and the tag pass the verification, the subsequent operation continues.

Then the backend sever generates random numbers DBx, calculates the messages $DB1, DB2, DB3, DB4$ in turn, and finally sends the messages $DB1, DB2, DB3, DB4$ to the reader/writer.

There are two ways to update the shared key:

If K^{new} passes the verification, update the shared key.

$$\begin{aligned} K^{old} &= K^{new}; \\ K^{new} &= Riro((K^{new}\&DBx) \oplus K^{new}\&(Tx \oplus Rx)). \end{aligned}$$

If K^{old} passes the verification, update the shared key.

$$\begin{aligned}
K^{old} &= K^{old};\\
K^{new} &= Riro((K^{old}\&DBx) \oplus K^{old}\&(Tx \oplus Rx));\\
R2 &= Riro(Rx\&(RID \oplus K))\\
&= Riro((RID \oplus R1)\&(RID \oplus K));\\
T3 &= Riro((K \oplus TID)\&Tx)\\
&= Riro((K \oplus TID)\&(K \oplus T1));\\
R3 &= Riro(T3\&(Rx \oplus Tx));\\
DB1 &= DBx \oplus Rx;\\
DB2 &= DBx \oplus Tx;\\
DB3 &= Riro(K \oplus (RID\&DBx));\\
DB4 &= Riro((K \oplus DBx)\&TID).
\end{aligned}$$

Step 5: The reader verifies the server mainly through messages $DB1, DB3$. The specific verification is as follows.

The reader/writer deform the message $DB1 = DBx \oplus Rx$ and gets $DBx = Rx \oplus DB1$, then, according to the same calculation rules to get $T2$. Compare the calculated result $DB3$ Compare the calculated result $DB3$, if the two are not equal, the s $DB3$ erver fails to pass the verification; if the two are equal, the server passes the verification, the subsequent operation continues.

The reader/writer calculates the message $R4$, and then the reader/writer starts to update the shared key K. After the shared key is updated, the message $DB2, R4$ is sent to the tag.

$$\begin{aligned}
DB3 &= Riro(K \oplus (RID\&DBx))\\
&= Riro(K \oplus (RID\&(Rx \oplus DB1)));\\
R4 &= Riro((DB4\&Rx) \oplus (Tx\&DBx));\\
K &= Riro((K\&DBx) \oplus K\&(Tx \oplus Rx)).
\end{aligned}$$

Step 6: The tag checks the value of $Count$ corresponding to the received $R4$ first. If it is not 0, no operation will be performed; If it is 0, subsequent operations will be performed.

The tag starts to verify the reader/writer and server. The specific verification methods are as follows.

The tag deform the message $DB2 = DBx \oplus Tx$ and gets $DBx = Tx \oplus DB2$, then, according to the same calculation rules to get $DB4$. Combine the calculated $DB4$ and use the same operation rules to calculate $R4$.

Compare the calculated result $R4$ with the received one $R4$, if the two are not equal, the server and the reader/writer fail to pass the verification; if the two are equal, the server and reader/writer pass the verification, the subsequent operation continues.

Then the tag starts to update the shared key K. After the shared key is updated, the authentication among tag, reader and server is completed.

$$\begin{aligned}
DB4 &= Riro((K \oplus DBx)\&TID)\\
&= Riro((K \oplus (Tx \oplus DB2))\&TID);\\
R4 &= Riro((DB4\&Rx) \oplus (Tx\&DBx));\\
K &= Riro((K\&DBx) \oplus K\&(Tx \oplus Rx)).
\end{aligned}$$

6 Security Analysis of Three Party Authentication Protocol

1) Three party authentication
The protocol used in mobile RFID system must be able to achieve the authentication among tag, reader and server. The protocol in this paper can meet this requirement.

In the message sent to the reader/writer by the tag, the reader/writer will authenticate the tag through $T1, T2$.

In the message sent by the reader/writer to the server, the server will first authenticate the reader/writer through $R1, R2$; Then, the reader/writer and tag will be verification at the same time again.

In the message sent by the server to the reader/writer, the reader/writer will authenticate the server through $DB1, DB3$.

In the message sent to the tag by the reader/writer, the tag will authenticate the server through $DB2, R4$ and the reader/writer at the same time.

According to the above description, it can be found that the design protocol in this paper can achieve third-party authentication.

2) Counterfeit attack
The mobile RFID system has three session entities: tag, server, and reader. An attacker can impersonate any of these session entities. Here, the attacker can impersonate a reader for analysis.

The attacker impersonates a reader/writer and sends a message to the tag. Since the message sent is only a request instruction $Query$, the tag will not verify the reader/writer after receiving the message, directly perform subsequent operations. After the attacker receives the message $T1, T2, T3$ sent by the tag, the attacker attempts to analyze some private data from the message $T1, T2, T3$, but the attacker cannot succeed. The attacker lacks K such private shared data, which makes it impossible for the attacker to crack the data implied in the message. At the same time, the attacker cannot forge the correct message $R1, R2, R3$. After the server receives the message $R1, R2, R3$, the server can identify the attacker's fake reader/writer through simple calculation, the protocol stops.

3) Replay attack
When the attacker fails to impersonate one of the entities, the attacker attempts to obtain the session messages in multiple complete sessions through continuous eavesdropping, and attempts to send the previous session messages obtained through eavesdropping again in a future session, so as to achieve the goal of replay attack.

In the design process of the protocol in this paper, the security risks of the above attackers is full considered so all sent messages are added with random numbers in the calculation process, and the number of random numbers added to all messages is different, and there is no rule to follow. When the attacker sends the previous round of messages again, the random number used in the current session message calculation has already changed. Therefore, after the receiver receives the message replayed by the attacker, a simple calculation can identify the message forgery source, and the attacker's replay attack will naturally fail.

4) Asynchronous attack
The reason why the protocol designed by Wu *et al.* cannot resist asynchronous attacks is that the tag side does not count the messages sent by the reader writer, resulting in the tag side continuously updating the privacy sharing data until the information between entities loses consistency.

In this protocol, a counter is introduced at the label end to record the number of tag messages sent by the reader writer. If the number of times corresponding to a message is non-zero, no operation is performed; When and only when the number of times corresponding to a message is zero, the corresponding subsequent operations will be performed at the label end. Therefore, if an attacker replays a message, he or she cannot make one end of the tag continuously update the private shared data. If the private shared data is not continuously updated, the shared data between session entities is still consistent.

At the same time, the protocol in this paper stores the shared data used in multiple sessions on the server side. When the server cannot identify and verify with the current shared data, the server will enable the shared data used in previous sessions to identify and verify again, so as to restore the consistency of private shared data between entities.

5) Tracking attacks
The message sent by the tag is continuously eavesdropped, and then, the obtained message is analyzed to try to locate the geographic location of the tag. This attack is called tracking attack.

In order to ensure the security of the tag's geographic location, the protocol in this paper mixes random numbers when encrypting every message sent by the tag, sometimes a random number, sometimes two random numbers, and sometimes three random numbers. These random numbers are not all generated by the tag itself, some are generated by the reader, and some are generated by the server, Therefore, the attacker could not obtain the correct value of each random number at all. The mixing of random numbers makes the session message values used by two closely adjacent sessions different. Therefore, the geographic location of the tags analyzed by the attacker is constantly changing, and the tracking attack launched by the attacker cannot succeed naturally.

6) An exhaustive attack
By eavesdropping on multiple rounds complete sessions, attackers can obtain session messages in multiple rounds sessions, and attempt to exhaust the correct values of some private data through direct enumeration. This method is invalid for the proposed protocol in this paper. Here, take the message $T1, T2$ as an example for analysis.

Attackers can obtain messages $T1, T2$ by eavesdropping. First, the message $T1 = K \oplus Tx$ is deformed and calculated $Tx = K \oplus T1$ according to the same calculation rules. In $T2$, it seems that the attacker does not know only one parameter K. The attacker thinks that the correct value K can be exhausted through exhaustive methods, but the attacker do not know the specific implementation steps of the original bitwise composition encryption algorithm in this paper.

The bitwise composition operation involves the Hamming weight values corresponding to the three parameters $K, K \oplus T1, K\&(K \oplus T1)$. The attacker does not know the Hamming weight values corresponding to the three parameters. The attacker does not know how the three parameters will be combined. Therefore, the attacker fails to attack in an exhaustive manner.

The protocol in this paper is compared with other classical protocols in terms of security. The comparison results are shown in Table 1.

7 Performance Analysis of Three Party Authentication Protocol

In this section, we will analyze the performance from the aspects of the computation and storage at the tag end, and the computation and communication of a complete process. The analysis results are shown in Table 2.

The meanings of some symbols involved in Table 2 are as follows: AND indicating the calculation amount of and operation; indicates the calculation amount of OR operation; SHI indicates the calculation amount of the

Table 1: Security Comparison Between Protocols

Type of Attack	**Ref. [1]**	**Ref. [14]**	**Ref. [19]**	**This protocol**
Three party authentication	√	√	√	√
Counterfeit attack	√	√	√	√
Replay attack	√	×	×	√
Asynchronous attack	√	√	×	√
Tracking attacks	×	√	√	×
An exhaustive attack	√	√	√	√

Table 2: Performance comparison between protocols

Comparison protocol	**Ref. [1]**	**Ref. [14]**	**Ref. [19]**	**This protocol**
Calculation amount at label end	$8XOR + 5AND+$ $9SHI$	$5XOR + 6ECC$	$4XOR + 1PUF() + 1mod$ $+1hash + 3PRNG$	$2XOR + 5Bco()$
Label end storage	$2L$	$4L$	$3L$	$2L$
Overall calculation amount	$18XOR + 12AND$ $21SHI$	$14XOR + 12ECC$	$10XOR + 1PUF() + 7mod$ $+2hash + 10PRNG$	$9XOR + 17Bco()$
Overall traffic	$2L + 2$	$11L + 1$	$15L + 1$	$13L + 1$

shift operation; $Riro()$ represents the amount of computation for the remainder, rounding and reordering operations; $PUF()$ indicates the computation amount of physical non-cloneable functions; mod indicates the calculation amount of modular operation; $hash$ represents the computation amount of the hash function; $PRNG$ indicates the calculation amount of non-random functions; ECC indicates the calculation amount of the elliptic curve; L indicates the length of the message (the request instruction only needs 1 byte space).

In view of the limited space, we only select the text protocol as the object. The calculation amount, storage amount, overall calculation amount, and overall traffic data of the protocol label side are given below.

Calculation amount at the tag end: XOR is used for the first time in the $T1$ calculation process, $Riro()$ is used for the first time in the $T2$ calculation process, $Riro()$ is used for the second time in the $T3$ calculation process, XOR is used for the second time in the $DB2$ deformation process, $Riro()$ is used for the third time in the $DB4$ calculation process, $Riro()$ is used for the fourth time in the $R4$ calculation process, and $Riro()$ is used for the fifth time in the update K calculation process. Based on the above, the calculation amount at tag end is $2XOR + 5Riro()$.

The storage capacity of the tag side: the shared key K and the identifier TID of the tag itself need to be stored, so the storage capacity is $2L$.

The calculation amount at the reader/writer end: XOR is used for the first time in the $T1$ calculation process, $Riro()$ is used for the first time in the $T2$ calculation process, XOR is used for the second time in the $R1$ calculation process, $Riro()$ is used for the second time in the $R2$ deformation process, $Riro()$ is used for the third time in the $R3$ calculation process, $Riro()$ is used for the fourth time in the $DB3$ calculation process, $Riro()$ is used for the fifth time in the $R4$ calculation process, and $Riro()$ is used for the sixth time in the update K calculation process. Based on the above, the calculation amount at the reader/writer end is $3XOR + 6Riro()$.

The calculation amount on the server side:

XOR is used for the first time in the $R1$ deformation process, $Riro()$ is used for the first time in the $R2$ calculation process, XOR is used for the second time in the $T1$ calculation process, $Riro()$ is used for the second time in the $T3$ calculation process, $Riro()$ is used for the third time in the $R3$ calculation process, $Riro()$ is used for the fourth time in the $DB3$ calculation process, $Riro()$ is used for the fifth time in the $DB4$ calculation process, and $Riro()$ is used for the sixth time in the update K calculation process.

Based on the above, the calculation amount on the server side is $4XOR + 6Riro()$.

Overall calculation amount: the overall calculation amount=the calculation amount on the tag side + the calculation amount on the reader/writer side + the calculation amount on the server side + the overall calculation amount is $9XOR + 17Riro()$.

In a complete communication process, the interactive messages are as follows: The data sent by the reader/writer are: 1 byte $Query$, one $R1$, one $R2$, one $R3$, one $T1$, one $DB2$, one $R4$, the total calculation amount is $6L+1$. The data sent by the tag are: one $T1$, one $T2$, one $T3$, the total calculation amount is $3L$. The data sent by the server are: one $DB1$, one $DB2$, one $DB3$, one $DB4$, the total calculation amount is $4L$.

Overall traffic: overall computation=data sent by tags + data sent by readers + data sent by servers, overall computation is $13L + 1$.

Compared with other protocols, the proposed protocol has significantly less computation and overall computation on the tag side. The main reason is that the encryption algorithm used in this protocol is an original super lightweight computation method based on bit operation, while other protocols use hash function encryption or elliptic ECC algorithm encryption or physical unclonable function encryption or pseudo-random function encryption, which makes it impossible to achieve the super lightweight level. In terms of the storage capacity and overall communication volume of the tag side, the proposed protocol is roughly equivalent to other protocols. Through comprehensive comparison and analysis, it is showed that the proposed protocol is superior to other protocols in terms of computational complexity, the proposed protocol can make up for the security problems existing in other protocols, indicating that the proposed protocol has the conditions for promotion and use.

8 Conclusion

By analyzing the difficulties faced by RFID system in application, a super lightweight three-party authentication protocol for mobile RFID systems is proposed. The main innovation of the protocol is to adopt a novel encryption algorithm, namely, the remainder, rounding and reordering operations. The proposed protocol skillfully combines the length variable of the encryption parameter itself and the Hamming weight variable, carries out the remainder rounding operation between them, and then, according to the result, performs reordering operations in the different ways, finally obtain the encryption result. The proposed protocol can increase the difficulty of cracking while reducing the introduction of parameters. In order to resist the replay attack and asynchronous attack initiated by the attacker, the proposed protocol introduces a message counter at the tag end to count the message number sent to the tag by the reader/writer. When the message counter is only 0, the tag will carry out subsequent operations, so as to prevent the attacker from replaying the message continuously, causing the tag end to constantly update information, resulting in the loss of information consistency between different entities. From the perspective of multiple attack types, the security of the proposed protocol is analyzed, it is shown that the protocol can resist multiple attacks. From different computing aspects, it is shown that the protocol is superior to other protocols in computational efficiency.

References

[1] S. F. Aghili, H. Mala, and P. Kaliyar, “Seclap: secure and light weight RFID authentication protocol for medical IoT,” *Future Generation Computer Systems*, vol. 101, pp. 621–634, 2019.

[2] Z. Cao and O. Markowitch, “Analysis of shim’s attacks against some certificateless signature schemes,” *International Journal of Network Security*, vol. 23, no. 3, pp. 545–548, 2021.

[3] Y. C. Chen, W. L. Wang, M. S. Hwang, “RFID authentication protocol for anti-counterfeiting and privacy protection”, in *The 9th International Conference on Advanced Communication Technology*, pp. 255-259, 2007.

[4] S. Y. Chiou and S. Y. Chang, “An enhanced authentication scheme in mobile RFID system,” *Ad Hoc Networks*, vol. 71, pp. 1–13, 2017.

[5] J. Chong and Z. Zhuo, “Constructions of balanced quaternary sequences of even length,” *International Journal of Network Security*, vol. 22, no. 6, pp. 911–915, 2020.

[6] R. Doss, S. Sundaresan, and W. Zhou, “A practical quadratic residues based scheme for authentication and privacy in mobile RFID systems,” *Ad Hoc Networks*, vol. 11, no. 1, pp. 383–396, 2013.

[7] Y. P. Duan, “Lightweight RFID group tag generation protocol,” *Control Engineering of China*, vol. 27, no. 4, pp. 751–757, 2020.

[8] K. Fan, W. Jiang, and H. Li, “Lightweight RFID protocol for medical privacy protection in iot,” *IEEE Transactions on Industrial Informatics*, vol. 14, no. 4, pp. 1656–1665, 2018.

[9] P. Gope, J. Lee, and T. Q. S. Quek, “Lightweight and practical anonymous authentication protocol for RFID systems using physically unclonable functions,” *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 11, pp. 2831–2843, 2018.

[10] Y. Y. Hsieh, L. H. Chang, and A. Y. H. Liao, “The system adoption evaluation of RFID safety management system on campus,” *International Journal of Network Security*, vol. 24, no. 1, pp. 176–180, 2022.

[11] Q. Jiang, Z. R. Chen, And B. Y. Li, “Security analysis and improvement of bio-hashing based three-factor authentication scheme for telecare medical information systems,” *Journal of Ambient Intelligence and Humanized Computing*, vol. 9, no. 4, pp. 1061–1073, 2018.

[12] W. Liang, S. Xie, And J. Long, “A double puf-based RFID identity authentication protocol in service-centric internet of things environments,” *Information Sciences*, vol. 50, no. 3, p. 129–147, 2019.

[13] D. W. Liu And J. Ling, “An improved RFID authentication protocol with backward privacy,” *Computer Science*, vol. 43, no. 8, pp. 128–130, 2016.

[14] S. Rostampour, M. Safkhani, and Y. Bendavid, “Ecbap: a secure ECC based authentication protocol for IoT edge devices,” *Pervasive and Mobile Computing*, vol. 67, pp. 101–194, 2020.

[15] M. Safkhani, S. Rostampour, and D. Y. Ben, "IoT in medical & pharmaceutical: designing lightweight rfid security protocols for ensuring supply chain integrity," *Computer Networks*, vol. 181, pp. 107–558, 2020.

[16] G. F. Shen, S. M. Gu, , and D. W. Liu, "An anti-counterfeit complete RFID tag grouping proof generation protocol," *International Journal of Network Security*, vol. 9, no. 4, pp. 889–896, 2018.

[17] M. Y. Wang, X. Zhang, W. J. Li, *et al.*, "Review of research on privacy protection technology for data publication," *Journal of Chinese Computer Systems*, vol. 41, no. 12, pp. 2657-2667, 2020.

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

[19] K. F. Wu And X. C. Yin, "Lightweight mutual authentication protocol between three communication agents," *Journal of Chinese Computer Systems*, 2022.

[20] R. Xie, J. Ling, and D. W. Liu, "A wireless key generation algorithm for RFID system based on bit operation," *International Journal of Network Security*, vol. 20, no. 5, pp. 938–950, 2018.

[21] X. Zhao, "Attack-defense game model: Research on dynamic defense mechanism of network security," *International Journal of Network Security*, vol. 22, no. 2020, pp. 1037–1042, 2018.

[22] F. Zhu, P. Li, H. Xu, and *et al.*, "A lightweight rfid mutual authentication protocol with puf," *Sensors*, vol. 19, no. 13, p. 2957–2978, 2019.

Biography

Sheng-hua Xu received a master's degree in School of Computers from Guangdong University of Technology (China) in June 2009. He is now a lecturer, working in Guangdong Polytechnic Normal University. At present, his research interests mainly include information security.

Spatial Crowdsourcing Optimal Task Allocation Algorithm Based on Tree Decomposition

Hui Xia[1], Shufeng Zhang[2], and Weiji Yang[3]

(Corresponding author: Shufeng Zhang)

[#] *Hui Xia and Shufeng Zhang contribute equally to the article.*

Shenyang Normal University, Shenyang 110034, China[1]

Suzhou Industrial Park Institute of Service Outsourcing, Suzhou 215123, Jiangsu, China[2]

Zhejiang Chinese Medical University, HangZhou 310000, China[3]

Email: zhangshuf@siso.edu.cn

(Received Nov. 7, 2022; Revised and Accepted Sept. 12, 2023; First Online Dec. 24, 2023)

Abstract

With the proliferation of mobile devices equipped with hi-Fi sensors and the rapid decline in wireless network rates, spatial crowdsourcing is being used as a problem-solving framework to solve the problem of assigning location-related tasks (such as road condition reporting food delivery) to workers (people equipped with smart devices and willing to perform the task). The key to studying the optimal task allocation problem in spatial crowdsourcing is to design a task allocation strategy that assigns each task to the most appropriate workers to maximize the total number of tasks completed so all workers can return to the starting point before the expected latest working time after completing the assigned tasks. Finding the global optimal allocation is a tricky problem because the problem does not amount to a simple summation of the optimal allocation for a single worker. It is noted that only part of the workers have task dependence, so the workers are divided into independent sets by using tree decomposition technology, and a depth-first search algorithm with a heuristic is proposed. The algorithm can quickly update the heuristic function boundary to efficiently prune the assignment scheme that is unlikely to be the optimal solution as soon as possible. Experimental results show that the proposed method is very effective and can solve the problem of optimal task allocation.

Keywords: Optimal Method; Spatial Crowdsourcing; Task Allocation; Task Dependency; Tree Decomposition

1 Introduction

With the rapid popularity of smart devices and the rapid decline of mobile charges, people carry a variety of sensors anytime and anywhere to participate in a variety of location-related activities, such as scenic spot photo collection, road condition monitoring and so on. In this context, the concept of spatial crowdsourcing [12] emerged. Spatial crowdsourcing requires participants (often called workers) to actually drive to a given location in order to complete a given task (such as taking photos of a scenic spot).

This paper studies the optimal task allocation scheme in spatial crowdsourcing based on the above scenarios. Specifically, given the location and expected latest working time of each worker, as well as the location and expiration time of each task, the optimal task allocation scheme is found to maximize the global task allocation quantity.

Compared to existing work, the main difficulty of this problem is that once it is necessary to consider the time cost for workers to travel to the task and the expiration time of the task, local optima may not necessarily produce a global optimal solution. The second challenge of this paper is that the range of tasks accessible to a worker is highly dependent on the worker's starting position and the worker's expected latest working hours. This makes it impossible to exclude unreachable tasks by setting a fixed range of work areas or the maximum number of tasks that can be accepted [1, 3, 4, 14].

To solve this problem, an exact solution algorithm is proposed in this paper, which is used to find the optimal allocation scheme that maximizes the number of global task allocation. The main idea of this paper is as follows: according to the task dependency relationship (if two workers can complete a certain task, there is task dependency between them), workers are divided into mutually independent workers by tree decomposition [2]. Then, the set of workers is indexed as a node into a search tree structure. Finally, a heuristic depth-first traversal algorithm is used to search for the optimal solution.

2 Related Work

In the present work [19], it has been proved that the global optimal task allocation scheme in space crowdsourcing is a difficult problem for NP. Therefore, the simplest method is to use greedy algorithm to find the maximum set of effective tasks for workers in turn, and then accumulate the number of assigned tasks. The main problem with this method is that for tasks that can be completed by more than one worker, simply assigning one worker at random may cause other tasks to be unallocated.

2.1 Computing Effective Task Sets

2.1.1 Finding Reachable Tasks

Restricted by the latest working hours and the expiration time of tasks, each worker can only accomplish a small number of tasks. Therefore, first of all, we should find out the set of tasks that each worker can reach without violating the constraints. The set of tasks that workers can reach should be recorded as meeting the following two conditions.

$$\forall s \in RS_w, c(w.l, s.l) \le s.e; \quad (1)$$

$$c(w.l, s.l, w.l) = c(w.l, s.l) + c(s.l, w.l) \le w.t. \quad (2)$$

$C(w.l, s.l, w.l)$ is the time for workers to return to $w.l$ from $w.l$ via $s.l$. These two conditions ensure that a worker can get from his starting point to the location of the task before it expires, while leaving enough time to return to his starting point before the worker's latest working time.

2.1.2 Searching for Maximum Effective Task Sets

When searching for the optimal solution, in order to speed up the search efficiency, this paper hopes that a worker can be assigned multiple simultaneous reachable tasks at one time, instead of one task at a time, and then judge whether the newly assigned tasks and the assigned tasks are reachable at the same time. Therefore, it is necessary to estimate the maximum effective task set of each worker. Given the reachable tasks of each worker. Sets, it can be proved that: find out the maximum effective task set of each worker (MVTS) is a NP-hard problem, and the proof process is similar to that in document [18]. However, because each worker's achievable task set is usually small, this means that this problem can be solved by efficient algorithms. Moreover, the calculation of MVTS for each worker is completely independent, so it can be calculated in parallel.

Next, this paper introduces the state transition equation of the dynamic programming algorithm to solve the maximal effective task set. By gradually increasing the size of the reachable task set, the algorithm constantly extends the worker's reachable task set, and finds out all the MVTS under the set in each iteration. Given a worker w and a set of tasks $Q \subset RS_w$, this paper defines $opt(Q, s)$ as a task after passing through the Q set. The maximum number of tasks that can be completed at the location where the task $s.l$ is located. R is the task scheduling sequence in the Q set. Representation of s_j in sequence R by s_i. The previous task, and R' denote the corresponding sequence of tasks for $opt(Q - \{s_j\}, s_i)$. $Opt(Q, s_j)$ can be calculated from Formula (3):

$$opt(Q, s_j) = \begin{cases} 1 & \text{if } |Q| = 1 \\ \max_{s_j \in Q, s_i \neq s_j} & \\ opt(Q - \{s_j\}, s_i) + \sigma_{i,j} & \text{otherwise} \end{cases} \quad (3)$$

$$\sigma_{i,j} = \begin{cases} 1 & \text{if } t(s_j.l) \le s_j.e, t(s_j.l) \\ & +c(s_j.l, w.l) \le w.t \\ 0 & \text{otherwise} \end{cases}$$

$\sigma_{i,j}$ indicates that the task s_i can still be completed at the end of adding the task to the sequence R', and that the worker s_j can return to the starting point before the latest expected time.

When Q only includes one task s_i, the problem is very simple. When $opt(s_i, s_j) = 1$ and $|Q| > 1$, it is necessary to search Q to check all possibilities of the effective task set s_i and find them so as to maximize $opt(Q, s_j)$. The time complexity of the MVTS set Q_w in Formula (3) is $O(n^3 \cdot 2^n)$, while the space complexity is $O(n \cdot 2^n)$.

2.2 Segmentation of Worker Set

The main challenge in finding the optimal solution lies in the large search space. When enumerating all possible effective task sets of all workers, the time complexity increase exponentially with the increase of the number of workers.

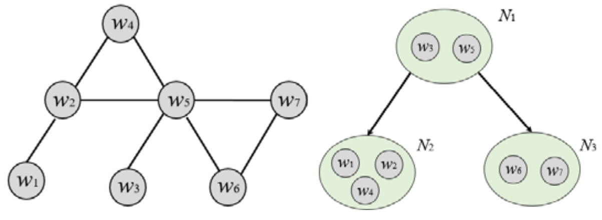
Definition 1. *(Task Dependence). Given two workers, WJ and their respective achievable task sets RS_w, if RS_w and RS_w mutual independence; otherwise, there is a task dependence between them.*

For example, in Figure 1, workers are only dependent on and exist task dependence, but not related to other workers. Obviously, if there is no task dependence among all workers, the optimal task allocation scheme only needs to simply add up the optimal scheme of each worker. Therefore, in order to reduce the computational cost of finding the optimal solution, it is necessary to use task dependence as much as possible to divide workers into unrelated sets. In each independent set, the optimal allocation scheme in the set is found.

2.2.1 Graph Decomposition

Definition 2. *(Task Dependency). Given two workers w_i, w_j and their reachable task sets RS_{w_i}, RS_{w_j}, if $RS_{w_i} \cap RS_{w_j} = \phi$, they are independent of each other; Otherwise, there is task dependency between them.*

Figure 1(a) shows the worker dependency diagram in the example.



(a) (b)

Figure 1: Worker dependency graph partition. (a)Worker Dependency graph, (b) Constructed search tree.

3 Problem Statement and Preliminaries

Definition 3. *(Space task). A spatial task can be defined by a binary group $s = < s.e, s.l >$. Where, $s.l$ is the position of the task, $s.e$ is the expiration time of the task, $s.l$ is a point (x, y) in a two-dimensional plane space. In spatial crowdsourcing, the task can be completed only when the worker actually reaches the specified position $s.l$ of task s. At the same time, considering the expiration time of the task. The worker can complete the task only when he arrives at the $s.l$ location before the task s expiration time $s.e$. Under the non redundant task allocation mode considered in this paper, the server will only assign one task to a single worker. Like the previous work [8,13], this paper assumes that the worker will immediately go to the next task after completing one task, and the time required to complete the task itself is negligible.*

Definition 4. *(worker). Workers generally refer to people who carry mobile equipment and voluntarily complete space tasks Workers can be represented by the binary group $s = <w.l, w.t>$, where wL is the worker's starting position, wT is the expected latest working time of the worker, and the worker needs to return to the starting position no later than the latest working time.*

The worker's working mode is divided into online mode and offline mode. When the worker is in online mode, he can accept space tasks. Once the worker is in online mode, he will send a task request to the server. The request contains the worker's location $w.l$. The latest expected working time of the worker $w.t$. The server will consider all workers and tasks acquired in a very short time interval at the same time, and then make the global optimal task allocation. Finally, the task sequence assigned to each worker is returned.

Definition 5. *(task sequence) Given any worker w and the task set s_w assigned to w, the task sequence of s_w is expressed as $R(s_w)$, representing the time sequence in which workers access each task. The time $t_{w,R}(s_i.l)$ when workers arrive at each task can be defined by Formula (1):*

$$t_{w,R}(s_i.l) = \begin{cases} t(s_{i-1}.l) + c(s_{i-1}.l, s_i.l) & \text{if } i \neq 1 \\ c(w.l, s_1.l) & \text{if } l = 1 \end{cases} \quad (4)$$

In Formula (4), $C(a, b)$ represents the travel time from task a to task b. Under the condition that w and R are not ambiguous, $t(s.l)$ is used to represent the time when worker w arrives at task s_i's location $s.l$, and $t(w.l)$ is used to represent the time when worker w returns to its original location. The time when a worker returns to the starting point after completing all tasks in the task set S_w is defined as:

$$t(w.l) = t(s_{|S|}.l) + c(s_{|S|}.l, w.l) \quad (5)$$

Because the driving speed is not within the scope of the main factors considered in this paper, for simplicity, this paper assumes that all workers have the same driving speed. Therefore, the travel time cost of workers can be defined by Euclidean distance at two locations. However, the method proposed in this paper does not rely on this assumption, and can be used when workers have different speeds.

Definition 6. *(valid task set VTS) When and only when the following conditions are true, the task set S_w is called the effective task set of worker w:*

1) *All tasks in Sw can be completed before their expiration time, that is, $\forall S_i \in S_w$, $t(S_i.l) \leq S_i e$;*

2) *Worker w can return to the starting point no later than the latest working time after completing all tasks in S_w, that is, $t(w.l) \leq w.t$.*

Definition 7. *(MVTS) If any superset of the effective task set S_w is not the effective task set, then S_w is called the maximum effective task set.*

Definition 8. *(Space Task Assignment) Given worker set W and task set S, spatial task assignment A consists of a series of workers, VTS tuples, such as $A = \{< w_1, VTS(w_1) >, < w_2, VTS(w_2) >, \cdots, < w_{|S|}, VTS(w_{|S|}) >\}$. Let $A.S$ represent the task set assigned to all workers, that is, $A.S = \cup_{w \in W} S_w$.*

The problems raised in this paper are summarized as follows:

Problem definition. Given a worker set W, a task set S, and a space crowdsourcing task allocation problem with the latest working time constraints, the goal is to find a global optimal task allocation scheme A_{opt}, so that $\forall A_i \in A, |A_i.S| \leq |A_{opt}.S|$, where A represents all space task allocation schemes.

4 Algorithm Analysis

4.1 Introductions of Algorithm

The purpose of graph decomposition is to divide irrelevant workers into different sets, and at the same time ensure that the used partition set (hereinafter referred to as cut set) can divide the connected graph as evenly as possible (see Algorithm 1 for decomposition algorithm). In each step of decomposition, take a point and connect to it as a cut set (row 7), and try to use the cut set to cut the connected graph (rows 8 and 9), and record the size of the cut set (the cut set contains multiple workers) and the sum of the number of workers which is in the maximum connected subgraph after the graph is cut (row 10). In each step, select the smallest cut set (lines 11 14) to segment the original graph, and recursively call the algorithm on the connected subgraph after segmentation until the number of workers in the connected subgraph is less than the threshold (lines 3 5).

Algorithm 1 Tree decomposition algorithm of graph

```
 1: Input: worker dependency graph $G$, sequence $C$ of point cut set (worker set), sequence $I$ generated by point cut set (initialized to 0, representing the sequence number of the first point cut set), and global variable index, representing the current point cut set sequence number $cur$;
 2: Output: point cut set sequence $C$ and sequence $I$ generated by each point cut set.
 3: $TreeDecomposition(G, C, I, index, cur)$;
 4: Best_h = infinity, Best_S=empty set;
 5: if $|G| < threshold$ then
 6:     $C[cur] \leftarrow$ nodes in $G$;
 7: end if
 8: for each node $w_i \in G$ do
 9:     $Set(V_i) \leftarrow$ node connected to $w_i$;
10:     $Set(V_i) \leftarrow$ node not connected to $w_i$;
11:     $G' \leftarrow$ build graph of $Set(V_i)$;
12:     $h \leftarrow |Set(V_i)| + \max |SubGraph(G')|$;
13:     if $h < Best\_h$ then
14:         $Best\_h \leftarrow h$;
15:         $Best\_S \leftarrow Set(V_i)$;
16:     end if
17: end for
18: $C[cur] \leftarrow Best\_S$;
19: $G'' \leftarrow$ build graph for node not in Best_S;
20: for each sub graph $G_s \in G''$ do
21:     Sub_index = ++index;
22:     $I[cur] \leftarrow$ Sub_index;
23:     $TreeDecomposition(G_s, C, index, Sub\_index)$;
24: end for
```

Because the optimal graph decomposition problem itself is NP hard, this scheme does not solve the optimal cut set in each iteration of tree decomposition, but attempts to take each node in the dependency graph and its surrounding nodes as the cut set in turn, and records the best cut set that has been found (lines 7 to 15). As shown in Figure 2, when w_3 and w_5 connected with w_3 are used as cut sets, other workers are divided into two relatively average groups, and the resulting point cut set sequence C is $[\{w_3, w_5\}, \{w_1, w_2, w_4\}, \{w_6, w_7\}]$. At this time, the cut set size is 2, w_1, w_2, and w_4, and the maximum size of the connected subgraph is 3, and the sum of the two is 5. It can be found that the sum of the cut set size and the maximum size of the connected subgraph is not smaller than 5 in other segmentation methods. Therefore, Algorithm 1 will preferentially use w_3 and w_5 as cut sets.

Given the worker set W and the task set S, we first calculate the maximum task subsequence Q_{W_i} (the second and third lines) for each worker W_i, and S_i is the reachable task set, and then establish the corresponding task dependency graph G (the fourth line). For each connected subgraph $g \in G$ in the task dependency graph, use the tree decomposition algorithm to divide the workers into different worker sets (row 6), and then build the sequence of worker sets into a search tree structure (row 7). Finally, use the depth first search algorithm to find the optimal task allocation scheme for the search tree established in the previous step. The initial heuristic function value can be calculated by the greedy algorithm. Because different connected subgraphs of G are not related to each other, the final task allocation scheme only needs to add up the schemes of different connected subgraphs (line 8).

Algorithm 2 Search framework

```
 1: Input: worker set $W$, task set $S$;
 2: Output: Optimal task allocation scheme $Opt$
 3: $Solve(W, S)$
 4: for each $w_i \in W$ do
 5:     $Q_w = MVTS(w_i, S_i)$;
 6: end for
 7: $G \leftarrow$ build $WDG$;
 8: for each connected component $g \in G$ do
 9:     $X_g \leftarrow$ TreeDecomposition() of $g$;
10:     $N_g \leftarrow$ build search tree;
11:     $Opt \leftarrow Opt + DFSearch(N_g, S, W_N, LB(N_g))$;
12: end for
13: Return $Opt$;
```

The depth first algorithm is described in detail below 3. The four key parameters of the search process are as follows: (1) the root node of the N^{th} subtree; (2) Unassigned task set S; (3) The worker set W_N that has not been searched in the N^{th} subtree; (4) The heuristic function value h used to prune the search space represents the minimum number of tasks to be allocated without pruning the subtree.

Algorithm 3 will recursively call itself to search for all task allocation schemes, and will exclude the allocation schemes that will not become the optimal solution. Therefore, when the algorithm exits, the algorithm can obtain the allocation scheme that maximizes the number of global tasks allocated.

4.2 Upper Bound of Estimation

The upper bound of the number of tasks that can be completed by node N is recorded as $UB(N)$, which represents the maximum number of tasks that can be completed by all workers in the subtree with node N as the root. The basic method for estimating the upper bound is to accumulate the maximum effective task set of all workers in the subtree with node N as the root, and the maximum

Algorithm 3 Heuristic search algorithm

```
Input: the current node serial number $N$, the unassigned task set $S$, the workers set $W_N$ that has not been searched in the $N^{th}$ subtree, and the heuristic function value h used to cut the search space;
Output: Optimal task allocation scheme $Opt$
$DFSearch(N, S, W_N, h)$
$Opt \leftarrow 0$;
$UB(N) \leftarrow$ Calculate upper bound of sub tree rooted at $N$;
if $UB(N) < h$ then
    return 0;
end if
if $W_N \neq \phi$ then
    for each worker $w_i \in W_N$ do
        for each MVT set $Q \in MVTS(w_i, S)$ do
            $Opt \leftarrow \max\{DFSearch(N, S-Q, W_N-w, h-|Q|)+|Q|, Opt\}$
            $h \leftarrow Opt$;
        end for
    end for
else
    for each child node $N_i$ of $N$ do
        $Opt+ = DFSearch(N_i, S, W_N, h)$;
    end for
end if
Return $Opt$;
```

effective task set with the largest number of workers. The calculation formula is as follows:

$$UB(N) = \sum_{i=1}^{|W|}(|\max R_{w_i}|) \tag{6}$$

W in the formula represents all workers of the current subtree; max R_{w_i} represents the maximum effective task set of worker w_i, and the set with the largest number of set elements. For example, when the search algorithm starts to search in Figure 3, $UB(N_3) = |\max R_6| + |\max R_7| = 1 + 3 = 4$, and the value of $|\max R|$ can be obtained by looking up Table 2.

For all task allocation schemes A (including the optimal task allocation scheme), the number of tasks that each worker can complete will not exceed $|\max R|$, so the following inequality is true:

$$|A.S| = |\cup_{w\in W} S_w| \leq \sum_{w\in W} |S_w| \leq \sum_{w\in W} |\max R_w| = UB(N).$$

4.3 Heuristic Function

In order to prune the scheme that is not expected to be the optimal solution as soon as possible, the algorithm will calculate the heuristic lower bound h and transfer it to the recursive function as a parameter H indicates the minimum number of tasks to be completed by the subtree with node N as the root. Only when it is not less than h, it is possible to produce a more promising solution than the currently searched optimal case. Easy to get: When the upper bound (the maximum number of tasks that can be allocated) of a subtree is less than the lower bound, you can safely exclude this scheme.

The following describes how to estimate the heuristic function value of the subtree with node N as the root. Assuming that node N contains m sub nodes, such as $N_1, N_2, \cdots, N_m$, the depth first search algorithm will be used to search each sub tree in turn to find a better solution than the currently found optimal solution Opt. The heuristic function value is updated by the following formula:

$$h' = h - \sum_{j=1}^{i-1} Opt(N_j) - \sum_{j=i+1}^{m} UB(N_j) \tag{7}$$

In Formula (5), h represents the minimum number of tasks to be allocated to all sub nodes of node $\sum_{j=1}^{i-1} Opt(N_j)$ represents the sum of the maximum number of tasks that can be allocated to the traversed subtree, and $\sum_{j=i+1}^{m} Opt(UB(N_j))$ represents the sum of the estimated upper bounds of the unserved subtree. That is, the number of tasks that the current subtree N_i needs to complete at least is equal to the heuristic function value h. Subtract the subtree $(N_1 - N_{i-1})$ that has been searched before the current subtree to determine the number of tasks that can be completed, and then subtract the estimated upper bound $UB(N)$ of the subtree $(N_{i+1} - N_m)$ that has not been searched after the current subtree.

4.4 Optimization Strategy

Three optimization strategies are introduced here to further reduce the search cost:

1) $UB(V)$ optimization: before calling the depth first search algorithm, start from the leaf node of the search tree, from the bottom to the top, and use the greedy algorithm to find the upper bound of the tasks that can be assigned to the subtree with each node N_i as the root (each worker only takes the maximum effective task set, regardless of the task being assigned), which can limit the heuristic function to a smaller extent;

2) Single worker search optimization: sort the maximum effective task set ($|MVTS|$) of each worker from large to small, and search the set with large base first. Because if the search for the larger effective task set that is more likely to produce the optimal solution is completed, the smaller effective task set is easier to prune;

3) Search sub tree order optimization: all sub nodes of a node are sorted from small to large according to the number of workers contained in the sub tree whose sub node is the root. Because small subtrees can be searched quickly, the heuristic function value h can be updated quickly, which will become more compact and have better pruning effect for large subtrees.

5 Experiments and Results

5.1 Experimental Setup

Because of the lack of benchmark experimental data in the space crowdsourcing field, this paper uses the simulation data set to generate experimental data. The rules are as follows:

First, use uniform distribution on a two-dimensional plane of 100 × 100, and randomly generate 100 point coordinates to represent the position of workers; Then, change the average number of tasks (T/W) of each worker, randomly generate points representing tasks around each worker, and the value range of T/W is [11, 12] (experimental verification: when conducting experiments on a larger scale, the depth of search usually exceeds 20, and the search time increases exponentially, exceeding the allowable range of the experimental machine configuration);

Secondly, given any worker and the tasks around him, the expiration time of the task is defined as follows: starting from the current position of the worker, the greedy algorithm is used to select the tasks closest to the current position of the worker in turn. For the task sequence calculated by the greedy algorithm, the total travel time t is calculated and used as the upper bound of the task expiration time; Then define a range $[e^l, e^u](0 < e^l < e^u < 1)$. The expiration time of tasks is defined as $[e^l.t, e^u.t]$. This paper uses five groups of task expiration ranges [0.2, 0.3], [0.3, 0.4], [0.4, 0.5], [0.5, 0.6] and [0.6, 0.7];

Finally, the expected latest working time of workers is defined as follows: the distance t_w between the last task and the worker's starting point is obtained by adding the worker's traveling time t calculated by the greedy algorithm as the upper bound of the expected latest working time of workers; Then define a range $[d^s, d^t](0 < d^s < d^t < 1)$. The expected latest working time of workers is defined as $[d^s.t_w, d^t.t_w]$. This paper uses five groups of this range [0.2, 0.3], [0.3, 0.4], [0.4, 0.5], [0.5, 0.6], [0.6, 0.7].

Basically, the task expiration time range and the expected latest working time range of workers determine the percentage of tasks that workers can complete. For the results of each parameter change, 50 groups of experiments were carried out. The reported results are the average results of 50 groups of experiments. All experiments were conducted on a machine with Core i5-2400, 3.1G HZ CPU and 8GB RAM.

5.2 Experimental Result

This paper evaluates the performance of worker partitioning stage and the impact of task partitioning results on search, and compares random tree construction algorithm (RTA) with balanced tree construction algorithm (BTA) proposed in this paper RTA algorithm randomly selects a worker cluster as the current node of the search tree. BTA algorithm always selects a better worker cluster as the current node when constructing the current search node. This section compares two dimensions: (1) Search depth: search the maximum number of workers enumerated from the root node to the leaf node using depth first; (2) Search time: CPU time spent searching the optimal allocation scheme using the constructed search tree.

Figure 3 shows the influence of the average number of workers' tasks T/W, the task expiration time coefficient e^l, and the worker's latest working time coefficient d^s on the search depth Figure 4 shows the effect of the above parameters on search time. As shown in Figure 2(a), the search depth of the two tree construction algorithms increases with the increase of T/W. However, the BTA algorithm can produce a more balanced tree, which makes it more efficient than RTA, as shown in Figure 4(a).

As shown in Figure 2(b): when the task expiration time coefficient is small, the number of reachable tasks for each worker is also small, and the BTA algorithm and RTA algorithm have no difference in the ability to construct a search tree; However, with the increase of $[e^l, e^u]$, the number of accessible tasks for workers also increases rapidly, and the advantages of BTA algorithm become more obvious. As shown in Figure 3(b), although the search time increases exponentially with the increase of task expiration time coefficient, the performance of BTA algorithm is an order of magnitude higher than that of RTA.

Based on the search tree constructed by BTA algorithm, this section compares the performance of three different search algorithms: (1) depth first search (DFS) without any optimization; (2) Use sorting based search to optimize DFS+W (depth first search+worker sort) (the effective task set of a single worker is in the order of large to small and the number of workers in the subtree is in the order of small to large); (3) On the basis of (2), the algorithm DFS+W&U (depth first search+worker sort and upper bound estimate) is added. For the final task allocation quantity, this paper compares the task allocation quantity of the algorithm proposed in this paper with that of two basic methods, namely, greedy algorithm (GA for short) and iterative greedy algorithm (IGA for short) [5, 7, 9, 10]. GA calculates the maximum number of tasks that can be assigned to the worker inthe unassigned task set for each worker in turn until all workers have been assigned or the task set to be assigned is empty. IGA performs task allocation and task scheduling iteratively until no better solution can be found within 1000 steps, and finally selects the best allocation as the result.

Figure 4 shows the efficiency of different search algorithms It can be seen from the figure that the number of reachable tasks for each worker is very small, and the task optimization strategy will not achieve significant effect, no matter the average number of tasks, the task expiration time coefficient and the latest working time coefficient of workers are small However, with the increase of the

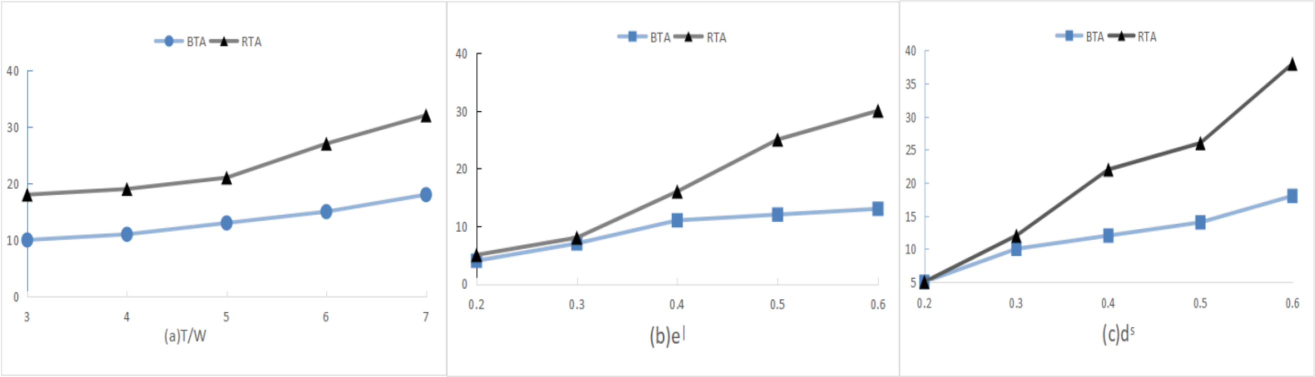


Figure 2: Effect of T/W, e^l and d^s on the depth of constructed search trees

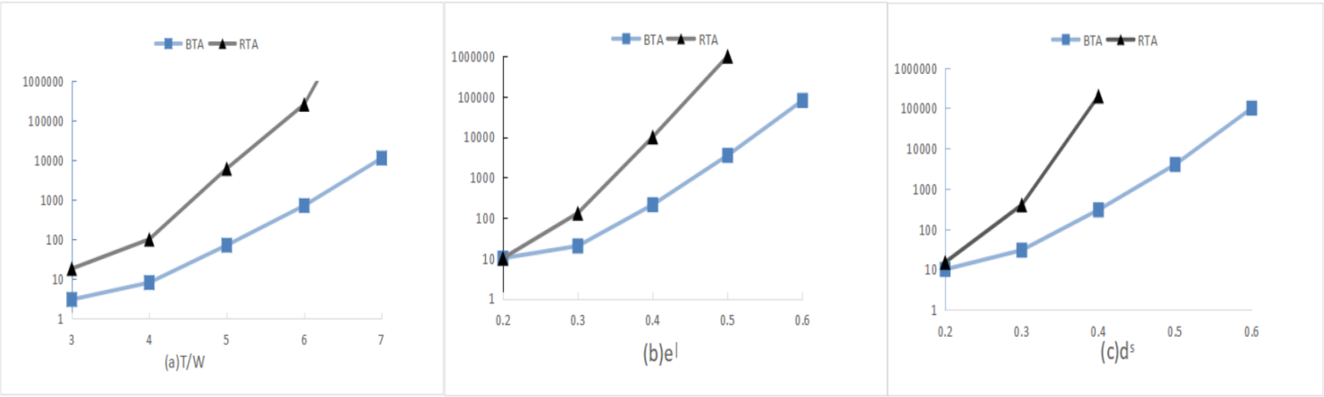


Figure 3: Effect of T/W, e^l and d^s on the Search Time

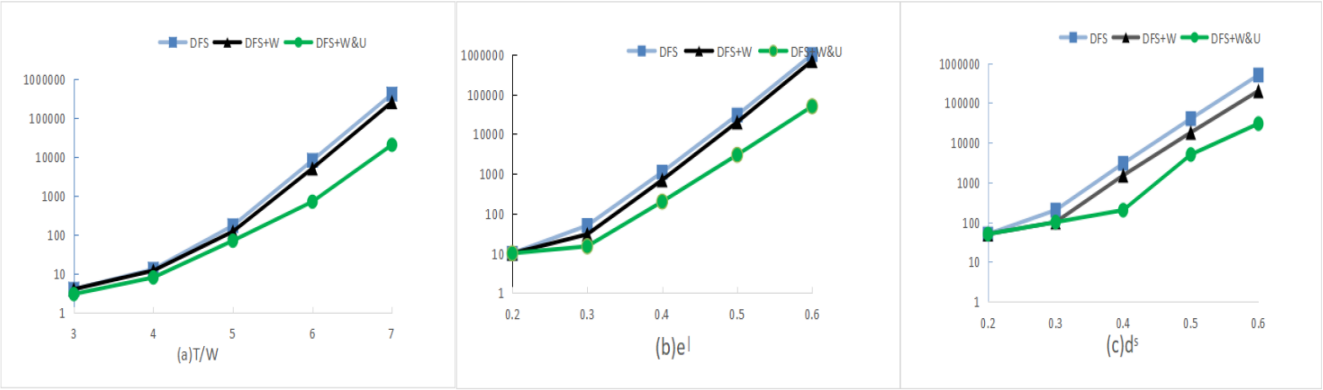


Figure 4: Effect of T/W, e^l and d^s on different search strategies

above three parameters, the number of reachable tasks for workers will increase, and the problem will become more complex.

The sorting based search algorithm DFS+W first searches for the scheme that can be completed quickly, and quickly modifies the value of the heuristic function, so as to prune the later scheme that requires a lot of search time as early as possible Since the ordinate in Figure 4 increases exponentially, it is known that the performance of DFS+W algorithm is improved by constant coefficients compared with DFS algorithm. The DFS+W&U algorithm [11, 15–17] has already calculated a static upper bound for searching each node in advance in the pre calculation process, and combined with the search upper bound dynamically calculated in each step, the heuristic function value is more tightly restricted. It can be seen from Figure 4 that DFS+W&U implements more efficient pruning, and with the problem scale growing, the search performance is improved by at least one order of magnitude compared with DFS.

Table 1 shows the search results of GA, IGA and OPT search algorithms as the number of workers' tied tasks increases. The OPT algorithm is the optimal case under the assumption. Compared with GA, the larger the data size is, the greater the difference between the number of tasks that can be allocated by OPT and the GA algorithm is. Table 2 and Table 3 show that IGA algorithm has less than 20 differences with OPT under the conditions of $e^l < 0.5$ and $d^s < 0.5$, and has obvious advantages over GA. But in more complex cases, even if the local optimization is achieved, there is still a gap of about 10% between the OPT algorithm and the global optimization algorithm.

Table 1: Effect of T/W on theTotal Number of Assigned Tasks, $e^l = d^s = 0.6$

T/W	GA	IGA	OPT
3	97	106	121
4	125	136	167
5	176	189	223
6	222	243	284
7	255	282	327

6 Discussion and Conclusions

Tasks in spatial crowdsourcing have specific location requirements, and only when workers actually travel to the designated location can they complete the task. This paper studies the task assignment problem with the latest working time constraint for workers This paper creatively proposes a tree decomposition method, which divides workers without task dependency into mutually independent worker sets, and uses the best effort search tree construction algorithm to construct a balanced search tree as much as possible. Finally, this paper designs a depth first search algorithm, which combines the optimization strategy to quickly tighten the upper and lower bounds, and can effectively cut the scheme that is not likely to be the optimal solution. The experimental results show that the optimal assignment algorithm proposed in this paper has more advantages in complex task assignment scenarios.

The next research direction is to mine the parallelizable part of this algorithm, and use parallel algorithms and distributed algorithms to achieve more efficient allocation schemes, so as to facilitate the promotion of this algorithm to the practical application scenarios with larger and more complex data.

Table 2: Effect of e^l on the Total Number of Assigned Tasks, $T/W = 5$, $d^s = 1$

e^l	GA	IGA	OPT
0.2	113	120	124
0.3	145	154	168
0.4	174	194	216
0.5	220	243	279
0.6	263	285	328

Table 3: Effect of d^s on the Total Number of Assigned Tasks, $T/W = 5$, $e^l = 1$

d^s	GA	IGA	OPT
0.2	140	143	148
0.3	165	172	183
0.4	218	226	242
0.5	259	273	295
0.6	298	317	342

Acknowledgments

This work is supported by 2023 Zhejiang Traditional Chinese Medical Science and Technology Program(Fund No. 2023ZF010). This article has been funded by the 2021 School Level Curriculum Ideological and Political Demonstration Course Construction Project (Project No.:KCSZ202108), the Jiangsu Province University Philosophy and Social Science Research General Project "Research on Deep Learning of Students after Class under the School Enterprise Collaborative Education Model" (Project No.: 2021SJA1443), and the Jiangsu Province University "Blue Project" project. The authors gratefully acknowledge the anonymous reviewers for their valuable comments.

References

[1] A. A. Alabbadi, M. F. Abulkhair, "Multi-objective task scheduling optimization in spatial crowdsourcing," *Algorithms*, vol. 14, no. 3, pp. 77, , 2021.

[2] N. Bhaskar, "Optimal processing of nearest-neighbor user queries in crowdsourcing based on the whale optimization algorithm," *Soft Computing: A fusion of foundations, methodologies and applications*, vol. 24, no. 17, 2020.

[3] S. S. Bhatti, Y. Chang, X. Gao X, *et al.*, "Affinitive diversity-aware task allocation in spatial crowdsourcing," in *IEEE International Conference on Web Services (ICWS'20)*, 2020.

[4] U. Hassan, E. Curry, "Efficient task assignment for spatial crowdsourcing: A combinatorial fractional optimization approach with semi-bandit learning," *Expert Systems with Applications*, vol. 58, pp. 36-56, 2016.

[5] M. Li, J Wang, L. Zheng, *et al.*, "Privacy-preserving batch-based task assignment in spatial crowdsourcing with untrusted server," in *30th ACM International Conference on Information & Knowledge Management*, pp. 947-956, 2021.

[6] Y. Li, L. Chang, L. Li, *et al.*, "TASC-MADM: Task assignment in spatial crowdsourcing based on multiattribute decision-making," *Security and Communication Networks*, vol. 2021, pp. 1-14, 2021.

[7] X. Liu, J. Fu, Y. Chen Y, *et al.*, "Trust-aware sensing quality estimation for team crowdsourcing in social IoT," *Computer Networks*, vol. 184, no. 6, pp. 107695, 2021.

[8] Z. Liu, K. Li, X. Zhou, *et al.*, "Multi-stage complex task assignment in spatial crowdsourcing," *Information Sciences: An International Journal*, vol. 586, pp. 119-139, 2022.

[9] L. Qian, G. Liu, F. Zhu, *et al.*, "Enhancing user experience of task assignment in spatial crowdsourcing: A Self-Adaptive Batching Approach," *IEEE Access*, vol. 7, no. 99, pp. 132324-132332, 2019.

[10] Y. Sun, M. Liu, L. Huang, *et al.*, "An embedding-based deterministic policy gradient model for spatial crowdsourcing applications," in *IEEE 24th International Conference on Computer Supported Cooperative Work in Design (CSCWD'21)*, pp. 1268-1274, 2021.

[11] Y. Sun, W. Tan, "Combining spatial optimization and multi-agent temporal difference learning for task assignment in uncertain crowdsourcing," *Information Systems Frontiers*, vol. 22, pp. 1447-1465, 2020.

[12] Q. Tao, Y. Tong, Z. Zhou, *et al.*, "Differentially private online task assignment in spatial crowdsourcing: A tree-based approach," in *IEEE International Conference on Data Engineering*, 2020.

[13] Y. Tian, B. Song, T. Ma, *et al.*, "Bi-tier differential privacy for precise auction-based people-centric IoT service," *IEEE Access*, vol. 9, pp. 55036-55044, 2021.

[14] H. To, I. Fan, T. Luan, *et al.*, "Real-time task assignment in hyperlocal spatial crowdsourcing under budget constraints," in *IEEE International Conference on Pervasive Computing & Communications*, 2016.

[15] Y. X. Tong, Y. Yuan, Y. R. Cheng, L. Chen, G. R. Wang, "Survey on spatiotemporal crowdsourced data management techniques," (in Chinese with English abstract) *Journal of Software*, vol. 28, no. 1, pp. 35-58, 2017.

[16] K. Xiong, Y. Dong, Z. Guo, *et al.*, "Exploring the ranking, classifications and evolution mechanisms of research fronts: A method based on multiattribute decision making and clustering," *International Journal of Information Technology & Decision Making*, vol. 22, no. 01, pp. 157-185, 2023.

[17] C. Zhang, Y. Guo, P. Lin, *et al.*, "Location prediction-based task assignment in spatial crowdsourcing," *Journal of Nanjing University (Natural Science)*, vol. 2018, no. 2, pp. 471-480, 2018.

[18] C. Zhang, J. A. Shah, "Co-optimizating multi-agent placement with task assignment and scheduling," *AAAI Press*, pp. 3308-3314, 2016.

[19] S. Zhang, T. Zhang, S. Z. Li, *et al.*, "Geo-indistinguishable mechanisms for spatial crowdsourcing via multi-objective evolutionary optimization," *arXiv e-prints*, 2022.

Biography

Wei Sun received his B.S. and M.S. degrees from the School of Information Engineering, Zhengzhou University, China, in 2003 and 2008, respectively. He is currently working in the School of Computer and Information Technology, Xinyang Normal University. His current research interests include access control and system security.

Jun Lu is currently working in the School of Computer and Information Technology, Xinyang Normal University. His current research interests include network and system security.

Huacheng Xie is currently working in the School of Computer and Information Technology, Xinyang Normal University. His current research interests include network and system security.

Research on the Detection of Illegal Transactions in Currency Transactions Based on Blockchain Technology

Shuang Wu[1], Quanzhi Liu[2], and Jie Li[3]
(Corresponding author: Shuang Wu)

Department of Planning and Finance, Cangzhou Normal University, Cangzhou, Hebei 061001, China[1]
Department of Mathematics and Statistics, Cangzhou Normal University, Cangzhou, Hebei 061001, China[2]
Department of Publicity, Cangzhou Normal University, Cangzhou, Hebei 061001, China[3]
Email: wusw87@gmail.com

(Received Oct. 9, 2022; Revised and Accepted Nov. 26, 2023; First Online Dec. 24, 2023)

Abstract

Currency transactions in blockchain are affected by the characteristics of blockchain, and many illegal transactions need to be detected to improve security. This paper first briefly introduces currency transactions and illegal transactions in blockchain. Then, a detection method was designed to combine the Bagging algorithm and graph convolutional neural network (GCN) algorithm, i.e., Bagging-GCN. Experiments were conducted on the Elliptic dataset. The results showed that compared with traditional machine learning methods such as logistic regression and support vector machine, the Bagging-GCN method better distinguished legal and illegal transactions. In addition, regarding feature input, the effect of using all features was better than that of only using the first 94 features. In the case of sample imbalance, the Bagging-GCN method always maintained a stable detection performance. In contrast, the detection performance of the GCN method declined rapidly with the aggravation of sample imbalance. Finally, when the number of weak classifiers used was 6, the Bagging-GCN method had the best detection performance, with an accuracy of 0.937, an F_1 value of 0.972, and an area under the curve of 0.956. The results prove the reliability of the Bagging-GCN method in detecting illegal transactions and its application potential in the actual blockchain.

Keywords: Blockchain; Currency Transaction; Graph Convolutional Neural Network; Illegal Transaction

1 Introduction

Blockchain technology is a distributed ledger system with the characteristics of decentralization, non-tampering and traceability [4, 5, 8]. It is also a virtual currency trading platform. By using the characteristics of transactions such as non-tampering, the virtual currency represented by Bitcoin is widely used and circulated in more and more fields [10]. With the continuous development of blockchain technology, the type and quantity of virtual currency have also developed [14, 22], and Bitcoin, Ethereum, Ripple, etc. [15] have appeared, which has brought a certain impact on the financial market [17]. Due to the high anonymity and decentralization of virtual currency, illegal transactions such as money laundering and fraud occur frequently. At the same time, due to the characteristics of blockchain, it is more difficult to track and detect illegal transactions [6, 7].

With the improvement of regulations and regulations, illegal transactions have been suppressed to a certain extent. With the progress of machine learning, more and more technologies have been used to detect illegal transactions on the blockchain [12]. Zhang *et al.* [26] established a temporal Bitcoin network model to detect local abnormal users. Through experiments on real data sets, it was found that the method improved the recall rate and F2 value. Mittal *et al.* [18] proposed a method to identify suspicious users based on reputation scores through collaborative centrality measurement and machine learning techniques. Experiments were conducted on two cryptocurrency datasets. It was found that the proposed method provided accurate results. Wang *et al.* [25] proposed a method based on oversampled long short-term memory for Ponzi scheme in Ethereum and proved the effectiveness of the method through experiments on the XBlock dataset.

Oad *et al.* [20] proposed a blockchain-based transaction scanning method to detect abnormal behaviors. Through experiments, it was found that this method had a certain limiting effect on money laundering events. However, faced with the anonymity and concealment of blockchain, as well as the continuous innovation of criminal methods, the detection of illegal transactions is still facing great challenges. It is still necessary to constantly update the existing detection methods to ensure the security of currency transactions in blockchain. Therefore, this paper designed a Bagging-graph convolutional neural network (GCN) detection method and proved its effectiveness through experiments on the Elliptic data set. It provides a new available method for the detection of illegal transactions in blockchain, which is conducive to promoting the security of the virtual currency market.

2 Analysis of Illegal Transactions in Blockchain

2.1 Monetary Transactions in the Blockchain

Blockchain is composed of a series of blocks, which realizes the distributed consistency of each node through a consensus mechanism, and ensures the security of the distributed ledger through encryption algorithms and digital signatures. Virtual cryptocurrency is one of the key technologies [21], which has attracted wide attention from the society [19]. Bitcoin is a representative of the first generation of blockchain technology [24]. This paper mainly takes Bitcoin as an example to analyze the currency transactions in the blockchain.

The transaction of Bitcoin can be regarded as a process of completing accounting, and a transaction can have multiple parties involved, and then a series of transactions are organized into blocks, and blocks are linked into chains, so that a continuously growing sample can be obtained. The block body is used to record transaction information, the block header contains the hash value of the previous block to realize transaction traceability, the markle root stores the synthesis of all transactions in the block, and the reliable data transmission is realized through the Proof of work (PoW) mechanism. In this data structure, a currency transaction can be represented by a network, and the transaction process can be represented as a directed acyclic graph composed of nodes and edges, and each directed edge contains the input and output information of the transaction.

2.2 Illegal Transaction Analysis

Monetary transactions in blockchain have the following characteristics.

Decentralization

Transactions in the blockchain are anonymous, and anyone can inquire [1]. However, due to the lack of centralized management, it is more convenient for criminals to conduct illegal transactions.

High Liquidity

Compared with the traditional physical currency, the transaction of virtual currency is more convenient, so the scope of circulation is also wider. In the face of such rapid advancements, illegal transactions have also been facilitated.

High Secrecy

The two sides of the virtual currency transaction can not know each other's identity, which has brought great convenience to the use of dark networks, underground banks and other organizations, and has become a common transaction method for criminals.

Under the influence of these characteristics, there are more and more illegal transactions in virtual currency, and money laundering is one of them. Because the currency transaction in the blockchain does not require the intervention of a third party, the transaction is public and cannot be tampered with, and the real identity information does not need to be uploaded, it is difficult to crack the identity of both sides, which brings great difficulty to the tracking of regulators. In addition, the circulation of virtual currency is very fast, not affected by foreign exchange control, etc., and virtual currency can be transferred to countries with loose supervision to complete black money cleaning. This illegal transaction is manifested in the transaction network. Nodes with frequent transfer, centralized transfer, decentralized transfer, and other phenomena need to be paid special attention to, and the risk of illegal transactions is high.

2.3 Experimental Data Set

The experimental data for this study is the Elliptic Data Set published by Elliptic Company in 2019 [11], which is a real Bitcoin transaction data set, containing 203,769 transaction nodes and 234,355 transaction edges. The nodes are divided into three categories, one is illegal transactions, including money laundering, fraud, Ponzi scheme, etc. A total of 4,545 nodes (2%) are marked as 1, a total of 42,019 nodes (21%) are legal transactions, marked as 0, and the remaining nodes are unknown, marked as unknown.

In this dataset, all nodes contain 166 features, the first 94 are transaction information, including input/output number, transaction fee, time step, etc., and the last 72 are aggregated features, which are derived from the neighbor information of the transaction node, including the maximum, minimum, standard deviation, and correlation coefficient of neighbor nodes.

3 Detection Method Based on Graph Convolutional Neural Network

3.1 Graph Convolutional Neural Network

Currency trading network has a graph structure, and GCN has many successful applications in the processing of graph structure [9]. Therefore, based on GCN, this paper designs an illegal trading detection method.

Assume that the currency transaction graph is written as $G = (E, V)$, where E is the edge feature and V is the node feature. There is a Laplacian matrix:

$$L_G = D - A,$$

where D is the degree matrix of G and A is the adjacency matrix. After normalizing the above formula, the normalized Laplacian matrix is obtained as follows:

$$L = I_N - D^{-1/2}AD^{-1/2}.$$

In the above equation, I_N is a $N \times N$ unit matrix. Eigenvalue g_θ and eigenvector U are obtained after decomposing the above formula. By simplifying the convolution operation on the graph using Chebyshev polynomials, the following equation is obtained.

$$g_\theta * x \approx \theta(I_N + D^{-1/2}AD^{-1/2})x.$$

The final calculation formula of single-layer GCN can be written as:

$$\begin{aligned} H^{(l+1)} &= \sigma(\widetilde{D}^{-1/2}\widetilde{A}\widetilde{D}^{-1/2}H^{(l)}W^{(l)}), \\ \widetilde{A} &= A + I_N, \end{aligned}$$

where σ is the ReLU activation function, $H^{(l)}$ is the output of the l^{th} layer of GCN, $W^{(l)}$ is the weight matrix of the l^{th} layer of GCN, and $\widetilde{D}$ is the degree matrix corresponding to $\widetilde{A}$.

3.2 Bagging Algorithm

In the Elliptic data set, there is a very serious imbalance between positive and negative samples. Legal transactions are the main part of the currency transaction network, and the proportion of illegal transactions is very small, which will affect the effect of algorithm training.

Ensemble learning refers to the aggregation of multiple weak classifiers to output the final classification results, which can often obtain better results than a single classifier. The commonly used methods are Bagging, Boosting and Stacking [23], and the Bagging algorithm [13] is selected in this paper. The principle of Bagging algorithm is shown in Figure 1. A training subset is randomly sampled from the original training set. It is repeated many times, and different base classifiers are obtained after training using different training subsets. Then, through a certain combination strategy, the classification results of different base classifiers are aggregated to obtain the final result.

The Bagging algorithm is introduced into GCN to obtain the Bagging-GCN method, which is applied to illegal transaction detection. The specific process is as follows.

1) The majority class samples in the training set, namely legal transaction samples, are randomly undersampled, and merged with the minority class samples, namely illegal transactions, to obtain a balanced training set, which is repeated M times to obtain M balanced training sets.

2) Each balanced training set is used to train a GCN weak classifier, and M GCN weak classifiers are obtained.

3) The M GCN weak classifiers are combined to obtain the Bagging-GCN ensemble classifier.

4) The samples in the test set are input into the Bagging-GCN ensemble classifier to obtain the final classification result:

$$G(x) = argmax_{y \in Y} \sum_{M=1}^{m} (y = G_M(x)).$$

4 Results and Analysis

4.1 Experimental Setup

The experiment was carried out under Windows 10 system, and all algorithms were implemented in PyTorch and Python 3.6. The experimental data was Elliptic data set, training set: validation set: test set = 2:1:2. In the Bagging algorithm, the number of rounds was set to 100, the number of negative examples sampled in each round was 1,818, which was the same as the number of positive examples, and the maximum number of iterations in each round was set to 20. In the GCN algorithm, the three-layer GCN structure was used, the number of hidden units was 16-16-8, the learning rate was 0.01, the Dropout was 0.5, and the Adam optimization algorithm was used to train six weak classifiers. As it is a binary classification problem, the evaluation of the illegal trade detection algorithm was based on the confusion matrix shown in Table 1, and the specific indicators are given in Table 1.

Table 1: Confusion matrix

Actual Value	Test Results	
	Positive example	Negative example
Positive example	TP	FN
Negative example	FP	TN

1) Accuracy: $ACC = \frac{TP+TN}{TP+FN+FP+TN}$.

2) F_1 value: $F_1 = \frac{2PR}{P+R}$, where P stands for precision, $P = \frac{TP}{TP+FP}$, R stands for recall rate, $R = \frac{TP}{TP+FN}$.

3) Area under the curve (AUC): The area between the receiver operator characteristic (ROC) curve and the axis; the larger the value, the better the classification performance of the algorithm.

4.2 Results Analysis

Firstly, the Bagging-GCN method designed in this paper was compared with some traditional machine learning

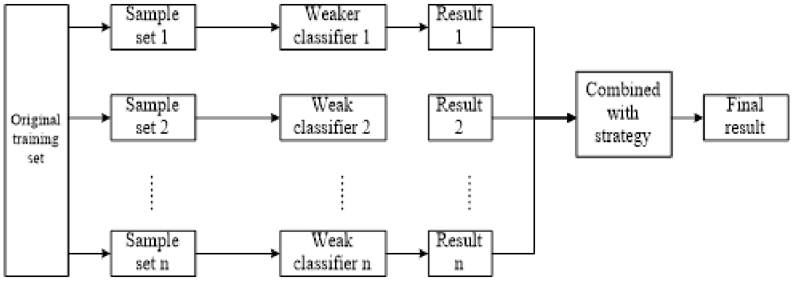


Figure 1: Bagging algorithm

methods, including logistic regression (LR) [2], support vector machine (SVM) [3], and random forest (RF) [16].

In addition, the influence of feature selection on detection results was also compared. The first 94 transaction information features (PF) and ALL features (ALL) were used as input, and the results are presented in Table 2.

Table 2: Comparison with other machine learning methods

	ACC	F_1 value	AUC
LR-PF	0.313	0.587	0.469
LR-ALL	0.412	0.623	0.511
SVM-PF	0.473	0.647	0.597
SVM-ALL	0.497	0.692	0.612
RF-PF	0.592	0.703	0.684
RF-ALL	0.632	0.764	0.712
Ours-PF	0.914	0.954	0.932
Ours-ALL	0.937	0.972	0.956

From Table 2, firstly, from the point of view of feature selection, the detection effect of different methods was always unsatisfactory when PF was used as input. Taking LR as an example, the ACC of LR-ALL was 0.412, which was 0.099 higher than that of LR-PF, the F_1 value was 0.623, which was 0.036 higher than that of LR-PF, and the AUC value was 0.511, which was 0.042 higher than that of LR-PF. These results showed that the aggregated features contained a lot of useful information for distinguishing legal and illegal transactions, and the features of neighbor nodes as a supplement to the original transaction node information could further improve the detection performance of the algorithm. Therefore, using all features could obtain better detection accuracy in illegal transaction detection.

Then, from the comparison of different methods, the LR method had the worst performance in illegal transaction detection, which may be because the data set is not linearly separable and the amount of data is large. The SVM method achieved slightly better results than the LR method, but it also suffered from the problem of overfitting, which led to poor classification effect. When ALL was used as input, the AUC of the RF approach reached 0.712, which showed that RF had a good performance in the processing of high-dimensional data and was capable of distinguishing between legal and illegal transactions even in cases of imbalanced samples. Finally, the Bagging-GCN method solved the problem of sample imbalance more properly and had excellent generalization ability. Ours-ALL had the best overall performance with an ACC of 0.937, an F_1 value of 0.972, and an AUC of 0.956. The reliability of the Bagging-GCN method in the detection of illegal transactions in blockchain was proved.

In order to discuss the influence of the Bagging algorithm on detection results, several combinations of samples were selected for comparison, as follows:

1) negative sample: positive sample (legal transaction sample: illegal transaction sample) =5:1;

2) negative sample: positive sample (legal transaction sample: illegal transaction sample) =10:1;

3) negative sample: positive sample (legal transaction sample: illegal transaction sample) =50:1.

Comparing the detection effect of the GCN and Bagging-GCN methods, the results are displayed in Table 3.

From Table 3, firstly, with the aggravation of unbalanced samples, the Bagging-GCN method always maintained a relatively stable detection performance. Whether it was 5:1 or 50:1, the ACC, F_1 value, and AUC always remained above 0.9, indicating that the Bagging-GCN method could handle unbalanced samples well, so as to maintain a stable ability to distinguish between legal and illegal transactions. When the negative sample: positive sample = 50:1, the ACC of the GCN method was only 0.424, the F_1 value was 0.567, and the AUC dropped to 0.497, showing a very large gap with the Bagging-GCN method. In the case of very unbalanced positive and negative samples, the features learned by the GCN method were not enough to accurately distinguish legal transactions from illegal transactions in the blockchain. However, after processing the samples with the Bagging algorithm, the performance of the GCN method could be guaranteed, which proves the advantage of introducing the Bagging method.

Table 3: Comparison between the GCN and Bagging-GCN methods

Negative samples: positive samples (legal transaction samples: illegal transaction samples)	Method	ACC	F_1 value	AUC
5:1	GCN	0.897	0.912	0.874
	Bagging-GCN	0.922	0.964	0.941
10:1	GCN	0.519	0.712	0.627
	Bagging-GCN	0.931	0.969	0.951
50:1	GCN	0.424	0.567	0.497
	Bagging-GCN	0.929	0.971	0.954

Six weak classifiers were used in the above experiments, and the number of weak classifiers was changed to compare the detection results, as shown in Figure 2.

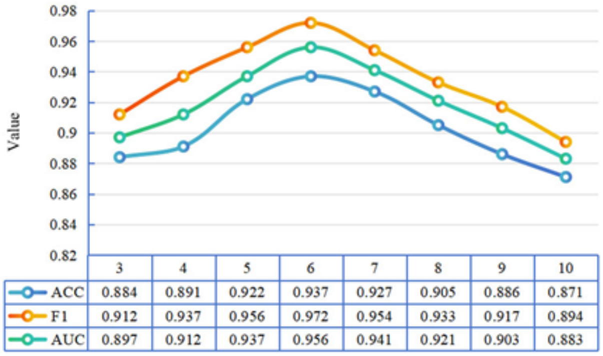


Figure 2: Comparison under different numbers of weak classifier

From Figure 2, it can be found that when the number of weak classifiers increased from 3 to 6, the ACC, F_1 value, and AUC of the Bagging-GCN method showed an upward trend, indicating that the performance of the algorithm was improved with the increase of the number of weak classifiers. Through the combination of more weak classifiers, the ability of the model to distinguish between legal and illegal trading samples was strengthened, and better classification results were obtained. However, with the further increase of the number of weak classifiers, the performance of the Bagging-GCN method decreased. When the number of weak classifiers was 10, the results of all indicators were lower than the situation when the number of weak classifiers was 3. This result showed that in the case of too many weak classifiers, it will lead to overfitting of the algorithm, resulting in the decline of detection performance. Therefore, in the illegal transaction detection, the best effect was obtained when six weak classifiers were used, which could make a more accurate judgment on whether the transaction in the sample is illegal transaction.

5 Conclusion

This paper mainly studied the illegal transaction detection methods in blockchain currency transactions and designed a Bagging-GCN detection method for the imbalance problem of the data set. Through experiments on the Bitcoin transaction dataset, it was found that the Bagging-GCN method had better performance in illegal transaction detection than traditional machine learning methods such as LR. The introduction of the Bagging algorithm effectively solved the problem of sample imbalance. At the same time, the number of weak classifiers had a certain impact on the detection results. When the number of weak classifiers was 6, the performance was the best, and the AUC reached 0.956. The Bagging-GCN approach can well distinguish legal and illegal transactions and can be further applied in practice.

References

[1] C. G. Akcora, M. F. Dixon, Y. R. Gel, M. Kantarcioglu, "Bitcoin risk modeling with blockchain graphs," *Economics Letters*, vol. 173, no. DEC., pp. 138-142, 2018.

[2] S. Bagherzadeh, D. Shahbazi-Gahrouei, F. Torabinezhad, S. Mahdavi, P. Fadavi, S. Salmanian, "Binary logistic regression modeling of voice impairment and voice assessment in Iranian patients with nonlaryngeal head-and-neck cancers after chemoradiation therapy: Objective and subjective voice evaluation," *Journal of Medical Signals and Sensors*, vol. 13, pp. 40-48, 2023.

[3] F. Camastra, V. Capone, A. Ciaramella, A. Riccio, A. Staiano, "Prediction of environmental missing data time series by Support Vector Machine Regression and Correlation Dimension estimation," *Environmental Modelling & Software*, vol. 150, pp. 1-7, 2022.

[4] P. Y. Chang, M.S. Hwang, C.C. Yang, "A blockchain-based traceable certification system", in *Security with Intelligent Computing and Big-data Services*, pp. 363-369, 2018.

[5] L. Chen, Q. Fu, Y. Mu, L. Zeng, F. Rezaeibagha, M. S. Hwang, "Blockchain-based random auditor com-

mittee for integrity verification", *Future Generation Computer Systems*, vol. 131, pp. 183-193, 2022.

[6] Y. H. Chen, L. C. Huang, I. C. Lin, M. S. Hwang, "Research on the secure financial surveillance blockchain systems", *International Journal of Network Security*, vol. 22, no. 4, pp. 708-716, 2020.

[7] Y. H. Chen, L. C. Huang, I. C. Lin, M. S. Hwang, "Research on blockchain technologies in bidding systems", *International Journal of Network Security*, vol. 22, no. 6, pp. 897-904, 2020.

[8] H. Covington, Y. B. Choi, "Blockchain and bitcoin: Concept, functionality, and security," *International Journal of Cyber Research and Education*, vol. 1, no. 1, pp. 27-37, 2019.

[9] L. Deng, Z. Fan, H. Xu, S. Yu, "PDSM-LGCN: Prediction of drug sensitivity associated microRNAs via light graph convolution neural network," *Methods*, vol. 205, pp. 106-113, 2022.

[10] L. V. T. Duong, N. T. T. Thuy, L. D. Khai, "A fast approach for bitcoin blockchain cryptocurrency mining system," *Integration: The VLSI Journal*, vol. 74, pp. 107-114, 2020.

[11] Elliptic Co., *Know exactly what happens on any blockchain*, Dec. 11, 2023. (`https://www.elliptic.co/`)

[12] S. Fan, S. Fu, H. Xu, X. Cheng, "Al-SPSD: Anti-leakage smart Ponzi schemes detection in blockchain," *Information Processing & Management*, vol. 58, no. 4, pp. 102587, 2021.

[13] N. Ghoggali, F. Douak, W. Ghoggali, "Towards a NIR spectroscopy ensemble learning technique competing with the standard ASTM-CFR: An optimal boosting and bagging extreme learning machine algorithms for gasoline octane number prediction," *Optik*, vol. 257, pp. 1-19, 2022.

[14] A. Hughes, A. Park, J. Kietzmann, C. Archer-Brown, "Beyond bitcoin: What blockchain and distributed ledger technologies mean for firms," *Business Horizons*, vol. 62, no. 3, pp. 273-281, 2019.

[15] N. Kyriazis, S. Papadamou, S. Corbet, "A systematic review of the bubble dynamics of cryptocurrency prices," *Research in International Business and Finance*, vol. 54, no. 3, pp. 101254, 2020.

[16] J. Liang, "Problems and solutions of art professional service rural revitalization strategy based on random forest algorithm," *Wireless Communications and Mobile Computing*, vol. 2022, no. 1, pp. 1-11, 2022.

[17] M. H. Miraz, K. Imran, M. G. Hassan, "Trust impact on blockchain & bitcoin monetary transaction," *Journal of Dynamical and Control Systems*, vol. 12, no. 3, pp. 155-162, 2020.

[18] R. Mittal, M. P. S. Bhatia, "Detection of suspicious or un-trusted users in crypto-currency financial trading applications," *International Journal of Digital Crime and Forensics*, vol. 13, no. 1, pp. 79-93, 2020.

[19] S. M. Mohammad, "Blockchain and bitcoin security in IT automation," *SSRN*, vol. 68, no. 3, pp. 103-110, 2020.

[20] A. Oad, A. Razaque, A. Tolemyssov, M. Alotaibi, B. Alotaibi, C. Zhao, "Blockchain-enabled transaction scanning method for money laundering detection," *Electronics*, vol. 10, no. 15, pp. 1-18, 2021.

[21] A. G. Parker, "Blockchain, bitcoin and the rise of new money," *ITNOW*, vol. 60, no. 4, pp. 8-13, 2018.

[22] Z. P. Peng, M. L. Chiang, I. C. Lin, C. C. Yang, M. S. Hwang, "CBP2P: Cooperative electronic bank payment systems based on blockchain technology", *Journal of Internet Technology*, vol. 23, no. 4, pp. 683-692, 2022.

[23] F. Serafino, G. Pio, M. Ceci, "Ensemble learning for multi-type classification in heterogeneous networks," *IEEE Transactions on Knowledge and Data Engineering*, vol. 30, no. 12, pp. 2326-2339, 2018.

[24] J. Taskinsoy, "Blockchain: Moving beyond bitcoin into a digitalized world," *SSRN*, pp. 1-21, 2019. (`https://ssrn.com/abstract=3471413`)

[25] L. Wang, H. Cheng, Z. Zheng, A. Yang, X. Zhu, "Ponzi scheme detection via oversampling-based long short-term memory for smart contracts," *Knowledge-based Systems*, vol. 228, no. Sep.27, pp. 1-12, 2021.

[26] R. Zhang, G. Zhang, L. Liu, C. Wang, S. Wan, "Anomaly detection in bitcoin information networks with multi-constrained meta path," *Journal of Systems Architecture*, vol. 110, pp. 1-8, 2020.

Biography

Wu Shuang was born in Cangzhou, Hebei Province, China, in November 1987. She studied in Griffith University from 2011 to 2012 and obtained a master's degree in 2012. Since 2012, she has been working in Huaxia bank. From 2017 to the present, she works in Cangzhou Normal University. She has published several papers and has presided over 5 projects at the municipal level. Her research interests are accounting and applied finance. She works in Cangzhou Normal University.

Quanzhi Liu was born in Cangzhou, Hebei, China, in 1987. From 2012 to 2015, he studied in Hebei University and received his Master's degree in 2015. Currently, he works in Cangzhou Normal University. He has published several papers and has presided over 5 projects at the municipal level. His research interests are statistical analysis and financial management.

Jie Li was born in Fuping, Hebei Province, China, in June 1989. She studied at Hebei University from 2012 to 2015 and obtained a master's degree in journalism. Since 2015, she has been working in Cangzhou Normal University. Her research field and specialty are new media communication practice and digital information communication. She has presided over a number of projects and published 5 relevant academic papers.

Research on Education Information Security Sharing Based on Blockchain and Bayesian Network

Daihong Feng[1] and Hang Li[2]
(Corresponding author: Hang Li)

Art School affiliated to Shenyang Normal University[1]
No. 253, Huanghe North Street, Shenyang, 110034 China
Software College & Shenyang Normal University[2]
No. 253, Huanghe North Street, Shenyang, 110034 China
Email: lihangsoft@163.com

(Received July 23, 2023; Revised and Accepted Dec. 1, 2023; First Online Dec. 24, 2023)

The Special Issue on Data Fusion, Deep Learning and Optimization Algorithms for Data Privacy Protection

Special Editor: Prof. Shoulin Yin (Harbin Institute of Technology)

Abstract

Due to the convenience of online education platforms, its business has been developing rapidly. More and more educational institutions provide original offline course content to paying (free) users through online means. The increase in the number of users not only increases business income for educational institutions but also promotes the integration and optimization of online education platform business and data. In order to improve the effect of the information security sharing mechanism on online education platforms, this paper proposes a new education information security sharing method based on blockchain and Bayesian networks. The topological structure of the network is established based on dependency analysis, and the parameters of the network are learned by the expectation maximum (EM) algorithm. The dynamic Bayesian network (DBN) aims to establish an online education-sharing model. Through the smart contract provided by Ethereum, data encryption and bidirectional verification between different platforms are realized, and the data transmitted through the network after the chain can be prevented from being cracked and tampered with. Experiments with other advanced algorithms show that the proposed method can effectively transmit data and encrypt education data securely.

Keywords: Bayesian Network; Blockchain; Education Information; Expectation maximum; Security Sharing Mechanism

1 Introduction

The online education platform is the education platform of the online network, the essence is the sharing of educational resources facing the Internet, and it is a new type of tool platform [22, 35]. On the premise of improving efficiency, the online education platform makes use of all tools to carry out educational activities, and changes the communication mode between teachers and students through the use of advanced network technology to further improve the efficiency of students' mastery of knowledge [33, 34].

Online education platform not only realizes information sharing, but also brings data security problems [2, 19, 39]. First, Internet data transmission between different entities may cause conventional network attacks, and the platform data can be stolen through network monitoring and data decryption functions. Secondly, direct data sharing among users between platforms will lead to the disclosure of unauthorized paid content [25]. Users' private transmission of learning content without clear content ownership will cause economic losses to knowledge providers, and it is difficult to obtain evidence of such incidents, so the platform cannot monitor user behavior in real time. Thirdly, the platform, as a data intermediary, cannot effectively monitor the compliance of personal data transmission behaviors, such as customers recording and disseminating platform educational products without the permission of the platform, tampering with platform data information, etc.; Finally, the expansion of the platform is accompanied by the integration and docking of different application systems and business ecosystems, and the differences in data standards between different systems also restrict the effective transmission of platform data between different business ecosystems [18,26,31,38].

The above problems highlight the shortcomings of information security management of school education platform, which also restricts the development of the platform. In the past, for such problems, engineers used digital watermarking and data bus and other technologies for data processing, but in actual operation, data watermarking technology can not block the illegal transmission of data. Although the data bus technology can solve the problem of inconsistent data standards between different systems to a certain extent, it needs to be reformed when the system is connected, and the technology cannot guarantee the information security of the data center. Therefore, the academic community proposed to learn from the business model and underlying application framework of Bitcoin production and trading, and tried to introduce blockchain technology into the system development and data integration of online education platforms [17,30].

This paper is organized as follows. In Section 2, we introduce the Blockchain technology. Section 3 displays the data distribution structure and key construction. Experiments are conducted in Section 4. There is a conclusion in Section 5.

2 Blockchain Technology

Blockchain technology is a distributed shared accounting technology based on P2P network communication, asymmetric encryption and consistency algorithm, and its essence is a distributed database in which participants keep data accounting according to the consensus mechanism [5,6]. Unlike traditional data storage, blockchain records data in blocks, and the hash value of a Block before a single block is organized into a chain structure as a table header. This structure not only makes it more difficult to tamper with on-chain data and trace the source of data, but also verifies the reliability of the entire data chain according to the key provided by the data holder [7,8]. Blockchain consensus mechanisms mainly include proof of work (PoW) [14], proof of authorization (PoA) [20], proof of stake (PoS) [13], proof of entrusted stake (DPoS) [11], and practical Byzantine fault tolerance (PBFT) [12]. Proof of work (PoW) means that nodes competing for accounting rights calculate accounting rights through random hash, the node that calculates the corresponding value first packages the calculation results to the network, and other nodes are verified and credited to the local account to form a network consensus. Proof of Authorization (PoA) means that a group of nodes in the consortium chain are authorized to be responsible for the generation of new blocks and the verification of existing blocks. Proof of Equity (PoS) is an improvement of proof of work, in which equity is a statistic of the number of bookkeeping rights held and the number of days held, and an individual who has not been bookkeeping for a long time will increase the probability of the next bookkeeping over time. Proof of Equity Authorization (DPoS) means that each node in the system can use its equity as a vote to authorize any node to carry out transaction package calculation, and the node with the most equity votes will automatically become the accounting node. Practical Byzantine fault tolerance (PBFT) is a method of determining accounting nodes by node voting, and the node with the most votes in a specified time slice will be responsible for the accounting work in the next time period. The above methods differ in data exchange rate and technology maturity [24,28].

The network topology of the blockchain system is based on the decentralized structure and adopts the communication controller (CM) [4,36] to realize the direct access of neighbor addresses between nodes and the transit access of non-neighbor addresses. There is no unified distribution and unified storage data center in the system, instead of independent point-to-point communication between nodes. The technical characteristics of blockchain determine that it has the following system characteristics: 1) distribution. Blockchain technology physically stores the data separately, but logically the data is unified, the data has a unified data standard in different branches, the data is stored in the form of a full copy in all participating nodes, and the copy of the data stored by each node is theoretically completely consistent. 2) Fairness. The system using blockchain technology is open to all data access rights within the system that recognizes the consistency algorithm, the copy of the data that the user sees is all the data generated by the system, and the data exchange record between any two nodes in the system is distributed to all other nodes by broadcast, ensuring the unity and fairness of the data record. 3) Autonomy. The blockchain system operates under the guidance of a unified consensus algorithm, the communication between nodes does not require the intervention of a data center, a single node has all the rights to exchange its own data, and the system can realize the business processing between internal nodes without relying on external supervision. 4) Scalability. The blockchain system allows nodes to join or exit the system at any point in time, and for nodes that quit and join, they only need to re-download the missing data from other nodes to access the normal transactions of the current system. The above technical features can well solve the information security problems involved in the current school education platform: the data structure of the blockchain ensures that even if the data is stolen, under the framework of the existing technology, hackers need to encrypt and modify the hash value corresponding to each block, and the complete accepted modification needs to obtain the public key of all uploaded data nodes, which is the premise itself It's extremely difficult [3,16,23].

The scalability of the blockchain system ensures that it can expand the network without limitation under the premise of network bandwidth and node computing capacity, but in practical applications, not all blockchain networks must disclose all information to the Internet, and some nodes only need data to be disseminated within a limited range. Based on this, the industry divides the blockchain applicable network into three parts: public chain, private chain and hybrid chain according to different access rights and user management scope. Public chain refers to the blockchain system that is open to the whole network, the operation rules of the entire system are open and transparent, all nodes in the system do not need to disclose their identities, and the anonymity and privacy of each node in the system are protected. Typical public chain systems are Bitcoin, Ethereum, etc.; Private chain is owned by private individuals or private institutions, only the use of blockchain technology as the underlying bookkeeping technology, bookkeeping rights belong to private individuals or private institutions, not open to the public; Alliance chain is a block chain used within a group or organization, nodes need to register permission to access the block chain and alliance chain is limited to alliance members, the system needs to compete in advance to elect some nodes as bookkeeping roles, the generation of blocks by all pre-selected bookkeepers jointly decided, other non-pre-selected nodes can be traded, but there is no bookkeeping right.

The system based on alliance chain can maximize and efficiently extend the boundaries of the system on the basis of ensuring limited individuals to control system access and secure use of internal data according to permissions, so it is currently the most widely used in the industry. In the financial sector, the use of blockchain can eliminate third-party intermediaries and achieve point-to-point direct docking, thus greatly reducing costs while quickly completing transaction payments. Take the bank KYC(Know Your Customer) as an example. By using the hash algorithm, the bank can encrypt the customer's account and identity information in the bank and generate a unique data block to join the banking alliance chain. Other banks only need to confirm the identity through two steps of hash calculation and blockchain query when creating a second-class account for identity authentication for customers, which significantly reduces the authentication cost of financial institutions and also provides users with a good service experience. In the field of copyright storage, China Copyright Center jointly issued a DCI standard alliance chain system with a number of head Internet media and core institutions. The operation of the system involves copyright operators, copyright owners, consumers and trusted institutions, and can achieve decentralized and trusted real-time transmission of information in the aspects of copyright storage, copyright use detection and tracking, infringement storage and copyright resource sharing On the basis of reducing the difficulty of obtaining evidence and reducing the cost of obtaining evidence, the establishment of the copyright alliance chain has opened up the copyright information island, built the copyright mutual trust between different subjects, and laid the foundation for the formation of an effective digital market. In the field of social services, blockchain can mainly reduce the cost of trust, promote the tracking of social credit data and the encrypted record of the actual credit behavior of social subjects. The Xiongan New Area Management Committee will build a housing rental platform based on the alliance chain with departments such as education, finance, social security and housing operation enterprises. Personal credit, transaction records, rental vouchers, transfer information, etc. of homeowners and tenants are formed through technical services to form a data link, and all aspects of rental housing can verify each other and mutual vouchers, and finally realize complete real-time online transaction of rental housing, ensuring that the masses rent ”only run once” requirements.

Summarizing the current implementation of blockchain applications, it can be found that the emergence of blockchain has solved the problem of information security and trust in Internet transactions (communication), and its characteristics in information security have improved the robustness of the system. This feature can play a strong role in optimizing the information docking of online education platforms and strengthening the safe sharing of online education resources. On the one hand, the decentralized information storage mechanism gets rid of the dependence of the current school education platform on the headquarters database, and massive teaching videos do not need to be stored and forwarded twice through limited central storage nodes, reducing the waste of network resources and storage media. At the same time, the decentralized data transmission mode can increase the ductility of the system, and the data docking between the systems no longer requires the intervention of a centralized database, and the transmission of data is no longer limited by platform technology. On the other hand, the tamper-proof and traceable data structure ensures that the source of data can be found in time even if the data is stolen and maliciously modified, and the combination of electronic ink and other technical means can maximize the protection of the copyright interests of the platform information and reduce the human and scientific costs paid by the platform to prevent infringement. In addition, the distributed ledger function ensures the consistency of data saved between different nodes, avoiding the problem that data loss caused by system causes cannot be recovered. The authentication node that re-connects to the platform only needs to access the node one by one according to the data link record address, and then the complete recovery network of all data under the node can be realized. Based on the above considerations, this paper focuses on the establishment of information sharing mechanism and code implementation of online education platform under the framework of blockchain technology.

3 Data Distribution Structure and Key Construction

3.1 Spatial Dimension Network Information Distribution Structure

Suppose a quadruple $G = (V, E, W, C)$ is used to represent the center of information fusion in a spatial dimensional cyclic awareness network. The information dimension of the spatial dimension cyclic awareness network is d, and the number of distribution sets of information security assessment of the $k-th$ spatial dimension cyclic awareness network is t_0. The activity of the $i-th$ node of the $(k+1)-th$ layer in the data sampling is W^k. The spatial dimension cycle perceives the network time series as x_i^{k+1}, and takes its time series x_i^{k+1} as the amount of information input of the layer $(k+1)-th$ blockchain fusion node $i=1,2,\cdots,N_k$. z represents the energy threshold of the sampling node of the spatial dimension cyclic sensing network. The cloud information flow model of spatial dimensional cyclic sensing network is established as follows:

$$x_n=\sqrt{\frac{(t_0-n\Delta t)}{hz}}-w_n. \tag{1}$$

Where, h is the spatial dimension cyclic perception network information distribution sequence. w_n is the characteristic quantity of information distribution in spatial dimension cyclic sensing network. Based on this, it is determined that the information distribution function in the l layer of the input layer of the spatial dimension cyclic awareness network is:

$$x_j^{k+1}(n-1)=y_j^{k+1}(n-1). \tag{2}$$

$$s_j^k(n-1)=\sqrt{\frac{W_{ij}^k(n-1)}{x_j^k(n-1)}}. \tag{3}$$

$$y_j^{k+1}(n-1)=\frac{n-1}{f(s_j^{k+1})}. \tag{4}$$

Where, y_j^{k+1} is the $k+1$ encryption key at time $n-1$. n is the length of the sequence. $W_{ij}^k(n-1)$ is the weighted component of the safety assessment. $x_j^{k+1}(n-1)$ is $k+1$ torque at time $n-1$. $f(s_j^{k+1})$ is the kernel function. Based on the above analysis, and on the basis of building the information distribution structure model of cyclic sensing network, the key distribution of principal component features of sensing network information is analyzed [19].

3.2 Security Assessment Output of Spatial Dimension Network Information

Based on the principal component feature distribution structure of spatial dimensional cyclic awareness network information and combined with the feature distribution of security assessment protocols [10,27], the sample set of spatial dimensional cyclic awareness network information is set as $x_i, y_i, i=1,2,\cdots,k$. Where k is the number of samples of network information time series, the number of samples collected is normalized, the network information test sequence is input into the linear combination sequence, and the connected graph structure model is obtained.

$$f(x)=\frac{w^{T+1}}{(\phi)_X}-b. \tag{5}$$

Where, w is the autocorrelation distribution moment. b is the deviation vector of information fusion scheduling. The spatial dimension cyclic perception network history information is selected as the initial feature quantity of the information fusion scheduling model, and the error term of the network information fusion is adaptively corrected. The ambiguity function of the spatial dimension cyclic perception network information is obtained as follows:

$$K(x_i,x_j)=\sqrt{\frac{(x_i,y_i)^2}{2(\sigma-1)^2}}. \tag{6}$$

Where, $(x_i,y_i)^2$ is the distribution distance of the Euclidean clustering center. σ is the root mean square error. The key security evaluation protocol is determined according to the ambiguity function, and a linear combination model is obtained as follows:

$$x_k=\sqrt{\frac{a_{n-1}N}{sin(2\pi kn)}}+\sqrt{\frac{b_{n-1}N}{cos(2\pi kn)}}. \tag{7}$$

Where, N is the number of network information nodes of DBN. a_{n-1} is the amplitude of the network information linear programming model. b_{n-1} is the output code of DBN for information security evaluation. Assuming that there are m DBN nodes $A_1A_2\cdots A_n$ that detect spatial dimensional network information distribution, the linear programming problem of spatial dimensional network information security is constructed with the following mathematical expression:

$$max(f)=\int_{i=1}^{m}f(C_{ij})d(ij)\int_{j=1}^{n}f(X_{ij})d(j). \tag{8}$$

Here, $\sum_{j=1}^{m+1}X_{ij}=(a+1)_i$, $(a+1)_i$ is the number of network information nodes. It is assumed that the number of information distribution nodes in the current spatial dimensional cyclic awareness network is $n+1$. It reconstructs the spatial dimensional network information time series phase space as $N_1,\cdots,N_n$, and obtains the association rules of spatial dimensional network information as follows:

$$x_i(n+1)=\sqrt{\frac{h_{ij}(n+1)^T}{s_j(n+1)}}-v_i(n+1). \tag{9}$$

Where $h_{ij}(n+1)$ is the cross-correlation feature set, $s_j(n+1)$ is the principal component of PCA, and $v_i(n+1)$ is the information entropy. Combine information fusion theory. The key security evaluation protocol distribution of text, location, picture, audio, video and other information in network information is perceived in a spatial dimension to realize the fusion of information security evaluation [36].

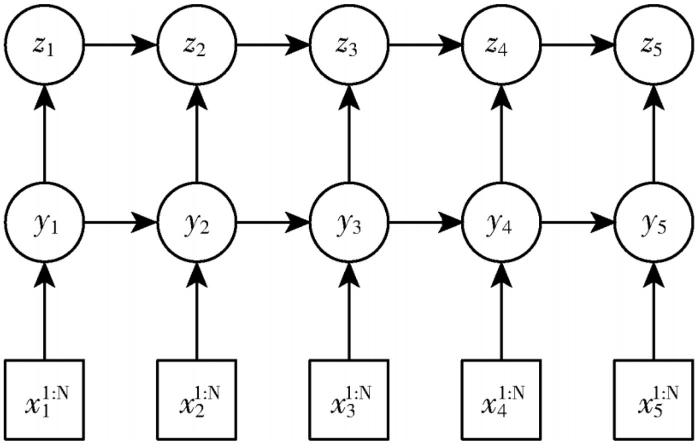


Figure 1: DBN model

3.3 DBN Network Structure

In order to more accurately describe the interdependence of various data, the DBN topology S in this paper adopts three-layer discrete DBN [1,29,37], and the network structure is shown in Figure 1.

The first layer is the observed variable layer. The observed variable is determined when the monitoring data for the current moment is known. A Bayesian network is used to judge the observed variable X_t, output the hidden state $y_t(y_t \in y^1, y^2, \cdots)$ at the current moment, and combine the hidden state $y_\tau(\tau = 1, 2, 3, \cdots, t-1)$ at the historical moment to form a state sequence $Y = y_1, y_2, \cdots, y_t$. The first layer of the network represents the joint probability relationship between the observed variable and the hidden state, and the network structure corresponding to the observed variable is different in different hidden states. As for the topology of the first layer network of DBN, this paper adopts the method of dependency analysis to establish the network topology.

The second layer is the hidden state variable layer. The second layer decodes the hidden state sequence $Y = y_1, y_2, \cdots, y_t$ and outputs the state sequence $Z = z_1, z_2, \cdots, z_t$. The joint probability of adjacent time slices is connected by the implicit state transition probability.

The third level is the decision-making level. According to the status sequence Z, the running state of the bearing is judged comprehensively. If the bearing fails, the fault is further diagnosed online. In the second and third layers of the DBN model, the state loop of state variables is allowed to describe the state fluctuation of the fault process. In an application, different numbers of state Spaces can be assigned to Layer 2 and Layer 3 networks.

4 Simulation Experiment and Result Analysis

In this paper, HCA [32], BPRP [21], Waldo [9] and proposed method are compared from the aspects of throughput, delay, transaction request completion rate and CPU utilization, so as to verify the effectiveness, applicability and energy saving of the algorithm. The PC used in the simulation experiment is configured with Intel Core i7-9750H 2.6GHz CPU and Intel Core i5-6500 3.2GHz CPU for dual-computer connection to realize the master-slave discrete simulation test. Java multithreading mechanism is used to simulate the communication interaction process of consensus nodes in the network.

Transaction delay refers to the time interval for the client to send a transaction request to the master node to confirm the completion of the consensus [37]. The transaction delay of this experiment takes the average value of 200 transactions. Figure 2 shows the comparison of transaction delays of different algorithms.

As can be seen from Figure 2, when there are evil nodes, with the increase of the number of nodes, compared with HCA algorithm, the growth rate of BPRP transaction delay is effectively slowed down. However, because BPRP adds dynamic processing mechanism, the delay is slightly higher than Waldo. Although the proposed method increases part of the delay, it has better flexibility and stability in data processing.

Table 1 shows the comparison of effective completion rates of transaction requests for different algorithms. With the increasing number of nodes, the effective completion rate of HCA and BPRP showed a significant downward trend, while BPRP decreased to a certain extent, but the overall change was relatively stable. Therefore, Waldo algorithm has better stability in the consensus process, which greatly reduces the probability of the evil node being elected as the master node. At the same time, the mechanism of actively deleting the evil node can eliminate the evil node from the network, improve the overall security and stability of the node in the network, and make each request reach a consensus smoothly. The method in this paper can reduce the waste of network resources caused by repeated transmission, so as to reduce the network overhead.

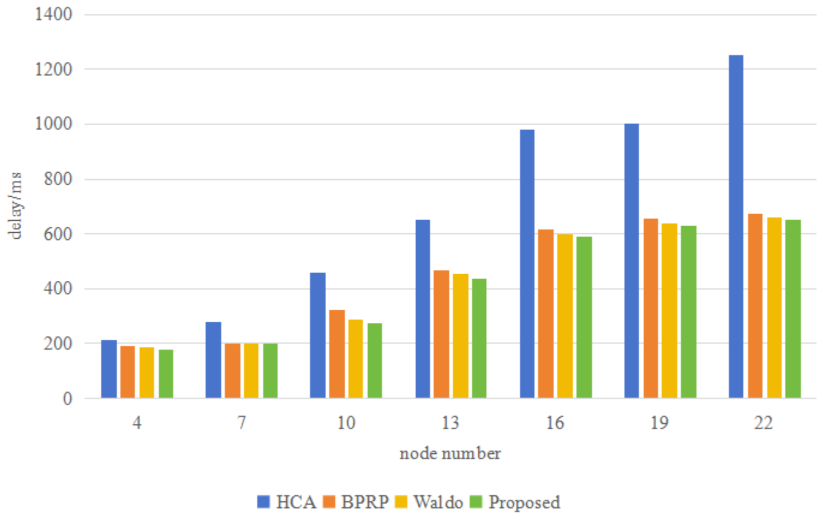


Figure 2: Time delay comparison of different algorithms

Table 1: Effective completion rate of transaction request for different algorithms/%

Node number	HCA	BPRP	Waldo	Proposed
4	93.5	92.1	90.1	87.9
7	91.9	88.3	84.9	79.3
10	92.9	85.5	81.5	75.9
13	93.5	81.5	79.5	71.3
16	89.3	80.1	75.9	63.3
19	87.1	77.5	74.3	54.1
22	84.5	73.5	69.7	45.9

Table 2 shows the time cost comparison. Based on the analysis of Table 2, it can be seen that the time cost of spatial dimensional network information scheduling by the proposed method is small, and the real-time performance of spatial dimensional network information scheduling and mining is improved.

Table 2: Typical states of SEIR model/ms

Iteration no.	HCA	BPRP	Waldo	Proposed
100	12.698	10.473	6.159	1.342
200	16.446	12.851	8.106	2.584
300	18.499	15.674	10.454	3.521
400	22.165	19.258	17.361	4.593

5 Conclusions

The introduction of blockchain technology facilitates the secure sharing of data on online education platforms. The decentralized (multi-centralized) blockchain system reduces the network security and data storage pressure of the data center of the traditional online education platform, while the data is not easy to tamper with and the characteristics of distributed ledger can increase the security and fault tolerance of information transmission between nodes. Through the development and testing of system functions, this paper verifies that blockchain technology, as the basis of information security transmission, can optimize the information sharing mode of the existing school education platform. The method of node + platform dual verification can avoid data leakage caused by the attack of a single node, simplify the process of platform docking, and open up the "data barriers" between platforms. Lay the technical foundation for the subsequent multi-channel integration development and more comprehensive and specific big data analysis of the online education platform, but it should also be noted that blockchain technology in the current application also has shortcomings such as the inability to roll back the on-link data, and the exponential increase in computing and transmission costs over time. Therefore, it is suggested that subsequent research can be considered from the perspective of the integration and development of blockchain technology, artificial intelligence technology and 5G communication technology, give full play to the advantages of 5G transmission bandwidth and the characteristics of artificial intelligence self-learning, and guide blockchain technology to enter more traditional industries and help industrial upgrading.

Acknowledgments

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

References

[1] H. B. Ajmal and M. G. Madden, ”Dynamic Bayesian network learning to infer sparse models from time series gene expression data,” *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 19, no. 5, pp. 2794-2805, 2022.

[2] A. Alam, ”Platform utilising blockchain technology for eLearning and online education for open sharing of academic proficiency and progress records,” in *Smart Data Intelligence: Proceedings of ICSMDI 2022. Singapore: Springer Nature Singapore*, pp. 307-320, 2022.

[3] M. N. M. Bhutta, A. A. Khwaja, A. Nadeem, H. F. Ahmad, M. K. Khan, ”A survey on blockchain technology: evolution, architecture and security,” *IEEE Access*, vol. 9, pp. 61048-61073, 2021.

[4] B. Chang, W. Tang, X. Yan, X. Tong and Z. Chen, ”Integrated scheduling of sensing, communication, and control for mmWave/THz communications in cellular connected UAV networks,” *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 7, pp. 2103-2113, 2022.

[5] P. Y. Chang, M.S. Hwang, C.C. Yang, “A blockchain-based traceable certification system”, in *Security with Intelligent Computing and Big-data Services*, pp. 363-369, 2018.

[6] L. Chen, Q. Fu, Y. Mu, L. Zeng, F. Rezaeibagha, M. S. Hwang, “Blockchain-based random auditor committee for integrity verification”, *Future Generation Computer Systems*, vol. 131, pp. 183-193, 2022.

[7] Y. H. Chen, L. C. Huang, I. C. Lin, M. S. Hwang, “Research on the secure financial surveillance blockchain systems”, *International Journal of Network Security*, vol. 22, no. 4, pp. 708-716, 2020.

[8] Y. H. Chen, L. C. Huang, I. C. Lin, M. S. Hwang, “Research on blockchain technologies in bidding systems”, *International Journal of Network Security*, vol. 22, no. 6, pp. 897-904, 2020.

[9] E. Dauterman, M. Rathee, R. A. Popa and I. Stoica, ”Waldo: A private time-series database from function secret sharing,” *IEEE Symposium on Security and Privacy (SP), San Francisco, CA, USA*, pp. 2450-2468, 2022, doi: 10.1109/SP46214.2022.9833611.

[10] Y. Feng, J. Yang, L. Huang, B. Ji, J. Su, ”Intellectualization and reliability evaluation of distribution network based on principal component analysis,” *Applied Mechanics and Materials*, vol. 672, pp. 1400-1404, 2014.

[11] D. Fu and L. Fang, ”Blockchain-based trusted computing in social network,” *2016 2nd IEEE International Conference on Computer and Communications (ICCC), Chengdu, China*, pp. 19-22, 2016, doi: 10.1109/CompComm.2016.7924656.

[12] S. Gao, T. Yu, J. Zhu and W. Cai, ”T-PBFT: An EigenTrust-based practical Byzantine fault tolerance consensus algorithm,” *China Communications*, vol. 16, no. 12, pp. 111-123, 2019.

[13] P. Gaži, A. Kiayias and D. Zindros, ”Proof-of-stake sidechains,” in *2019 IEEE Symposium on Security and Privacy (SP), San Francisco, CA, USA*, pp. 139-156, 2019, doi: 10.1109/SP.2019.00040.

[14] I. G. A. K. Gemeliarana and R. F. Sari, ”Evaluation of proof of work (POW) blockchains security network on selfish mining,” in *2018 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), Yogyakarta, Indonesia*, pp. 126-130, 2018, doi: 10.1109/ISRITI.2018.8864381.

[15] H. Guo, X. Yu, ”A survey on blockchain technology and its security,” *Blockchain: research and applications*, vol. 3, no. 2, pp. 100067, 2022.

[16] H. M. Hussien, S. M. Yasin, N. I. Udzir, M. I. H. Ninggal, S. Salman, ”Blockchain technology in the healthcare industry: Trends and opportunities,” *Journal of Industrial Information Integration*, vol. 22, pp. 100217, 2021.

[17] M. Javaid, A. Haleem, R. P. Singh, S. Khan, R. Suman, ”Blockchain technology applications for Industry 4.0: A literature-based review,” *Blockchain: Research and Applications*, vol. 2, no. 4, pp. 100027, 2021.

[18] A. Jisi and S. Yin, ”A new feature fusion network for student behavior recognition in education,” *Journal of Applied Science and Engineering*, vol. 24, no. 2, pp. 133-140, 2021.

[19] J. Kang, J. Zhang, W. Song, X. Yang, ”Friend relationships recommendation algorithm in online education platform,” in *Web Information Systems and Applications: 18th International Conference, WISA 2021, Kaifeng, China, September 24-C26, 2021, Proceedings 18. Springer International Publishing*, pp. 592-604, 2021.

[20] J. Lauinger, J. Ernstberger, E. Regnath, M. Hamad and S. Steinhorst, ”A-PoA: Anonymous proof of authorization for decentralized identity management,” in *IEEE International Conference on Blockchain and Cryptocurrency (ICBC), Sydney, Australia*, pp. 1-9, 2021, doi: 10.1109/ICBC51069.2021.9461082.

[21] T. Li, H. Wang, D. He and J. Yu, ”Blockchain-based privacy-preserving and rewarding private data sharing for IoT,” *IEEE Internet of Things Journal*, vol. 9, no. 16, pp. 15138-15149, 2022.

[22] A. Y. H. Liao, Y. Y. Hsieh, C. Y. Yang, M. S. Hwang, “Research on a trustworthy digital learning roll call system”, *International Journal of Network Security*, vol. 24, no. 4, pp. 681-688, 2022.

[23] M. K. Lim, Y. Li, C. Wang, M. L. Tseng, ”A literature review of blockchain technology applications in supply chains: A comprehensive analysis of themes, methodologies and industries,” *Computers & industrial engineering*, vol. 154, pp. 107133, 2021.

[24] C. Y. Lin, L. C. Huang, Y. H. Chen, M. S. Hwang, “Research on security and performance of blockchain with innovation architecture technology”, *International Journal of Network Security*, vol. 23, no. 1, pp. 1-8, 2021.

[25] C. W. Liu, W. F. Hsien, C. C. Yang, and M. S. Hwang, “A survey of public auditing for shared data storage with user revocation in cloud computing”, *International Journal of Network Security*, vol. 18, no. 4, pp. 650-666, 2016.

[26] H. Liu, Y. Liu, L. Xu, S. Abdullah, ”Multi-attribute group decision-making for online education live platform selection based on linguistic intuitionistic cubic fuzzy aggregation operators,” *Computational and Applied Mathematics*, vol. 40, pp. 1-34, 2021.

[27] J. Liu, X. Xu, F. Zhang, Y. Gao and W. Gao, ”Modeling of spatial distribution characteristics of high proportion renewable energy based on complex principal component analysis,” *2020 IEEE Sustainable Power and Energy Conference (iSPEC), Chengdu, China*, pp. 193-198, 2020, doi: 10.1109/iSPEC50848.2020.9351293.

[28] Z. P. Peng, M. L. Chiang, I. C. Lin, C. C. Yang, M. S. Hwang, “CBP2P: Cooperative electronic bank payment systems based on blockchain technology”, *Journal of Internet Technology*, vol. 23, no. 4, pp. 683-692, 2022.

[29] L. Qu, Z. Wang, C. Li, S. Guo, J. Xin, Y. Zhou, W. Wang, ”Dynamic Bayesian network modeling based on structure prediction for gene regulatory network,” *IEEE Access*, vol. 9, pp. 123616-123634, 2021.

[30] A. S. Rajasekaran, M. Azees, F. Al-Turjman, ”A comprehensive survey on blockchain technology,” *Sustainable Energy Technologies and Assessments*, vol. 52, pp. 102039, 2022.

[31] A. C. Sales, E. B. Prihar, J. A. Gagnon-Bartsch, N. T. Heffernan, ”Using auxiliary data to boost precision in the analysis of A/B tests on an online educational platform: New data and new results,” *arXiv preprint arXiv:2306.06273*, 2023.

[32] P. William, A. Choubey, G. S. Chhabra, R. Bhattacharya, K. Vengatesan and S. Choubey, ”Assessment of hybrid cryptographic algorithm for secure sharing of textual and pictorial content,” in *2022 International Conference on Electronics and Renewable Systems (ICEARS), Tuticorin, India*, pp. 918-922, 2022.

[33] C. C. Wu, C. Y. Yang, M. S. Hwang, C. Y. Lee, “A trustworthy web-based system platform for teaching evaluation and STEM education”, *The International Journal of Electrical Engineering & Education*, 2019. (`https://doi.org/10.1177/0020720919853427`)

[34] C. C. Wu, C. Y. Yang, M. S. Hwang, M. Y. Lin, “The design and application of a web-based teacher evaluation system for STEM education”, in *The International Journal of Electrical Engineering & Education*, 2019. (`https://doi.org/10.1177/0020720919852783`)

[35] C. Y. Yang, T. Y. Chung, M. S. Hwang, C. Y. Li, J. F. Yao, “Learning performance evaluation in eLearning with the web-based assessment”, in *Information Science and Applications*, pp. 645-651, 2017.

[36] M. Yang, I. Tjuawinata, K. Y. Lam, J. Zhao and L. Sun, ”Secure hot path crowdsourcing with local differential privacy under fog computing architecture,” *IEEE Transactions on Services Computing*, vol. 15, no. 4, pp. 2188-2201, 2022.

[37] F. Ye, Y. Mao, Y. Li and X. Liu, ”Target threat estimation based on discrete dynamic Bayesian networks with small samples,” *Journal of Systems Engineering and Electronics*, vol. 33, no. 5, pp. 1135-1142, 2022.

[38] S. Yin, H. Li, A. A. Laghari, et al. , ”A bagging strategy-based kernel extreme learning machine for complex network intrusion detection,” *EAI Endorsed Transactions on Scalable Information Systems*, vol. 21, no. 33, 2021.

[39] J. Zhou, ”Deep learning-driven distributed communication systems for cluster online educational platform considering human-Ccomputer interaction,” *International Journal of Communication Systems*, vol. 35, no. 1, pp. e5009, 2022.

Biography

Daihong Feng biography. Daihong Feng is with Art School affiliated to Shenyang Normal University. Research interests are Information security, Intelligent education, artificial intelligence+computer education.

Hang Li biography. Hang Li is with Software College, Shenyang Normal University. Research interests are Information security, Computer network, information security, digital image encryption processing, artificial intelligence, computer education.

CSDN: A Compressed Sensing Dynamic Network-based on Deep Learning for Nonlinear Scrambling Diffusion Synchronous English Image Encryption

Qingxiang Duan
(Corresponding author: Qingxiang Duan)

School of Foreign Languages, Zhengzhou University of Science and Technology
No. 1 Xueyuan Road, Mashai Industrial Park, Erqi District, Zhengzhou, 450064 China
Email: duanqingxx2023@163.com

(Received Sept. 21, 2023; Revised and Accepted Dec. 1, 2023; First Online Dec. 24, 2023)

The Special Issue on Data Fusion, Deep Learning and Optimization Algorithms for Data Privacy Protection

Special Editor: Prof. Shoulin Yin (Harbin Institute of Technology)

Abstract

At present, most image encryption algorithms directly encrypt the plaintext image into the ciphertext image without visual significance, and this encrypted image is easy to be discovered by hackers in the transmission process, so it is subject to various attacks. The encryption process is weak and nonlinear, resulting in poor algorithm security. Therefore, in this paper, we propose a compressed sensing dynamic network based on deep learning for nonlinear scrambling diffusion synchronous English image encryption. Compressed sensing (CS) is used to preprocess the image. In the preprocessing, the measurement matrix is constructed by Kronecker product (KP), and the original image is scaled down. Secondly, a new sine-cos chaotic map is constructed to broaden the range of control parameters and improve the randomness of sequence distribution. Then, the XOR sum of the plaintext pixel and chaotic sequence is used as the initial chaotic value to generate the chaotic sequence, which is used to construct the network structure of different plaintext pixels, and the diffusion value is used to update the network value to make the network dynamic dynamically. Finally, a single-pixel serial scramble-diffusion is used to make the scrambles and diffuses cross, and the scrambles and diffuses are synchronized to resist the separation attack effectively. Experimental results show that the new algorithm has high encryption security and strong plaintext sensitivity, and it is particularly effective in resisting statistical attacks, differential attacks, and selective plaintext attacks.

Keywords: Compressed Sensing; Deep learning; Dynamic Network; Image Encryption; Nonlinear Scrambling Diffusion Synchronous

1 Introduction

With the popularization of the Internet, people will transmit all kinds of data through the network, and the risk of information leakage also increases. Image encryption is one of the technologies to provide security for multimedia data [4, 7, 15, 21]. Because chaotic system is sensitive to initial value and chaotic sequence is pseudorandom, it is very suitable for image encryption. Cavusoglu *et al.* proposed an image encryption scheme based on chaotic S-box. Hashimoto *et al.* [11] combined Logistic mapping and Chua circuit to define complex mapping and generate chaotic sequences for encryption. Zhou *et al.* [30] used fractional chaos for image encryption. These encryption schemes are only designed from the perspective of information security, without considering the characteristics of the image itself.

There are many similarities between chaos and cryptography [10]. Therefore, chaos is widely used in image encryption, and many image encryption based on chaos have been proved to have security problems. In reference [3], a 3D matrix replacement mechanism was proposed, which enabled each person to move to any position to obtain the scrambled bit matrix [5]. However, the encryption process had nothing to do with plaintext, and the scrambling of different images was the same as the diffused key stream. Wang *et al.* [24] used mathematical language analysis and summarized the main properties of such encryption algorithms, showing that there was a correspondence between the plaintext and plaintext of the plaintext independent algorithm, which could be selected to crack the plaintext attack [14, 27]. To solve this problem, plaintext correlation algorithm is proposed. Alghafis *et al.* [2] proposed an image encryption algorithm based on bit-level arrangement and Deoxyribo Nucleic Acid (DNA) coding, in which plaintext and chaotic system were used to generate sequences. Liu *et al.* [20] extended Arnold mapping and proposed a new encryption algorithm, in which both scrambling and diffusion operations were related to plaintext. Song *et al.* [22] proposed a parallel fast image encryption algorithm, in which the diffusion process was related to the plaintext and could be encrypted in parallel to reduce the time complexity of the algorithm.

Gopalakrishnan *et al.* [9] proposed a single-round diffusion encryption algorithm based on hyperchaotic system, which used the sum of positioned pixels as the plaintext feature value. Xie *et al.* [25] first scrambled the plaintext pixel position, then cross-embedded the binary into another two groups of chaotic sequences, and finally performed the XOR operation on the two groups of chaotic sequences. Talhaoui *et al.* [23] proposed an image encryption algorithm with variable block size. The algorithm first used Arnold transform to scramble the whole, then blocked according to Baker mapping, and finally used pixel and to encrypt the blocks. But some plaintext correlation algorithms also have security problems. Hedayati *et al.* [12] analyzed the algorithm proposed in reference [23], and used the selective plaintext attack to crack successfully, and obtained the key stream to decrypt the ciphertext graph of other plaintexts without the key. Liang *et al.* [17] used special images to restore the scrambled image, and successfully cracked it through multiple special images. Ahmad *et al.* [1] conducted an in-depth analysis of the algorithm in reference [25] and found that the generation of chaotic sequences was closely related to the average value of plaintext pixels. Firstly, the selective plaintext attack was used to obtain the diffusion equivalent key, and then the special image was used to obtain the scrambled equivalent key, whose equivalent key could decrypt other images encrypted with the same key. The decrypted plaintext correlation algorithm has the problem of scrambling diffusion separation and independent operation. When the plaintext image is a special image, the security of the encryption algorithm only depends on the diffusion operation. Once the diffusion is cracked, the plaintext statistics will be exposed. Therefore, some researchers put forward the scrambling diffusion synchronization algorithm. Zhao *et al.* [29] proposed a simultaneous scrambling and diffusion encryption algorithm based on DNA. When using chaotic sequences to obtain new positions, DNA sequences and DNA algorithms were used for pixel diffusion. Liug *et al.* [19] proposed a symmetric color image encryption algorithm, which transformed the color image into a three-dimensional matrix, obtained the scrambled position through chaotic mapping, and synchronized plaintext related diffusion operations. However, the above scrambling and diffusion synchronization algorithm has low nonlinearity, and the diffusion proceeds according to the inherent order. The scrambling position information remains unchanged before and after different plaintext diffusion, which is easy to cause the exposure of the diffusion and scrambling information. Jasra *et al.* [13] pointed out that if diffusion operated in a fixed order and the pre-ciphertext value was used to increase the avalanche effect, the attacker could obtain the variables in the diffusion decoding equation through the diffusion order, simplify the diffusion equation, and facilitate the attack. They carried out a strict analysis and suggested that nonlinear diffusion was introduced into image encryption. Yang *et al.* [26] proposed a general cracking algorithm by taking advantage of the relative position invariability of different plaintext scrambling in encryption, and also suggested adding nonlinearity in scrambling.

Aiming at the above problems, this paper proposes a scrambling diffusion synchronous encryption algorithm based on compressed sensing and strong nonlinearity. The contributions of this paper are as follows:

1) The CS is used to preprocess the image. In the preprocessing, the measurement matrix is constructed by Kronecker product (KP) and the original image is scaled down.

2) A novel sine-cos chaotic map is constructed to broaden the range of control parameters, improve sequence distribution and randomness, and be more suitable for encryption.

3) The combination of plaintext pixel and chaotic map makes the structure of different plaintext key streams and pixel networks different, and the network values are updated dynamically to make the network change dynamically, so as to achieve the effect of one picture one dense.

4) It uses adjacent node pixels for diffusion to make the diffusion process dynamic and increase the plaintext relevance of the algorithm. Moreover, single-pixel diffusion and scrambling are carried out alternately, which makes it impossible to separate the scrambling and spreading and synchronizes the scrambling and spreading.

5) The pixel operation is transferred according to the network structure, resulting in the network structure of the encryption path, which has strong nonlinear, so that the attacker can not know the diffusion order and scrambling order according to the clear (secret) text.

2 Preliminaries

2.1 Novel Sine-cos Chaotic Map

Sine chaotic map and kent chaotic map are simple and easy to implement, and are widely used in image encryption. sine and kent chaotic maps are:

$$z_{n+1} = \phi sin(\pi z_n) \tag{1}$$

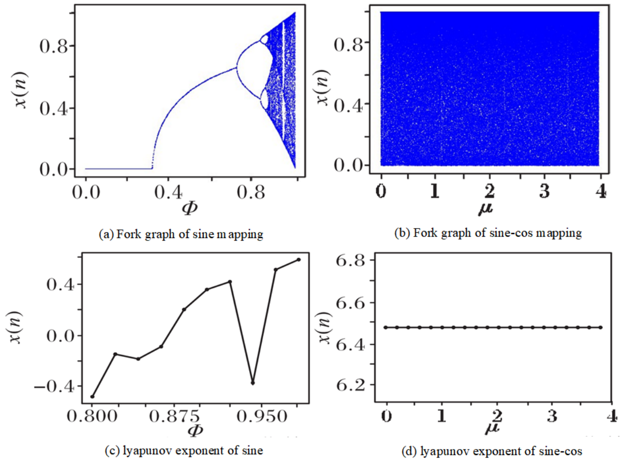
$$y_{n+1} = \begin{cases} y_n/\sigma, y_n \in (0, s] \\ (1-y_n)/(1-\sigma) \end{cases} \tag{2}$$

Where the parameters $\phi \in (0.85, 1)$, $\sigma \in (0, 1)$, the sequence $z, y \in (0, 1)$.

To overcome the defects that sine mapping key space is small, stable area is narrow and blank, a new chaotic system sine-cos is obtained by modifying sine. sine-cos chaotic mapping is:

$$x_{n+1} = cos(\pi x_n sin(\mu exp(10^{x_n})\pi/2)/2) \qquad (3)$$

The fork of sine chaotic map and sine-cos chaotic map with lyapunov exponent is shown in Figure 1. The parameter μ range of sine-cos chaotic mapping is wider than that of sine mapping, reaching the whole set of positive numbers. There is no blank region in chaotic region. Meanwhile, the lyapunov exponent is stable, which is more suitable for image encryption.



(a) Fork graph of sine mapping

(b) Fork graph of sine-cos mapping

(c) lyapunov exponent of sine

(d) lyapunov exponent of sine-cos

Figure 1: Bifurcation of sine chaotic mapping and sine-cos chaotic mapping and lyapunov exponential map

2.2 Compressed Sensing Technology

Compressed sensing is a signal sampling technique that compresses data simultaneously during the sampling process [24]. Image is a kind of two-dimensional or three-dimensional signal, and the algorithm in this paper deals with two-dimensional gray image. Assuming that the size of the original image I is $N \times N$, if I can be sparsely represented, then:

$$CS - I = \Phi_1 I \Phi_2^T \qquad (4)$$

Where Φ_1 and Φ_2 are measurement matrices with size $M \times N$, corresponding to the subsampling process. The measurement matrix projects the high-dimensional signal I into the low-dimensional space. $CS - I$ is the measurement image, and the size is $M \times N$, which is the result after sub-sampling. In this algorithm, the measurement matrix is composed of random Gaussian matrix.

In Equation (4), since $M < N$, under the condition that the measurement matrix Φ_1, Φ_2 and the measurement image $CS-I$ are known, the original image I cannot be reconstructed accurately. To accurately reconstruct the original image I, the following two conditions should be met:

1) The original image I is sparse. The original image I is sparse represented on some sparse basis. Fourier transform basis, wavelet transform basis and discrete cosine transform basis are commonly used for sparse decomposition. In this paper, wavelet transform basis is used for sparse representation of image signals, as shown in Equation (5):

$$S = \Psi I \Psi^T \qquad (5)$$

Where Ψ is a sparse basis matrix with size $N \times N$. S is the sparse coefficient matrix with size $N \times N$.

2) Φ_1 and Φ_2 must meet Restricted Isometry Property (RIP). However, it is very difficult to judge whether the measurement matrix has RIP property. Related studies show that the equivalence condition of RIP is that the measurement matrices Φ_1 and Φ_2 are not correlated with the sparse basis Ψ. Therefore, the sparse image S can be transformed into a minimization problem for solving, as shown in Equation (6):

$$min||S||_0, subject \quad to \quad CS - I = \Phi_1 S \Phi_2^T. \qquad (6)$$

Where $||S||_0$ is the number of non-zero vectors in sparse coefficient matrix S.

Signal reconstruction is the process of using some algorithms to recover high-dimensional original signals from low-dimensional measured signals. The selected measurement matrix and reconstruction algorithm will directly affect the quality of the reconstructed signal. The measurement matrix used in this paper is random Gaussian matrix, and the reconstruction algorithm is Orthogonal Matching Pursuit (OMP) algorithm.

2.3 Kronecker Product

Kronecker product is a special form of tensor product, which represents the operational relationship between matrices of any two sizes. If A is the matrix of $m \times n$ and B is the matrix of $p \times q$, then $A \otimes B$ is a block matrix of $mp \times nq$, as shown in Equation (7):

$$A \otimes B = [a_{11}B, \cdots, a_{1n}B] \cdot \cdots \cdot [a_{m1}B, \cdots, a_{mn}B] \qquad (7)$$

If A and B are linearly independent orthogonal matrices, according to the properties of KP, the linearly independent orthogonal matrix of lower dimension can be obtained from the linearly independent orthogonal matrix of higher dimension. In this paper, CS measurement matrix is constructed by using this method.

3 Proposed Encryption and Decryption System

3.1 Image Compression

According to Equation (4), the original image is sparse decomposed by wavelet transform and represented in the frequency domain. Arnold transform is applied to the sparsified image. The representation of Arnold transform is shown in Equation (8):

$$[V \quad K]^T = [1 \quad 1]^T [1 \quad 2]^T [v \quad k]^T (mod \quad L) \tag{8}$$

Where L is the height of the sparsified image. (V, K) is the position coordinates of the pixels of the sparsified image. (v, k) is the position coordinates of pixels of the transformed sparse image.

In this paper, the Kronecker product property is used to construct the measurement matrix, and the proposed compressed sensing framework is shown in Equation (9):

$$CS - I = (\Phi_1 \otimes P)I - A(\Phi_2^T \otimes P) \tag{9}$$

Where $I - A$ is the Arnold transform image after sparsification. P is the random Gaussian matrix of $t \times t$. t is the dimensionality reduction multiple, as shown in Equation (10):

$$t = \frac{M}{m} = \frac{N}{n} \tag{10}$$

The low-dimensional measurement matrices Φ_1 and Φ_2 can be extended to high-dimensional measurement matrices by random Gaussian matrix P. According to Equation (6), Equation (11) can be obtained:

$$\Phi_1 \otimes P = [\varphi_{11}P, \cdots, \varphi_{1n}P]^T \cdots \cdots [\varphi_{m1}P, \cdots, \varphi_{mn}P]^T \tag{11}$$

So the size of $\Phi_1 P$ is $(m \times t) \times (n \times t)$, which is $M \times N$. Assuming that the original image size is 256×256 and the image compression rate $f = M/N = m/n = 0.75$, the dimension of the traditional compressed sensing measurement matrix is 192×256. In this paper, the dimension of the measurement matrix is reduced, and the dimension of the measurement matrix is 48×64 when the dimensionality reduction factor $t = 4$, which greatly reduces the storage space required by the measurement matrix and improves the transmission efficiency. According to Equation (9), a compressed $M \times N$ measurement image is obtained.

3.2 Encryption Process

The compressed plaintext image P is a 256-level grayscale image with $n \times m$. The public key of the encryption algorithm is σ, μ, μ', x_0, y_0. The encryption process includes pixel network construction and single-pixel serial scrambling and spreading. The specific encryption process is shown in Figure 2. The decryption public key is σ, μ, μ', x_0, y_0, S. S is XOR sum.

3.3 Encryption System

A. Building a pixel network

Step 1. Substitute the parameters μ and x_0 into sine-cos map and iterate $n \times m + 1000$ times, remove the first 1000 terms to obtain $L_x = L_{x_1}, L_{x_2}, \cdots, L_{x_{n\times m}}$.

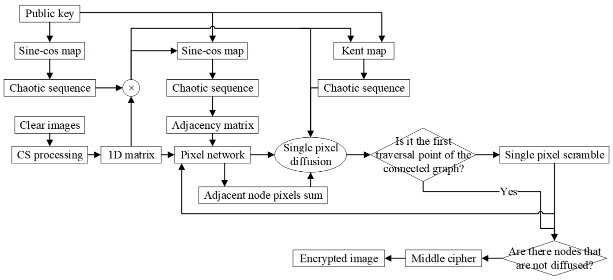


Figure 2: Encryption process

Step 2. Convert plaintext P into one-dimensional sequence $P' = P'_1, P'_2, \cdots, P'_{n\times m}$ according to line priority, and then obtain pixel XOR sum S through Equation (12).

$$S = \sum_{i=1}^{n\times m} bitxor(P'_i, L_{x_i}) \tag{12}$$

Where *bitxor* is the XOR operation.

Step 3. Use Equation (13) to find x', then substitute the parameter μ' and x' into sine-cos mapping and iterate $n \times m + 1000$ times, and remove the first 1000 terms to get $L_{x'} = L_{x'_1}, L_{x'_2}, \cdots, L_{x'_{n\times m}}$.

$$x' = mod(S/10^{14}, 1) \tag{13}$$

Step 4. Quantize $L_{x'}$ through Equation (14) to get integer sequence $K = K_1, K_2, \cdots, K_{n\times m}$, and then use Equations (4) and (16) to get $link = link_1, link_2, \cdots, link_{n\times m}$. Take P' as the node matrix, $link$ as the adjacency matrix, and generate the network G.

$$K_i = mod(L_{x'_i} \times 10^{14}, n \times m) \tag{14}$$

$$link_i = cat(link_i, K_i) \tag{15}$$

$$link_{K_i} = cat(link_{K_i}, i) \tag{16}$$

Where *mod* is the coremainder formula. *cat* is the merge operation.

B. Single pixel serial scrambling and diffusion

Step 1. Substitute the S into Equation (17) and y, σ into Equation (8) and iterate $n \times m + 1000$ times, remove the first 1000 terms to obtain $L_y = L_{y_1}, L_{y_2}, \cdots, L_{y_{n\times m}}$. L_y is quantized through Equation (18) and the integer sequence $l = l1_{,l_2,\cdots,l_{n\times m}}$ is obtained.

$$y = \mod(y_0 \times 10^{14}/S, 1) \tag{17}$$

$$l_i = \mod(l_{y_i} \times 10^{14}/S, 256) \tag{18}$$

Step 2. Randomly obtain node G_i from the network.

Step 3. Calculate the pixels sum T of adjacent nodes of G_i, quantified as t according to Equation (19).

$$t = mod(T, 256) \tag{19}$$

Step 4. Use Equation (20) to diffuse G_i and obtain C_i.

$$C_i = bitxor(t, bitxor(G_i, mod(S + l_i, 256))) \tag{20}$$

Step 5. Obtain C_i from Step 4 and switch with the neighboring source node C'. Update G_i, G' with C, C'. If G_i is the first traversal point of a connected graph, there is no exchange and update.

Step 6. Nodes C_i and G_i serve as C' and G'. Determine whether G_i has adjacent nodes that have not been scrambled or diffused. If yes, the adjacent node G_{i+1} of G_i is randomly obtained. If not, judge whether the connected graph of G_i has unprocessed nodes. If there are unprocessed nodes, the nearest node G_{i+1} is obtained. Otherwise, check whether there are unprocessed nodes in the network, obtain node G_{i+1} randomly, and perform Step 3 to Step 5 on node G_{i+1} to obtain intermediate ciphertext C.

Step 7. Convert C to the $m \times n$ ciphertext matrix C'.

3.4 Decryption System

Step 1. Use chaotic sequence and ciphertext to generate network G, and then perform breadth first traversal on network G to obtain access order L. Restore pixel nodes in reverse order of L.

Step 2. The node switches with the access source node. If it is the last pixel of the connected graph, the node does not switch.

Step 3. Calculate the sum of adjacent pixels of the node and convert it into the value of $0-255$. Then, use Equation (21) to carry out diffusion.

$$P'_i = bitxor(mod(S + l, 256), bitxor(G_i, t)) \tag{21}$$

Step 4. Continue to perform Steps 2 to 3 until there are no unprocessed pixels, and finally get P'.

Step 5. Convert P' into a two-dimensional plaintext matrix P.

3.5 Experimental Results and Safety Analysis

In order to verify the security and effectiveness of the algorithm, 512×512 standard gray scale is selected as the experimental object. Set the public key to $\sigma = 0.59999$, $\mu = 3.9994$, $\mu' = 3.8884$, $x_0 = 0.2375$, $y_0 = 0.3734$. The encryption and decryption effect of the algorithm is shown in Figure 3. All information in plaintext is hidden. The decrypted image is the same as the plaintext image.

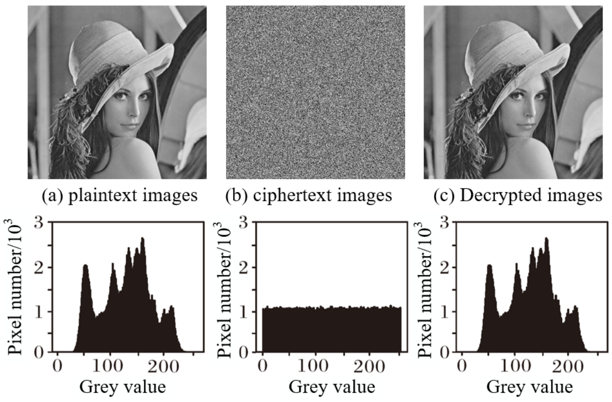


Figure 3: Encryption, decryption effects and corresponding histograms

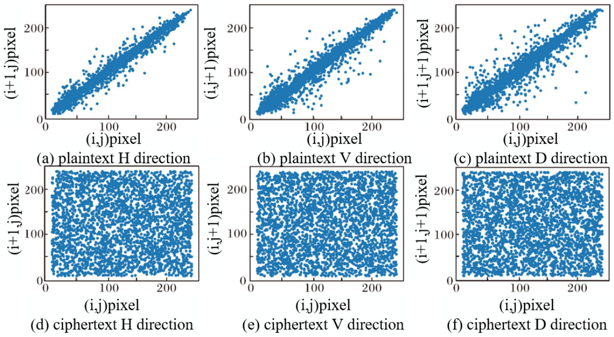
3.6 Statistical Characteristic Analysis

A. Histogram analysis

Image pixel distribution can be described by histogram. The distribution of pixel value and corresponding histogram of the encryption algorithm against statistical attack is uniform. As shown in Figure 3, the ciphertext histogram is evenly distributed and significantly different from the plaintext histogram. Therefore, the ciphertext pixels are evenly distributed, and the attacker cannot obtain plaintext statistical information from the ciphertext image, which can resist statistical attacks.

B. Adjacent pixel correlation

The adjacent pixels in the plaintext image generally have strong correlation. A secure encryption algorithm should be able to break the strong correlation and make the adjacent pixels of the ciphertext have no correlation. At the same time, the correlation coefficient r of adjacent pixels can quantitatively describe the correlation of adjacent pixels. To visualize the correlation, 4000 pairs of adjacent pixels were randomly selected in the horizontal (H), vertical (V) and diagonal (D) directions of Ming ciphertext to obtain the correlation distribution, as shown in Figure 4.



(a) plaintext H direction (b) plaintext V direction (c) plaintext D direction

(d) ciphertext H direction (e) ciphertext V direction (f) ciphertext D direction

Figure 4: Correlation of adjacent pixels

The adjacent pixel pairs of the plaintext image are distributed diagonally, and the ciphertext is evenly distributed on the plane. The correlation between the plaintext pixels is high, but the correlation between the ciphertext pixels is not high. The correlation coefficients of plaintext and ciphertext are shown in Table 1. The plaintext correlation coefficient is close to 1, while the ciphertext correlation coefficient is close to 0. The average correlation coefficient in Set14 dataset also has the same characteristics, which further verifies the low correlation of adjacent pixels in ciphertext.

Table 1: Pixel correlation coefficient of ciphertext image

Direction	H	V	D
plaintext (Lena)	0.9692	0.9841	0.9557
ciphertext (Lena)	0.0008	4.4×10^{-5}	-0.0006
plaintext (pepper)	0.9756	0.9808	0.9626
ciphertext (pepper)	0.0036	-0.0012	-0.0008
ciphertext Set14	-0.0004	0.0007	-0.003

C. Information entropy

Information drops reflect the uncertainty and randomness of image pixels. The greater information entropy denotes the stronger randomness. The calculation formula of information entropy is shown in Equation (22):

$$H(x) = -\sum_{i=1}^{n} p(x_i) lbp(x_i) \tag{22}$$

Where $p(x_i \in (0,1))$, and the sum of $p(x_i$ is 1.

If the image pixels of 256 gray levels are evenly distributed, the information entropy is close to the theoretical maximum value of 8. The entropy values of plaintext and ciphertext are shown in Table 2.

Table 2: Information entropy (IE)

Image	IE
Lena	7.9994
pepper	0.79993
Set14	7.9989

The ciphertext information entropy is close to the theoretical maximum value of 8, and the average information entropy in Set14 data set is also close to 8, indicating that the ciphertext has stronger uncertainty and randomness.

3.7 Differential Analysis

A. Plaintext sensitivity analysis

The Number of Pixel Change Rate (NPCR) and Unified Average Changing Intensity (UACI) can quantitatively evaluate the difference between two images. The larger NPCR value denotes that the encryption algorithm is more sensitive to the changes of the original image. The larger UACI value denotes the greater average intensity of the image change. NPCR and UACI are defined as follows:

$$NPCR = (\frac{\sum_{i=1}^{n} P_i}{M \times N}) \times 100/\% \tag{23}$$

$$UACI = \frac{(C(1)_i - C(2)_i)/256}{M \times N} \times 100/\% \tag{24}$$

Where $C(1)$ and $C(2)$ are ciphertext images. When $C(1)_i = C(2)_i$, $P_i = 0$, otherwise $P_i = 1$.

In this paper, the plaintext sensitivity of the algorithm is analyzed by the NPCR value and UACI value of the ciphertext image with one random pixel adding 1 and the ciphertext of the image with two random pixels exchanged. The results are shown in Table 3. Both NPCR and UACI reached 99.6% and 33.4% respectively, indicating that almost all the pixel values of ciphertext with any change show the strong plaintext sensitivity.

Table 3: NPCR and UACI values%

Change way	index	Lena	pepper	baboon
pixel+1	NPCR	99.62	99.61	99.61
pixel+1	UACI	33.44	33.44	33.42
Switch any two pixels	NPCR	99.64	99.62	99.61
Switch any two pixels	UACI	33.43	33.44	33.43

B. Key sensitivity analysis

The secure encryption algorithm is sensitive to the subtle change of the key, and the decryption effect of the key with slight difference is also different. When the decryption key changes slightly, the decryption effect is shown in Figure 5.

It can be seen from Figure 5 that the plaintext image cannot be restored when the difference between key μ and x_0 is 10^{-16}. When the difference between key μ and x_0 is 10^{-17}, the plaintext image can be restored, so the sensitivity of μ and x_0 is 10^{-16}. Similarly, the sensitivity of σ, y_0 and μ' is 10^{-16}. It can be seen that the new algorithm has strong key sensitivity.

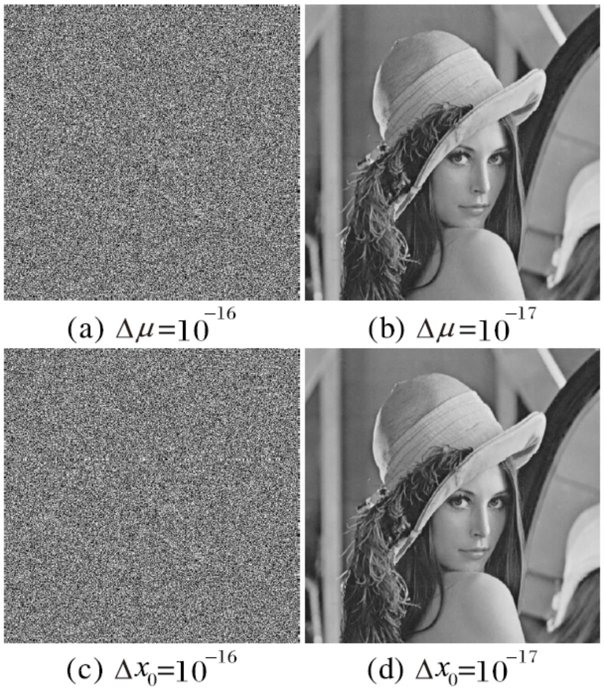
3.8 Key Space Analysis

The chaotic system parameters of the proposed algorithm are σ, σ', μ, x_0 and y_0 respectively, and the parameters are double precision data with 16 significant bits. The plaintext XOR sum S are also used as keys to further expand the key space. Therefore, the key space of the algorithm is at least 10^{96}. The proposed algorithm has sufficient security level to resist exhaustive attacks.

3.9 Attack Resistance Analysis

As an important step of cryptanalysis to evaluate the security of encryption algorithms, the anti-attack ability of the proposed algorithm is verified by special image experiment analysis, graph-by-graph analysis, stepwise attack analysis and nonlinear performance analysis.

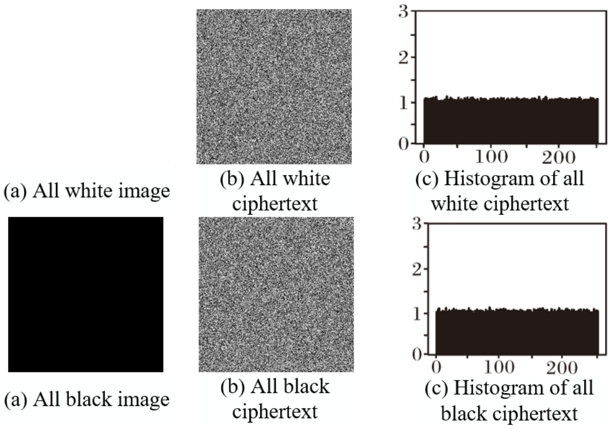
A. Experimental analysis of special images


(a) $\Delta\mu=10^{-16}$ (b) $\Delta\mu=10^{-17}$

(c) $\Delta x_0=10^{-16}$ (d) $\Delta x_0=10^{-17}$

Figure 5: Decrypted image with key deviation

In some image cryptanalysis, special images (all black or all white) are constructed to obtain the decryption information of the encryption algorithm. In this paper, all-black and all-white images are selected as the test objects of the algorithm, and the results are shown in Figure 6 and Table 4.


(a) All white image (b) All white ciphertext (c) Histogram of all white ciphertext

(a) All black image (b) All black ciphertext (c) Histogram of all black ciphertext

Figure 6: Experimental results of special images

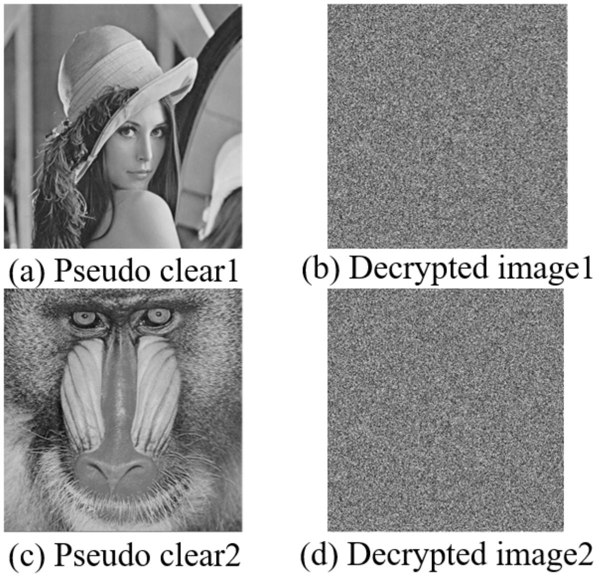
CC is correlation coefficient. It can be seen from Figure 6 and Table 4 that the low correlation of adjacent pixels, uniform histogram and ciphertext are noise-like, indicating that there is no obvious statistical information and other information in ciphertext of special image. The diffusion avalanche effect can be enhanced by network value updating and dynamic diffusion. At the same time, this paper uses XOR sum to associate with the plaintext, does not directly use the plaintext information, there is no special position, which can effectively resist special image analysis.

Table 4: NPCR and UACI values%

Image	IE	CC(H)	CC(V)	CC(D)
White	7.9993	0.0009	-0.0001	-0.0022
Black	7.9993	0.0006	0.0003	0.0012

B. One image one key analysis

The chaotic sequence generated during image encryption is obtained, the value of a pixel of the image is added by 1 as pseudoplaintext 1, and a different image is taken as pseudoplaintext 2 to encrypt the two images respectively. The obtained chaotic sequence is used for decryption, and the result is shown in Figure 7.


(a) Pseudo clear1 (b) Decrypted image1

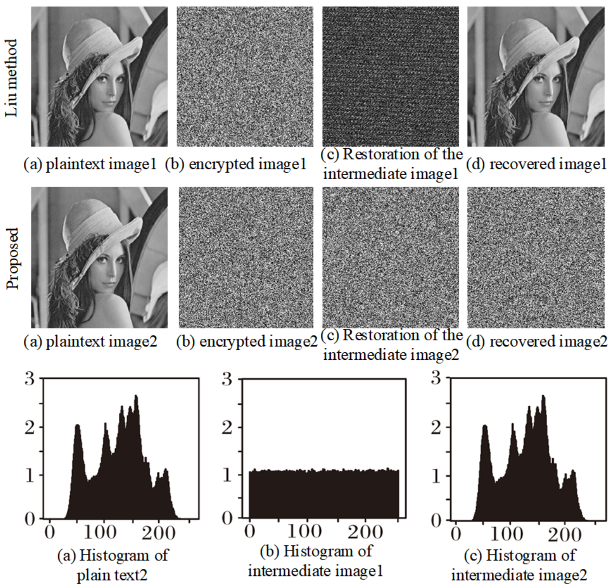
(c) Pseudo clear2 (d) Decrypted image2

Figure 7: One image one key analysis results

Pseudo plaintext 1 differs from the image by only one pixel, but no intermediate key can be used to restore any plaintext information. The pseudo-plaintext 2 which is far from it can not restore the plaintext information. In this paper, the XOR and of the plaintext pixel and chaotic sequence are used as chaotic initial values and diffusion confusion values. Even with the same chaotic sequence, as long as the plaintext images are different, the initial values and confusion values will be different, so that different plaintext has different intermediate keys and pixel network structure. The attacker cannot crack the encryption algorithm by obtaining other plaintext intermediate keys, so as to achieve the effect of one picture one secret.

C. Stepwise attack analysis

In order to verify the ability of the proposed algorithm to resist the stepwise attack, the stepwise cracking method is adopted as the test method, and the proposed algorithm is compared with Liu algorithm [19]. As shown in Figure 8, the two algorithms encrypt Lena image (a) and (e) respectively to obtain ciphertext (b) and (f). The intermediate image (c)(g) is obtained by the test method, and the special image is constructed to obtain the decrypted image (d)(h). The histograms of Lena and (c) and (g) are also given. The distributed attack results are shown in Figure 8.



(a) plaintext image1 (b) encrypted image1 (c) Restoration of the intermediate image1 (d) recovered image1

(a) plaintext image2 (b) encrypted image2 (c) Restoration of the intermediate image2 (d) recovered image2

(a) Histogram of plain text2 (b) Histogram of intermediate image1 (c) Histogram of intermediate image2

Figure 8: Step-by-step attack results

D. Nonlinear performance analysis

In order to verify the security of the nonlinearity of the algorithm proposed in this paper, the differential equation cracking method is used as the test method, as shown in Equation (25), and compared with the algorithm in reference [6].

$$\Delta C = C1 \oplus C2 = wb(M1 \oplus M2) = wb(\Delta M) \quad (25)$$

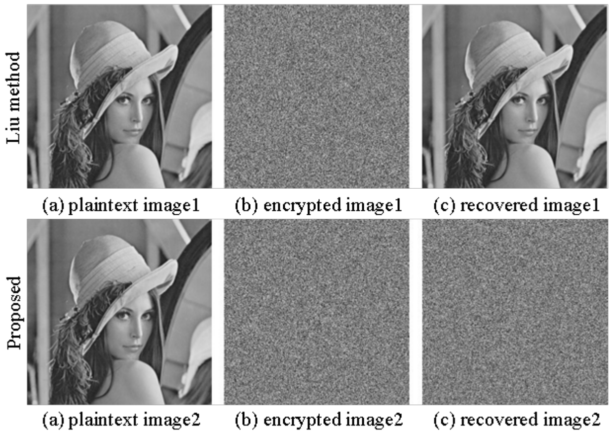
$$c_i = P_i \oplus t \oplus k_i \quad (26)$$

$$\Delta c_n = c_n \oplus c_{n-1} = P_n \oplus t_n \oplus k_n \oplus c_{n-1} \quad (27)$$

$$\Delta C = (h) = C(h) \oplus C(0) = \sum_{j=1}^{n} \Delta P_j'(h) \quad (28)$$

Where c_i is the former ciphertext pixel. P_i is the plaintext pixel. C_i is the ciphertext image. $M(i)$ is the plaintext image. wb is the scramble operation. k is the chaotic value. $C(i)$ is the ciphertext image. $\Delta C(h)$ is the XOR image.

As shown in Figure 9, the two algorithms encrypt the Lena image respectively, and then crack it with the selected attack method to get the ciphertext recovery image. The results show that the cracking method can recover the encrypted ciphertext of the algorithm in reference [23], but cannot recover the ciphertext of the new algorithm in this paper. It shows that the proposed algorithm can resist the cracking method. The method in [23] makes use of the feature that different plaintext scrambling in encryption has the same relative position, but the scrambling position information in this paper is locally adjusted by the network structure. According to the results of one figure-one encryption analysis, the network structure of different plaintexts is different, so the method proposed in this paper can resist the cracking methods in reference [23].



(a) plaintext image1 (b) encrypted image1 (c) recovered image1

(a) plaintext image2 (b) encrypted image2 (c) recovered image2

Figure 9: Attack test results

3.10 Encryption Time Analysis

As an index to evaluate the efficiency of the algorithm, the lower the encryption time, the more efficient the algorithm. The main time of the algorithm in this paper is spent on the extraction of network nodes [18, 28]. Since the adjacency list storage method is used in this paper, the number of vertices is the number of plaintext pixels, and the number of edges is twice the number of plaintext pixels, the time complexity of the algorithm is $O(3 \times n \times m)$, which is proportional to the size of the image.

Table 5 shows the encryption time of images with different sizes. As can be seen from Table 5, the encryption time of the proposed algorithm is low, and it is proportional to the image size, so it has certain high efficiency.

Table 5: Encryption time

Image	Time/s
Lena (128 × 128)	0.087
Lena (256 × 256)	0.337
Lena (512 × 512)	1.715
Set14	1.516

3.11 Comparison Analysis

To measure the performance of the algorithm in this paper, entropy, correlation coefficient, NPCR and UCAI are compared with the algorithms in references [8,16], and the results are shown in Table 6. As can be seen from Table 6, all indicators in this paper are better than the algorithms in references [8, 16], and have good performance on the whole.

Table 6: Encryption time

Method	[8]	[16]	Proposed
entropy	7.9992	7.9992	7.9993
CC(H)	0.0022	0.0017	0.0007
CC(V)	0.0008	0.0013	4.4×10^{-5}
CC(D)	0.0005	0.0009	-0.0006
NPCR	99.61	99.62	99.62
UCAI	33.38	33.41	33.42

4 Conclusion

In this paper, a highly secure image encryption algorithm with dynamic nonlinear encryption process and excellent anti-attack performance is proposed. Sine-cos chaotic mapping is designed to overcome the defects of sine chaotic mapping, such as narrow stable area and blank area, and increase the randomness of distribution, which is more suitable for image encryption. The initial value of chaotic map is generated by XOR sum of plaintext pixel and chaotic sequence, which makes the algorithm have one graph and one density. Chaotic and plaintext pixels generate a pixel network, and the network value updating mechanism makes the network change dynamically. Single-pixel alternate scrambling diffusion and pixel transfer operation according to the network structure make the path of scrambling diffusion synchronization be network structure, which makes the algorithm has strong nonlinearity and can resist the existing cracking methods. Based on the dynamic diffusion of the sum of adjacent nodes, the plaintext correlation is enhanced. Experimental results show that the algorithm has good encryption effect, no obvious statistical characteristics, strong plaintext correlation and anti-attack ability.

References

[1] I. Ahmad, S. Shin, ”A novel hybrid image encryption-compression scheme by combining chaos theory and number theory,” *Signal Processing: Image Communication*, vol. 98, 2021.

[2] A. Alghafis, F. Firdousi, M. Khan, S. I. Batool, M. Amin, ”An efficient image encryption scheme based on chaotic and Deoxyribonucleic acid sequencing,” *Mathematics and Computers in Simulation*, vol. 177, pp. 441-466, 2020.

[3] A. Brahim, A. Pacha, N. Said, ”Image encryption based on compressive sensing and chaos systems,” *Optics & Laser Technology*, vol. 132, 2020.

[4] C. C. Chang, K. F. Hwang, M. S. Hwang, “Robust authentication scheme for protecting copyrights of images and graphics”, *IEE Proceedings-Vision, Image and Signal Processing*, vol. 149, no. 1, pp. 43-50, 2002.

[5] T. Y. Chen, M. S. Hwang, and J. K. Jan, “Adaptive authentication schemes for 3D mesh models,” *International Journal of Innovative Computing, Information and Control*, vol. 5, no. 12, pp. 4561-4572, 2009.

[6] C. Chowdhary, P. Patel, K. Kathrotia, M. Attique, K. Perumal, M. F. Ijaz, ”Analytical study of hybrid techniques for image encryption and decryption,” *Sensors*, vol. 20, no. 18, pp. 5162, 2020.

[7] R. Dhanalakshmi, L. Kavisankar, S. Balasubramani, ”A Novel Technique using IoT Based Automated Irrigation System for Smart Farming,” *Journal of Applied Science and Engineering*, vol. 25, no. 4, pp. 641-648, 2021.

[8] M. Farah, A. Farah, T. Farah, ”An image encryption scheme based on a new hybrid chaotic map and optimized substitution box,” *Nonlinear Dynamics*, vol. 99, no. 4, pp. 3041-3064, 2020.

[9] T. Gopalakrishnan, S. Ramakrishnan, ”Image encryption using hyper-chaotic map for permutation and diffusion by multiple hyper-chaotic maps,” *Wireless Personal Communications*, vol. 109, no. 1, pp. 437-454, 2019.

[10] S. Han, E. Chang, T. Dillon, M. S. Hwang, C. C. Lee, “Identifying attributes and insecurity of a public-channel key exchange protocol using chaos synchronization”, *Chaos Solitons & Fractals*, vol. 40, no. 5, pp. 2569-2575, 2009.

[11] H. Hashimoto, H. Otsubo, H. Fujihara, S. Hitoshi, O. Toyoaki, Y. Toru, T. Yoshio, U. Yoichi, ”Centrally administered ghrelin potently inhibits water intake induced by angiotensin II and hypovolemia in rats,” *The Journal of Physiological Sciences*, vol. 60, no. 1, pp. 19-25, 2010.

[12] R. Hedayati, S. Mostafavi, ”A lightweight image encryption algorithm for secure communications in multimedia Internet of Things,” *Wireless Personal Communications*, vol. 123, no. 2, pp. 1121-1143, 2022.

[13] B. Jasra, A. Moon, ”Color image encryption and authentication using dynamic DNA encoding and hyper chaotic system,” *Expert Systems with Applications*, vol. 2061, 2022.

[14] M. Kamalesh, B. Chokkalingam, J. Arumugam, G. Sengottaiyan, S. Subramani, M. A. Shah, ”An intelligent real time pothole detection and warning system for automobile applications based on IoT technology,” *Journal of Applied Science and Engineering*, vol. 24, no. 1, pp. 77-81, 2021.

[15] M. Kaur, V. Kumar, ”A comprehensive review on image encryption techniques,” *Archives of Computational Methods in Engineering*, vol. 27, no. 1, pp. 15-43, 2020.

[16] X. Li, J. Mou, L. Xiong, Z. Wang, J. Xu, ”Fractional-order double-ring erbium-doped fiber laser chaotic

system and its application on image encryption," *Optics & Laser Technology*1, vol. 140, 202.

[17] Z. Liang, Q. Qin, C. Zhou, "An image encryption algorithm based on Fibonacci Q-matrix and genetic algorithm," *Neural Computing and Applications*, vol. 34, pp. 19313-19341, 2022.

[18] J. Liu, J. Zhang, S. Yin, "Hybrid chaotic system-oriented artificial fish swarm neural network for image encryption," *Evolutionary Intelligence*, vol. 16, pp. 77-87, 2023.

[19] X. Liu, X. Tong, Z. Wang, M. Zhang, "A new n-dimensional conservative chaos based on Generalized Hamiltonian System and its applications in image encryption," *Chaos, Solitons & Fractals*, vol. 154, 2022.

[20] X. Liu, D. Xiao, C. Liu, "Three-level quantum image encryption based on Arnold transform and logistic map," *Quantum Information Processing*, vol. 20, no. 1, pp. 1-22, 2021.

[21] M. Srinivas, M. Patnaik, "Fuzzy Extended Krill Herd Optimization with Quantum Bat Algorithm for Cluster Based Routing in Mobile Adhoc Networks," *Journal of Applied Science and Engineering*, vol. 25, no. 4, pp. 633-640, 2021.

[22] W. Song, C. Fu, M. Tie, C. Sham, J. Liu, H. Ma, "A fast parallel batch image encryption algorithm using intrinsic properties of chaos," *Signal Processing: Image Communication*, vol. 102, 2022.

[23] M. Talhaoui, X. Wang, "A new fractional one dimensional chaotic map and its application in high-speed image encryption," *Information Sciences*, 2021, 550: 13-26.

[24] X. Wang, Y. Su, "Image encryption based on compressed sensing and DNA encoding," *Signal Processing: Image Communication*, vol. 95, 2021.

[25] X. Yan, X. Wang, Y. Xian, "Chaotic image encryption algorithm based on arithmetic sequence scrambling model and DNA encoding operation," *Multimedia Tools and Applications*, vol. 80, no. 7, pp. 10949-10983, 2021.

[26] C. Yang, Y. Wang, P. Chen, S. You, "A Data Hiding Method Based on Partition Variable Block Size with Exclusive-or Operation on Binary Image," *International Journal of Informatics and Information Systems*, vol. 5, no. 1, pp. 1-15, 2022.

[27] S. Yin, H. Li, "GSAPSO-MQC:medical image encryption based on genetic simulated annealing particle swarm optimization and modified quantum chaos system," *Evolutionary Intelligence*, vol. 14, pp. 1817-1829, 2021.

[28] S. Yin, H. Li, L. Teng, A. A. Laghari, V. V. Estrela, "Attribute-based Multiparty Searchable encryption model for Privacy Protection of Text Data," *Multimedia Tools and Applications*, 2023. https://doi.org/10.1007/s11042-023-16818-4

[29] H. Zhao, S. Xie, J. Zhang, T. Wu, "A dynamic block image encryption using variable-length secret key and modified Henon map," *Optik*, vol. 230, 2021.

[30] Z. Zhou, W. Yu, "Studying Stochastic Resonance Phenomenon in the Fractional-Order Lorenz-Like Chaotic System," *International Journal of Bifurcation and Chaos*, vol. 32, no. 10, 2022.

Biography

Qingxiang Duan biography. Qingxiang Duan is with School of Foreign Languages, Zhengzhou University of Science and Technology. Her interests are content analysis, English data security analysis.

A Novel Four-dimensional Hyperchaotic System and DNA Encoding Method for Image Encryption

Jianjun Zhu and Jian'E Zhao

(Corresponding author: Jian'E Zhao)

School of Electronics and Electrical Engineering, Zhengzhou University of Science and Technology
Zhengzhou 450000 China
Email: xdwangxd@163.com

(Received Sept. 21, 2023; Revised and Accepted Dec. 1, 2023; First Online Dec. 24, 2023)

The Special Issue on Data Fusion, Deep Learning and Optimization Algorithms for Data Privacy Protection

Special Editor: Prof. Shoulin Yin (Harbin Institute of Technology)

Abstract

Aiming at the problems of poor speed and insufficient security with traditional image encryption algorithms when encrypting robot images, an image encryption algorithm based on a four-dimensional Hyperchaotic system and deoxyribonucleic acid (DNA) encoding is proposed. The newly constructed four-dimensional hyperchaotic system scrambles the image after bit decomposition to generate the intermediate ciphertext image. The randomness of the key sequence is ensured by selecting multiple high-dimensional chaotic systems and modifying the initial value of chaotic systems. Finally, DNA encryption is performed on the image using the key sequence, and the final ciphertext image is obtained. During DNA encryption, a DNA-S cassette is generated to make nonlinear substitutions to the DNA code. Simulation results show that the encrypted robot image by the proposed algorithm has better distribution characteristics, which can resist statistical analysis and differential attacks, and has the advantages of ample critical space and good sensitivity to initial values.

Keywords: Four-dimensional Hyperchaotic System; Image Encryption; Key Sequence

1 Introduction

The problem of information security in modern society has attracted much attention from experts. Because digital image is vivid, intuitive, so it is widely used, the research of encryption algorithm for digital image has become one of the hot spots. The main technological evolution of image encryption can be summarized into five categories [1,3,7,14,15,21].

1) Matrix transformation. Changing the original pixel position to realize image encryption;

2) Based on the transformation domain. The time domain is transformed to the frequency domain, the frequency domain image is encrypted, and then the inverse transformation is converted to the time domain to form the ciphertext image;

3) Based on chaos theory. According to the sensitivity of the initial value of chaotic sequence, image encryption is realized;

4) Based on DNA coding. According to the principle of base complementary pairing, image information is combined with DNA sequence to generate the corresponding ciphertext information;

5) Based on neural networks. Using neural networks to scramble pixel positions or replace pixel value to realize image encryption.

In order to improve the anti-attack ability and reduce the computational complexity, the advantages and disadvantages of the above five methods are considered comprehensively. This paper presents a new image encryption algorithm based on four-dimensional hyperchaos. The essential reason for chaotic-based encryption is that chaotic sequences are sensitive to initial values, leading to pseudorandomness and ergodicity of phase trajectories [8,13,16].

Traditional chaotic image encryption algorithms are mostly based on low-dimensional chaotic systems [9,11,22]. Because of the simple structure, small key space and low sequence complexity of the low-dimensional chaotic system, the encryption algorithm based on low-dimensional chaos has the advantages of fast running speed and high encryption efficiency, but the disadvantage is that the encryption security is not high.

Generally, compared with low-dimensional chaotic systems, hyperchaotic systems have more complex structure and higher sequence complexity. Therefore, the encryption algorithm based on hyperchaotic system has better encryption performance. For example, reference [27] proposed a hyperchaotic image encryption algorithm based on bit scrambling. Firstly, the chaotic sequence generated by Kent mapping was used to scramble the position of the plaintext pixel, and then the hyperchaotic sequence generated by Hyperhenon mapping was used to scramble the internal bit of each pixel, and finally the image pixel was diffused. Reference [18] proposed an image encryption algorithm based on bit-plane transformation. Firstly, the image was decomposed into 8 bit planes based on bits, and each bit plane was scrambled by a set of hyperchaotic sequences, and then merged into ciphertext images to realize image encryption. Comprehensive analysis of references [20, 24, 25] shows that they are all based on hyperchaotic systems and adopt bit-scrambling methods. However, the bit scrambling in reference [10] is limited to the pixel interior. The disadvantage is that bits cannot be exchanged between different pixels. In reference [23], bit scrambling is limited to the bit plane of each layer, and the disadvantage is that bits between different bit planes cannot be exchanged, which undoubtedly limits the degree of bit scrambling and further affects the effect of image encryption. Zhang *et al.* [29] proposed a parallel encryption algorithm based on a 5-dimensional hyperchaotic system, which divided image pixels into different levels and used parallel encryption between the same levels to improve the encryption speed. Cheng *et al.* [4] proposed a fast image encryption algorithm based on parallel system, which divides the image into blocks and computes in parallel to improve the encryption speed. The above algorithm improves the encryption efficiency to a certain extent. Because it is mainly based on central processing unit (CPU) parallelism [26], the number of parallelism is limited by the number of CPU threads, and parallelism is not considered in iterative chaotic sequences, the efficiency of robot image still cannot meet the requirements.

To further improve the efficiency and security of image encryption systems, some researchers combine chaotic-based schemes with deoxyribonucleic acid (DNA) rules. For example, Zhang *et al.* [28] used DNA coding and low-dimensional chaotic maps to design an image encryption system, but the randomness of the key stream was not high enough. To overcome this shortcoming, Cun *et al.* [6] proposed a color image encryption algorithm based on hyperchaotic systems and cellular automata. The algorithm utilized multiple high-dimensional chaotic systems to generate key streams and ensure the randomness of the key. Due to the limited DNA coding and operation rules, Wang *et al.* [17] proposed a secure Hash algorithm 256 based on chaos theory and Secure Hash algorithm 256, which used random number of DNA complementary rules of "XOR" operation to replace each pixel, it improved the security of the algorithm. However, the algorithm produced a large number of chaotic sequences and sorting operations, which could not meet the demand in efficiency.

In order to further improve the performance of image encryption, this paper extends the in-pixel bit scrambling and bit-plane bit scrambling to image bit global scrambling, and proposes a hyperchaotic image encryption algorithm based on bit-total scrambling. In this new algorithm, eight bit planes decomposed based on bit are spliced into a large bit plane from low to high level, and then the large bit plane is scrambled by hyperchaotic sequence to realize global scrambling of image bits. Finally, the image pixels are diffused forward and backward. Simulation results show that the new method not only has a large key space and is sensitive to the initial value, but also has a strong ability to resist external attacks and a good encryption effect.

The paper is organized as follows. In Section 2, the new four-dimensional hyperchaotic system is described in detail. Then, in Section 3 we present the DNA encoding, and in Section 4, Image encryption and decryption algorithm is shown. Experimental simulations and analysis are validated in Section 5. Finally, there is a conclusion in Section 6.

2 New Four-dimensional Hyperchaotic System

Cicek *et al.* [5] proposed a 3D continuous chaotic system, the equation was shown in Equation (1).

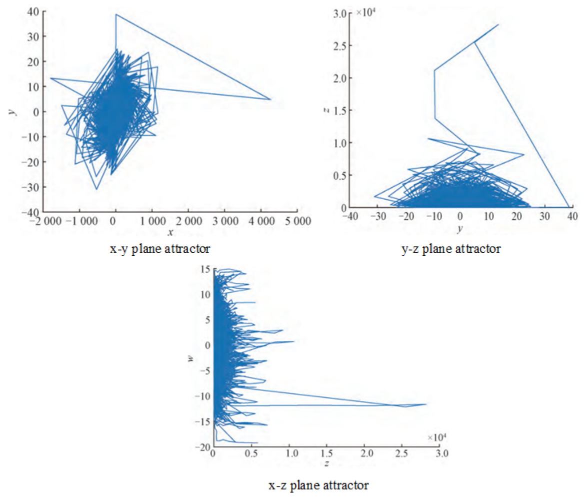
$$\left\{\begin{array}{l}\dot{x}=-ax+yz\\ \dot{y}=-x+by\\ \dot{z}=cy^2-dz\end{array}\right\} \tag{1}$$

When $a = 20$, $b = 10$, $c = 7$ and $d = 5$, the system is a three-dimensional chaotic system. In this paper, based on system (1), variable w is introduced into the second equation of the system. The y^2 in third equation is changed as xy. The variable y is added to the fourth equation to construct a new four-dimensional hyperchaotic system, whose equation is shown in Equation (2).

$$\left\{\begin{array}{l}\dot{x}=-ax+yz\\ \dot{y}=-x+by+w\\ \dot{z}=c|xy|-dz\\ \dot{w}=-ey\end{array}\right\} \tag{2}$$

When $a = 18$, $b = 10$, $c = 9$, $d = 7$, $e = 2$, it uses the Jacobian method to calculate the Lyapunov exponent: $LE1 = 2.3998$, $LE2 = 0.1354$, $LE3 = -0237$, $LE4 = -17.5109$. Two Lyapunov exponents are greater than 0, and one Lyapunov exponent is approximately equal to 0. The sum of all Lyapunov exponents is less than 0, so the system can be judged as a four-dimensional hyperchaotic system.

Therefore, the chaotic attractors of the new four-dimensional hyperchaotic system obtained in x-y plane, y-z plane and z-w plane are shown in Figure 1.



x-y plane attractor

y-z plane attractor

x-z plane attractor

Figure 1: A new four-dimensional hyperchaotic attractor

3 DNA Encoding

3.1 DNA Complementarity Rule and Algebraic Operation

The DNA sequence contains four nucleic acid bases, namely, adenine (A), cymetidine (C), ornipurine (G), and thymetidine (T). Where A and T, C and G are complementary. In a binary system, 0 and 1 are complementary. Similarly, the binary numbers 00 and 11 are complementary, and 01 and 10 are also complementary. Using four bases A, C, G and T to represent the binary numbers 00, 01, 10 and 11, there are 24 DNA coding schemes. However, only eight numbers satisfy the base complementation pairing rule. The encoding rules are shown in Table 1.

Table 1: DNAS encoding rule

Rule	1	2	3	4	5	6	7	8
A	00	00	01	01	10	10	11	11
C	01	10	00	11	00	11	01	10
G	10	01	11	00	11	00	10	01
T	11	11	11	10	10	01	00	00

In this paper, DNA sequences are used to encrypt images. By using A, C, T and G to represent the binary values 00, 01, 10, and 11, each 8-bit pixel value of the image can be encoded as a DNA sequence with length 4. For example, the pixel value 155 can be encoded as ”GCGT” by Rule 1. In order to facilitate the application of DNA computation in cryptography, some biological and algebraic operations are introduced to DNA sequences, such as addition (+), subtraction (-), XOR (exclusive OR) operation, etc.

3.2 DNA-S Box Replacement Procedure

In order to make up for the shortage of operation rules in DNA encryption, S-box in DES is used to encode and replace DNA encryption. In cryptography, S-box is the basic structure of symmetric key encryption algorithm to perform nonlinear replacement computation. The S-box in DES algorithm is a fixed S-box, which can map 6-bit input to 4-bit output. In this paper, DNA coding rule 1 is used to convert S-box into DNA-S box, as shown in Table 2. By analogy, eight fixed S-boxes in DES algorithm can be converted into corresponding DNA-S boxes by using eight DNA coding rules [2]. The resulting DNA-S cassette can replace three DNA codes with two DNA codes. The replacement steps are as follows:

1) Group the encoded DNA sequences, and each group consists of two DNA codes.

2) Two key sequences are calculated, and the first key sequence is used to fill each group of DNA sequences into three; The second key sequence is used to select DNA-S boxes for each group to be replaced.

3) Input each group of DNA sequences into the DNA-S box. The newly filled DNA codes are used to find the rows, and the remaining codes are found in the columns to obtain two new DNA codes and complete the replacement.

Table 2: DNA-S box

Rule	A	C	G	T
AA	TG	AA	CA	TT
AC	CA	TT	AC	TA
AG	TC	CT	TG	GA
AT	AC	CA	GA	AG
CA	AG	TG	TC	CA
CC	TT	AG	CG	GC
CG	GT	TC	AG	AC
CT	GA	AC	GT	CT
GA	AT	GG	TT	CC
GC	GG	CG	TA	GT

4 Image Encryption and Decryption Algorithm

The steps of the image encryption algorithm are as follows.

1) Set the size of the plaintext gray image as $M \times N$.

2) The plaintext image is decomposed into 8 bit planes according to bits, which are $a_1, \cdots, a_8$ respectively from low to high. The 8 bit planes are concatenated

into a large bit plane a in the order as shown in Figure 2. The large bit plane a is with 8M length and N width.

3) Set the equation coefficients of the new four-dimensional hyperchaotic system as a_0, b_0, c_0, d_0 and e_0. The initial values of the state variable is x_0, y_0, z_0, and w_0. The sequences generated by 1000 iterations of the hyperchaotic system are discarded, and then generations $(8M + N)$ are selected to generate 5 hyperchaotic sequences. It selects the hyperchaotic sequence $X = x_i$ whose length is truncated as $8M(i = 1, 2, \cdots, 8M)$ and hyperchaotic sequence with length truncated as $Y = y_i$.

4) The elements in sequence X are sorted in ascending order, and a position index sequence $Q = q_i$ is generated for storing the positions (subscripts) of elements in the ascending sequence in the original sequence X, as shown in figure 3. The complexity of four-dimensional hyperchaotic sequence is high, and the length of the selected sequence is relatively small, which ensures that the same sequence value will not appear in a sequence.

   Similarly, it performs the same operation on the sequence Y to obtain the second position index sequence $P = p_i$.

5) The position index sequence Q and P are jointly used to scramble the large bit plane, that is, the elements with coordinate (i, j) in the large bit plane are placed on the position with coordinate (q_i, p_j). At this point, it completes the scrambling of image bits.

6) The scrambled large bit plane is divided into 8 bit planes according to the order of the bit planes in step (2), and then the 8 bit planes are merged in bit order to transform into the intermediate ciphertext image. Then the intermediate ciphertext image is diffused. The diffusion operation is divided into two steps: forward diffusion and backward diffusion.

7) Forward diffusion of image pixels. The intermediate ciphertext image matrix is concatenated into a sequence $G = g_i, i = 1, 2, \cdots, M \times N$ according to rows. Let the image pixel sequence after forward diffusion be $H = h_i, i = 1, 2, \cdots, M \times N$. The forward diffusion operation is carried out according to Equations (2) and (3). The value of M_0 is an integer between 0 and 255.

$$h_1 = mod(g_1 + M_0 256). \tag{3}$$

$$h_i = mod(g_i + h_{i-1} 256), i \geq 2. \tag{4}$$

8) Backward diffusion of image pixels. Let the pixel sequence after back diffusion be $F = f_i, i = 1, 2, \cdots, M \times N$. According to Equations (4) and (5), back diffusion is performed. The value of M_0 is an integer between 0 and 255.

9) Transform the pixel sequence F after diffusion into a row of length M into a $M \times N$ image matrix, which is the ciphertext image. At this point, the image encryption is complete.

5 Experimental Simulation and Analysis

The simulation experiment software is MATLAB 2017a, the processor is Intel cool capacity i5-8250U, and the memory is 8GB. The selected image is Lena grayscale image with the size of 256 × 256. Setting the encryption algorithm key as $a_0 = 18$, $b_0 = 10$, $c_0 = 9$, $d_0 = 7$ and $e_0 = 2$, $x_0 = y_0 = z_0 = w_0 = 0.1$, $M_0 = M_1 = 35$. Experiment simulation results are shown in figure 4.

As can be seen from Figure 4, the plaintext image is encrypted by the proposed algorithm, and a completely different image is obtained. In order to evaluate the performance of the proposed encryption algorithm, this paper will analyze the encryption effect and security of the algorithm from the aspects of gray histogram, key space, key sensitivity, information entropy, correlation and differential attack.

5.1 Gray Histogram Analysis

Ciphertext images can effectively resist statistical analysis attacks only when they can hide their statistical features. Figure 5 shows the gray histogram of the original image and the encrypted image. The distribution characteristics of the two gray histogram are as follows: the gray histogram of plaintext image is very uneven, while the gray histogram of ciphertext image is relatively uniform. It shows that the algorithm can well conceal the statistical characteristics of plaintext images.

5.2 Key Space Analysis

An algorithm with good encryption performance must have a large enough key space to effectively resist the external exhaustive attack. The key of the algorithm in this paper can be divided into two parts: The first part is five equation coefficients and four initial values of state variables of the four-dimensional hyperchaotic system; In the second part, two parameter values are introduced from the outside when image pixels are diffused in the forward and reverse direction. Therefore, the key has 11 parameter values. Assuming that the computer processes data with a precision of 10^{15}, the key space size is $10^{(11 \times 15)} = 10^{165}$. To resist exhaustive attacks, the key space must be greater than 2^{100}. Since 10^{165} is much larger than 2^{100}, the proposed algorithm is sufficient to resist external exhaustive attacks.

a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8

Figure 2: Bit plane splicing sequence diagram

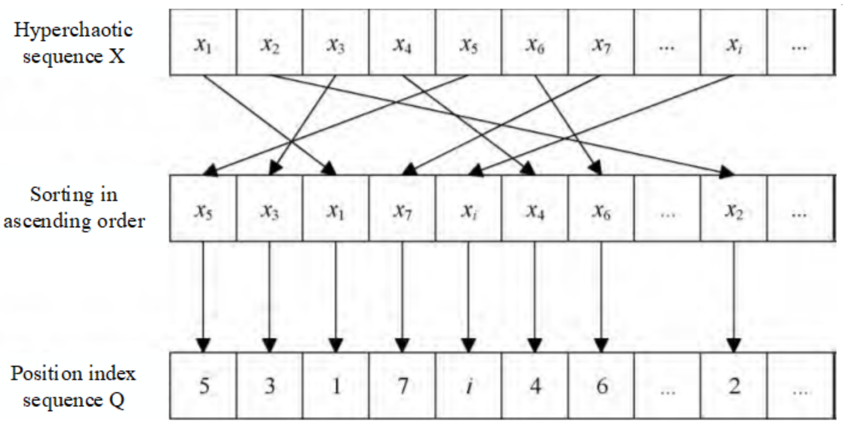
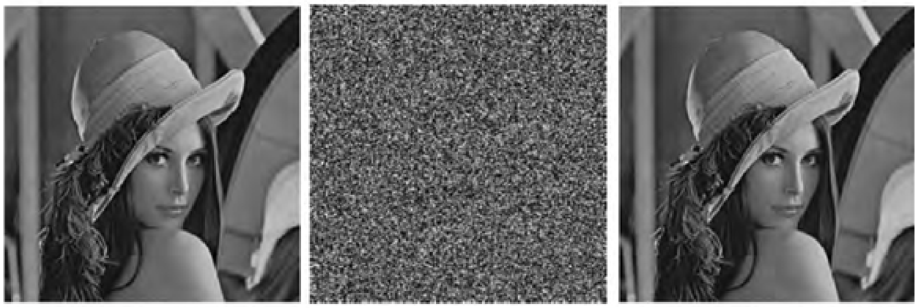
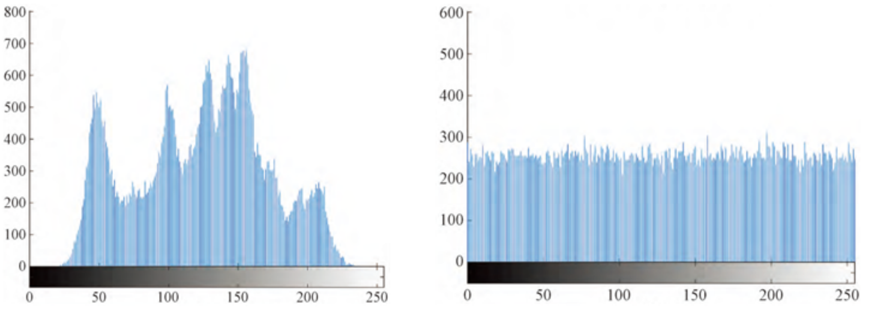


Figure 3: Generation position index sequence



(a) plaintext image (b) ciphertext image (c) Decrypted image

Figure 4: Experimental simulation result



(a) Histogram of plaintext image (b) Histogram of ciphertext image

Figure 5: Histogram analysis

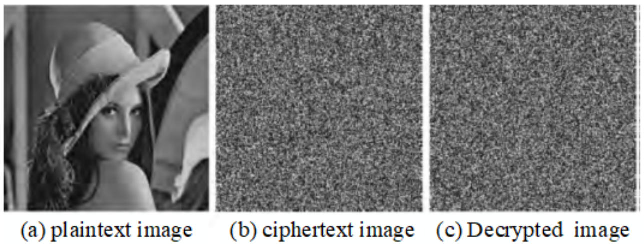


Figure 6: Key sensitivity test

5.3 Key Sensitivity Analysis

A good encryption algorithm must be highly sensitive to small changes in the key. In order to test the sensitivity of the proposed algorithm to the key, in the decryption process, the parameter value x_0 in the key is changed from 0.1 to $0.1 + 10^{-16}$, and other parameter values of the key remain unchanged.

The decrypted image by using the wrong key is shown in Figure 6(c), which is completely different from the original figure in Figure 6(a), indicating that the decryption is failed. It can be seen that the plaintext image cannot be successfully recovered even if the key changes only extremely slightly, which proves that the sensitivity of the algorithm to the key is very high.

5.4 Information Entropy Analysis

Information entropy reflects the chaotic state of pixels in an image and is often used to analyze the effect of image encryption. For an image with a gray level of 256, the image information entropy is equal to 8 in the ideal case. Therefore, the closer the information entropy of ciphertext image is to 8, the better the image encryption effect is. The information entropy of ciphertext image encrypted by the proposed algorithm is 7.9972, which is very close to the ideal value, indicating that the algorithm has a strong ability to resist entropy attack.

5.5 Correlation Analysis

An attacker can attack an image by analyzing the correlation between adjacent pixels. Therefore, a good encryption algorithm should be able to reduce or even eliminate the correlation between adjacent pixels. The correlation between adjacent pixels is defined as follows:

$$r_{xy} = \frac{cov(x,y)}{\sqrt{E(x)}\sqrt{D(x)}}. \quad (5)$$

Where,

$$cov(x,y) = \frac{1}{N}\sum_{i=1}^{N}(x_i - E(x))(y_i - E(y)).$$

$$E(x) = \frac{1}{N}\sum_{i=1}^{N}x_i.$$

$$D(x) = \frac{1}{N}\sum_{i=1}^{N}(x_i - E(x))^2.$$

To evaluate the ability of the proposed algorithm to reduce the correlation between adjacent pixels of an image, correlation coefficients were calculated in the horizontal, vertical and diagonal directions of Lena plaintext and ciphertext images. Meanwhile, the proposed algorithm is compared with the competitive algorithm, and the results are shown in Table 3.

The smaller correlation coefficient denotes the lower correlation. Therefore, this new algorithm can effectively reduce the correlation between adjacent pixels and enhance the ability of ciphertext image to resist statistical attacks.

5.6 Differential Attack Analysis

By making small changes to the plaintext, the attacker can analyze the difference between the ciphertext before and after the plaintext changes, and then obtain useful information. Algorithms with good encryption performance should be highly sensitive to plaintext changes. Even a very small change in the plaintext will make a huge change in the encrypted ciphertext. To evaluate the anti-differential attack capability of the proposed algorithm, NPCR (pixel change rate) and UACI (average change intensity) are used in this paper, which are defined as follows:

$$NPCR = \frac{\sum_{i=1}^{P}\sum_{j=1}^{Q} D(i,j)}{M \times N} \times 100\%.$$

$$UACI = \frac{\sum_{i=1}^{P}\sum_{j=1}^{Q} |S(i,j) - S'(i,j)|/255}{M \times N} \times 100\%.$$

Where $D(i,j) = 0, S(i,j) = S'(i,j)$, $D(i,j) = 1, S(i,j) \neq S'(i,j)$; M and N are the length and width of the image respectively. S and S' are ciphertexts of two images. Where S is the ciphertext of the plaintext image. S' is the ciphertext obtained by encrypting the pixel value of a pixel point in the plaintext image after slightly changing it.

Ideally, the NPCR and UACI values of 8-bit gray image are 99.6094% and 33.4635%, respectively. In the simulation experiment, the Lena image is first encrypted to obtain the first ciphertext image, and then a random pixel point is selected from the original Lena image (ensure that the pixel value of the point is less than 255) to increase the pixel value of the point by 1, and then the Lena image after slight change is encrypted to obtain the second ciphertext image. The corresponding NPCR and UACI are calculated from the two ciphertext images. In order to avoid the contingency of the simulation experiment, five pixels in different positions are selected for the experiment, and the results are shown in Table 4. The two indexes performance of the competitive algorithms is also given, as shown in Table 5. According to Tables 4 and 5,

Table 3: Typical states of SEIR model

Method	Horizontal correlation	Vertical correlation	Diagonal correlation
Original image	0.8569	0.8665	0.8387
Proposed	0.0101	-0.0046	0.0008
Wang *et al.* [19]	-0.0352	0.0202	0.0375
Man *et al.* [12]	-0.0148	0.0037	0.0332

Table 4: NPCR and UACI values by changing the value of a pixel with proposed method

Index	(15,15)	(35,50)	(75,60)	(145,155)	(235,200)
NPCR	99.6077	99.6043	99.5347	99.5805	99.6095
UACI	33.5179	33.4272	33.4569	33.4819	33.4465

NPCR and UACI calculated by the proposed algorithm are very close to the ideal values. Therefore, the proposed algorithm has strong ability to resist differential attack.

Table 5: Other methods for NPCR and UACI

Method	NPCR	UACI
Wang *et al.* [19]	9.1553×10^5	6.2231×10^6
Man *et al.* [12]	0.3493	0.1173

6 Conclusions

In order to improve the secrecy performance of robot image, a hyperchaotic image encryption algorithm based on bit total scrambling and DNA is proposed in this paper. The innovative points of the algorithm are as follows: 1) the new four-dimensional hyperchaotic system adopted by the algorithm has the advantages of simple structure, fast calculation speed, large Lyapunov index and complex chaotic sequence, which is more conducive to image encryption; 2) The algorithm splashes 8 bit-planes of the image into a large bit-plane, and uses the hyperchaotic sequence to scramble the large bit-plane, realizing the global scrambling of image bits and enhancing the effect of bit-scrambling. Simulation results show that the new method not only has large key space, but also can effectively resist statistical analysis and differential attacks, and has better encryption effect and higher security. Considering that nowadays the minimum addressing unit of computer is generally byte, that is, 8bit, the algorithm based on bit scrambling proposed in this paper may produce a large number of repeated computations in the process of implementation, which will affect the computational efficiency. In the future, we will consider to improve the computational efficiency of the algorithm from the hardware aspect.

References

[1] I. Ahmed, M. Kashmoola, "CCF based system framework in federated learning against data poisoning attacks," *Journal of Applied Science and Engineering*, vol. 26, no. 7, pp. 973-981, 2022.

[2] X. Chai, X. Fu, Z. Gan, Y. Yu, Y. Chen, "A color image cryptosystem based on dynamic DNA encryption and chaos," *Signal Processing*, vol. 155, pp. 44-62, 2019.

[3] C. C. Chang, K. F. Hwang, M. S. Hwang, "Robust authentication scheme for protecting copyrights of images and graphics", *IEE Proceedings-Vision, Image and Signal Processing*, vol. 149, no. 1, pp. 43-50, 2002.

[4] G. Cheng, C. Wang, H. Chen, "A novel color image encryption algorithm based on hyperchaotic system and permutation-diffusion architecture," *International Journal of Bifurcation and Chaos*, vol. 29, no. 9, 2019.

[5] S. Cicek, A. Ferikoglu, I. Pehlivan, "A new 3D chaotic system: dynamical analysis, electronic circuit design, active control synchronization and chaotic masking communication application," *Optik*, vol. 127, no. 8, pp. 4024-4030, 2016.

[6] Q. Cun, X. Tong, Z. Wang, M. Zhang, "Selective image encryption method based on dynamic DNA coding and new chaotic map," *Optik*, vol. 243, 2021.

[7] K. Dong, R. Ali, P. Dominic, S. Ali, "The effect of organizational information security climate on information security policy compliance: The mediating effect of social bonding towards healthcare nurses," *Sustainability*, vol. 13, no. 5, pp. 2800, 2021.

[8] K. Fares, A. Khaldi, K. Redouane, E. Salah, "DCT & DWT based watermarking scheme for medical information security," *Biomedical Signal Processing and Control*, vol. 66, 2021.

[9] S. Han, E. Chang, T. Dillon, M. S. Hwang, C. C. Lee, "Identifying attributes and insecurity of a public-channel key exchange protocol using chaos synchro-

nization", *Chaos Solitons & Fractals*, vol. 40, no. 5, pp. 2569-2575, 2009.

[10] M. Kaur, D. Singh, "Multiobjective evolutionary optimization techniques based hyperchaotic map and their applications in image encryption," *Multidimensional Systems and Signal Processing*, vol. 32, no. 1, pp. 281-301, 2021.

[11] Z. Liu, D. Jiang, C. Zhang, H. Zhao, Q. Zhao, B. Zhang, "A novel fireworks algorithm for the protein-ligand docking on the autodock," *Mobile Networks and Applications*, vol. 26, pp. 657-668, 2021.

[12] Z. Man, J. Li, X. Di, Y. Sheng, Z. Liu, "Double image encryption algorithm based on neural network and chaos," *Chaos, Solitons & Fractals*, vol. 152, 2021.

[13] M. Mirtsch, K. Blind, C. Koch, G. Dudek, "Information security management in ICT and non-ICT sector companies: A preventive innovation perspective," *Computers & Security*, vol. 109, 2021.

[14] R. Reshmi, "Information security breaches due to ransomware attacks-a systematic literature review," *International Journal of Information Management Data Insights*, vol. 1, no. 2, pp. 100013, 2021.

[15] Z. Shaikh, A. Khan, L. Teng, A. Wagan, A. Laghari, "BIoMT modular infrastructure: The recent challenges, issues, and limitations in blockchain hyperledger-enabled e-healthcare application," *Wireless Communications and Mobile Computing*, vol. 2022, 2022.

[16] Q. Shi, S. Yin, K. Wang, L. Teng and H. Li, "Multichannel convolutional neural network-based fuzzy active contour model for medical image segmentation," *Evolving Systems*, vol. 13, no. 4, pp. 535-549, 2022.

[17] S. Wang, Q. Peng, B. Du, "Chaotic color image encryption based on 4D chaotic maps and DNA sequence," *Optics & Laser Technology*, vol. 148, 2022.

[18] T. Wang, M. Wang, "Hyperchaotic image encryption algorithm based on bit-level permutation and DNA encoding," *Optics & Laser Technology*, vol. 132, 2020.

[19] X. Wang, Y. Su, "Image encryption based on compressed sensing and DNA encoding," *Signal Processing: Image Communication*, vol. 95, 2021.

[20] X. Wang, S. Yin, M. Shafiq, A. Laghari, S. Karim, O. Cheikhrouhou, W. Alhakami, H. Hamam, "A new V-Net convolutional neural network based on four-dimensional hyperchaotic system for medical image encryption," *Security and Communication Networks*, vol. 2022, 2022.

[21] W. Wei, J. Tang, "Cooperative output regulation by Q-learning for discrete multi-agent systems in finite-time," *Journal of Applied Science and Engineering*, vol. 26, no. 6, pp. 853-864, 2022.

[22] Y. Xian, X. Wang, "Fractal sorting matrix and its application on chaotic image encryption," *Information Sciences*, vol. 547, pp. 1154-1169, 2021.

[23] C. Xu, J. Sun, C. Wang, "An image encryption algorithm based on random walk and hyperchaotic systems," *International Journal of Bifurcation and Chaos*, vol. 30, no. 4, 2020.

[24] Q. Xu, K. Sun, C. Cao, C. Zhu, "A fast image encryption algorithm based on compressive sensing and hyperchaotic map," *Optics and Lasers in Engineering*, vol. 121, pp. 203-214, 2019.

[25] F. Yang, J. Mou, J. Liu, C. Ma, H. Yan, "Characteristic analysis of the fractional-order hyperchaotic complex system and its image encryption application," *Signal processing*, vol. 169, 2020.

[26] S. Yin, H. Li, "GSAPSO-MQC:medical image encryption based on genetic simulated annealing particle swarm optimization and modified quantum chaos system," *Evolutionary Intelligence*, vol. 14, pp. 1817-1829, 2021.

[27] J. Zeng, C. Wang, "A novel hyperchaotic image encryption system based on particle swarm optimization algorithm and cellular automata," *Security and Communication Networks*, vol. 2021, 2021.

[28] J. Zhang, D. Huo, "Image encryption algorithm based on quantum chaotic map and DNA coding," *Multimedia Tools and Applications*, vol. 78, no. 11, pp. 15605-15621, 2019.

[29] Y. Zhang, L. Zhang, Z. Zhong, L. Yu, M. Shan, "Hyperchaotic image encryption using phase-truncated fractional Fourier transform and DNA-level operation," *Optics and Lasers in Engineering*, vol. 143, 2021.

Biography

Jianjun Zhu biography. Jianjun Zhu is with School of Electronics and Electrical Engineering, Zhengzhou University of Science and Technology. His interests are deep learning, AI, image processing.

Jian'E Zhao biography. Jian'E Zhao is with School of Electronics and Electrical Engineering, Zhengzhou University of Science and Technology. His interests are deep learning, AI, image processing.

Attribute Network Representation Learning Based on Generative Adversarial Network and Self-attention Mechanism

Shanshan Li, Meiling Tang, and Yingnan Dong
(Corresponding author: Shanshan Li)

Shenyang Institute of Technology
No. 18 Puchang Road, Shenbei New District, Shenyang, 110136 China
Email: zsshanln@163.com

(Received Sept. 23, 2023; Revised and Accepted Dec. 1, 2023; First Online Dec. 24, 2023)
The Special Issue on Data Fusion, Deep Learning and Optimization Algorithms for Data Privacy Protection
Special Editor: Prof. Shoulin Yin (Harbin Institute of Technology)

Abstract

The attribute network has a complex topology structure, and its nodes contain rich attribute information. The existing learning methods of attribute network representation are not sufficiently enhanced with complementary structural information when learning attribute information, thus affecting the final representation. To solve this problem, we propose a novel attribute network representation learning method based on a generative adversarial network and self-attention mechanism. The traditional convolutional self-attention model makes it easy to waste computing resources because of the information redundancy in the attention graph. In this model, a double normalization method is used to alleviate the problem that the attention model is sensitive to input features, and a new single sample generation adversarial network model is built based on this model. By introducing a joint loss function, structure and attribute information can be represented in the same vector space. Node classification and link prediction experiments on three real attribute network data have obvious advantages over current network representation learning methods.

Keywords: Attribute Network Representation Learning; Double Normalization Method; Generative Adversarial Network; Self-Attention Mechanism

1 Introduction

In the information age of the interconnection of everything, ”network” as an important data form describing the connection between entities in daily life, the study of its representation learning has also received extensive attention. Network representation learning aims to learn low-dimensional, dense, real-valued vector representations for each node in the network, thereby solving the high-dimensional, sparse problems faced by traditional network representations such as adjacency matrix2 [24]. The low-dimensional vector obtained by network representation learning has certain inference ability, and can be directly applied to a variety of network analysis tasks to mine potential information in network data, such as node classification, link prediction, community discovery, etc.

Currently, there have been a lot of researches in the field of network representation learning, and these methods can be generally divided into three categories: factorization based methods, random walk based methods and deep learning based methods.

Factorization based approach. LLE (Locally Linear Embedding) [5] assumes that each node is a linear combination of neighboring nodes in the embedding space. The distance between the weighted sum vector represented by the neighbor node and the vector represented by the central node is used as a loss function, and the embedding vector of the node is obtained by minimizing the loss function. Laplacian Eigenmaps [25] maintains the structure of the network by adding penalty terms to the objective function so that two adjacent nodes are as close as possible in the embedded vector space. GraRep [1] uses the matrix decomposition method as a solution to network embedding, and captures the local and global topology information of the network simultaneously by defining the K-order nearest neighbor loss function.

The factorization method is difficult to be used in large networks because of its large computation and long time. Therefore, the network representation learning method with lower computational complexity and better performance becomes the demand of the current era. Inspired by Word2vec method [4], DeepWalk method [16] obtained the sequence of nodes through random walk, and used

Word2vec model to learn the embedding vector of nodes. The Node2vec method [6] used a preferred random walk strategy to balance the probability of node occurrence in the sequence obtained from the random walk between breadth-first and depth-first search to maintain the high order proximity between nodes, thus generating embedding vectors with higher quality and more information than DeepWalk. In order to get rid of the situation that training is easy to fall into local optimality, HARP [23] used graph coarse-aggregation of nodes in the upper layer of the hierarchy to create a node hierarchy. It was embedded into a rough graph and divides nodes and edges of the original network into a series of network graphs with smaller hierarchical structure through recursive coarse-granulation. Then DeepWalk or Node2vec methods were used for continuous feature extraction to obtain high-performance embedding vectors.

In recent years, network representation learning based on deep learning has become a key technology to solve network analysis tasks. SDNE [18] used deep autoencoders to maintain the similarity of first- and second-order network nodes, and highly nonlinear functions to obtain embedding vectors. SDNE consisted of an autoencoder and a Laplace feature map, where the autoencoder used second-order similarity between nodes to preserve the global structure of the network and the Laplace feature map used first-order similarity to preserve the local structure of the network. DNGR [15] combined random walk and deep autoencoder, used random walk model to generate a probabilistic co-occurrence matrix on the network, then converted the matrix into a PPMI matrix, and then input it into the superimposed denoising autoencoder to obtain the network embedding vector. GCN [7] solved the network embedding problem by defining convolution operators on the network, iterating the neighborhood embeddings of aggregation nodes, and combining the embeddings obtained in the previous iteration to obtain new embeddings. VGAE [8] adopted graph Convolutional network (GCN) encoder and inner product decoder, and used GCN to learn the higher-order dependency between nodes according to the adjacency matrix of the network, and used the inner product decoder to improve the embedding performance.

These graphs show that most of the learning models only encode the network structure information to get the node embedding vector, and do not consider the large number of attribute information of the nodes in the real network. In order to solve the problem of representation learning of attribute networks, scholars have proposed a series of representation learning methods of attribute networks. For example, Cen *et al.* [2] proposed the deep walk of node attribute association, improved the deep walk strategy of attribute information under the matrix decomposition framework, and introduced the attribute features of nodes into network representation learning. Wang *et al.* [20] proposed the maximum margin deep walk method, which uses the attribute information of nodes to learn network representation. Li *et al.* [10] took the attribute information of nodes into account, embed the network structure and the attribute information of nodes as nodes, and proposed a context-aware embedding method.

In this paper, the attribute information of network nodes is considered, the attribute information of nodes is encoded, the problem of overload of node attribute information is solved by introducing global attention mechanism, and the long-distance dependence of node attributes is improved. In this way, the attribute embedding of nodes is learned, and the joint embedding vector representation of nodes is obtained by combining the structural embedding of the network.

2 Related Works

2.1 Network Embedding

At present, a large number of network representation learning models have been proposed for learning effective low-dimensional representations of nodes. For example, DeepWalk combines a random walk and Skip-Gram model to learn network embedding. LINE carries out probabilistic modeling for all pairs of nodes with first-order similarity and second-order similarity in the network, and learns the node representation by minimizing the KL distance between the probability distribution and the empirical distribution. Node2vec comprehensively considers the random walks of depth-first and breadth-first neighborhoods, effectively preserving the network structure. SDNE uses a deep autoencoder to optimize the first and second order similarity simultaneously, and the learned embedding vector can preserve the local and global structure, and is robust to sparse networks. GCN proposed a convolutional neural network for non-European network data. By encoding the local structure of the network and the features of the nodes, the embedded vector of the nodes is obtained.

In real networks, many nodes have rich attributes, for example, Twitter has millions of active and connected users, all of whom can post multiple tweets, and the text, images, or videos of these tweets are attribute information. Node attributes are of great value for network mining tasks, such as solving the bubble problem of large network filters and predicting social behaviors of people [13, 19]. Node attributes contain rich information, which can be combined with network structure to improve the quality of node embedding. However, many characterization methods only encode the network structure information to get the node embedding vector, but do not consider the attribute information of the nodes in the real social network, which will affect the accuracy of the embedding vector and cause many problems. For example, it is difficult for a node to display its different attribute characteristics when interacting with different neighbors, while the embedding vectors of adjacent nodes with relatively different attributes are closer. In this regard, scholars have proposed a series of attribute network representation learning methods. For example, Minardi *et* *al.* [14] introduced packet enhanced network embedding (GEN) to integrate existing packet information in network embedding to strengthen network embedding. Liu *et al.* [11] regarded text content as a special node, learned network embedding by using structure and text information, and then proposed context-enhanced network embedding (CENE). Different from the algorithms proposed before, the algorithm in this paper combines the attention mechanism, network structure information and node attribute information to get the node embedding vector. The following is the relevant research basis.

2.2 Attention Mechanism

The attention mechanism can give a neural network the ability to focus on a subset of its input features, that is, to choose to focus on a particular input. In recent years, it has been widely used in different fields and types of tasks such as image processing, speech recognition, natural language processing and network representation learning.

Attention mechanism is a general approach that does not depend on a specific framework, but most current attention models include the Encoder-Decoder framework. The NMT model applies attention mechanism to Encoder-Decoder in machine translation. The output of the encoder is the weighted sum of the hidden layers in the encoder process, which can alleviate the problem of long distance dependence for very long inputs. With the widespread application of attention mechanisms, many variant models of attention have emerged to deal with more complex tasks. Xu *et al.* [22] designed three functions suitable for different downstream tasks, and proposed the Global Attention and Local Attention mechanisms. Global Attention considers all hidden states of the encoder when calculating hidden layer vectors. Local Attention only pays attention to the hidden state of a part of the encoder when calculating the hidden layer vector. Choudhary *et al.* [3] added a scaling factor to the attention weight function to overcome the minimal gradient problem that softmax function may have in conventional methods and accelerate model training. The Self-Attention method [9] calculates the attention weight by correlating the different positions of a single input in order to calculate the interactive representation of the input. GAT introduces SelfAttention method to calculate the hidden layer state of each node by focusing on the neighbor nodes in the network, and then trains the embedded vector. GeniePath [12] is a scalable graph neural network model capable of learning adaptive sensory paths. Its adaptive path layer includes two complementary functional units, which are used to search breadth and depth respectively. The former introduces attention mechanism to learn the weights of first-order neighborhood nodes, and the latter is used to extract and filter information aggregated in higher-order neighborhoods.

3 Proposed Attribute Network Representation Learning

The proposed model uses a unified framework for representation learning of network structure information and node attribute information. In the early stage, the adjacency of the attribute information of nodes is enhanced to better integrate the two information to learn the optimal network representation. The framework flow of this paper is shown in Figure 1. This is described in detail below.

3.1 Global Structure Information and Attribute Information Learning

The structure information of the network and the attribute information of the nodes belong to heterogeneous information, and the combination of vectors after separate representation learning can also play a role in combination, but the simple concatenation operation is not enough to describe the complex relationship between the structure and attribute of the nodes. Here, inspired by GCN, we aggregate the attributes of the neighbors of nodes through the adjacency matrix reflecting the global structure information and the attribute matrix of the node attribute information, and retain the attributes of the nodes themselves. Specifically, the mathematical representation is shown in Formula (1):

$$M = \tilde{D}^{-0.5}\tilde{A}\tilde{D}^{-0.5}Z. \tag{1}$$

Where, M represents the matrix after the attribute information of itself and its neighbor is aggregated. $\tilde{A}$ is the network adjacency matrix A plus self-connection, which is calculated as shown in Equation (2):

$$\tilde{A} = A + I_n. \tag{2}$$

Where I_n is an $n \times n$ identity matrix. $\tilde{D}$ is the degree matrix of $\tilde{A}$, which is a diagonal matrix whose mathematical representation is shown in Formula (3):

$$\tilde{D}_{ii} = \sum_j \tilde{A}_{ij}. \tag{3}$$

$\tilde{D}^{-0.5}\tilde{A}\tilde{D}^{-0.5}$ is called a normalized adjacency matrix. This is a normalization operation. The original adjacency matrix has not been normalized, and the multiplication of the attribute matrix will change the distribution of the original attribute features. In this way, the aggregation matrix M combining the global structure information of the network and the node attribute information can be obtained by multiplying the normalized adjacency matrix with the attribute matrix, so that the two information can be fused in the early stage of representation learning.

Each row A_i of A represents the first-order adjacency of the current node i. The corresponding position of the node directly connected to node i is 1 in A_i, and the other positions are 0. When the adjacency matrix A is multiplied by the attribute matrix Z, it is the sum of the

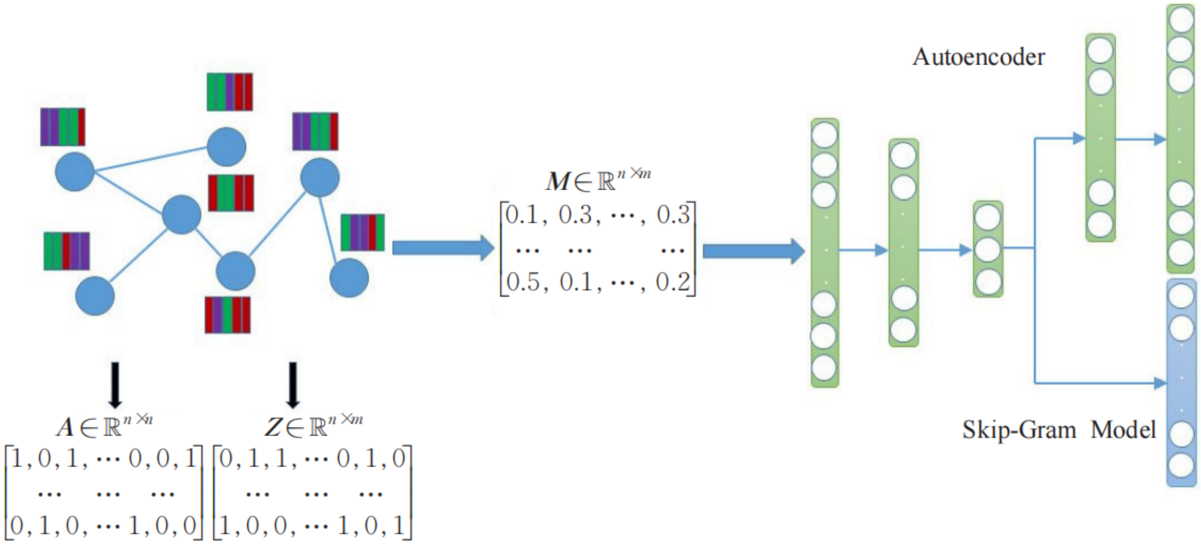


Figure 1: Proposed architecture

attributes of the neighbors of node i in all dimensions. Since the attribute of the node itself is also important, an identity matrix I_n is added to the basis of A to get $\tilde{A}$. In this way, after $\tilde{A}$ and Z are multiplied, the properties of the node itself are also taken into account. The new matrix M obtained after multiplying with its own degree matrix is actually the weighted average of the attributes of the node itself and its neighbors. Compared with the common attribute matrix Z, M contains its own and neighbor attribute information, which is the fusion of network structure information and attribute information.

It is not the ultimate goal to obtain a matrix M with its own attributes and adjacent attributes, but to reduce its dimension and retain the global structure and attribute information in it. In order to give full play to the advantages of automatic feature extraction in deep learning, this paper adopts generative adversarial network and self-attention mechanism for representation learning. Autoencoder is a typical deep learning model for representation learning, and its idea is as follows. The input data is mapped to a feature space through the encoder, and then the feature space compressed by the encoder is mapped back to the input space through the decoder, and the input data is reconstructed. In this way, the middle layer of the neural network saves the features of the input data to achieve dimensionality reduction. Corresponding to the task in this paper, the attribute aggregation matrix M, which enhances the structural information, is used as the input of the auto-encoder. The encoded part of the data is mapped to a low-dimensional vector space, and then the input data is reconstructed in the decoding part. The hidden layer is "forced" to preserve as many features as possible of the input data, the aggregate matrix M, which combines global structure information and attribute information. Thus, the implicit representation of each layer of the auto-encoder is as follows:

$$\left\{ \begin{array}{l} Y_i^1 = \sigma(W^1 \cdot M_i + b^1) \\ Y_i^l = \sigma(W^l \cdot Y_i^{l-1} + b^l) \end{array} \right\} \tag{4}$$

Here, $l = 1, 2, \cdots, L$. L indicates the number of layers of the auto-encoder. $\sigma(\cdot)$ is the activation function of each layer of the network, such as Tanh, ReLU [21], and so on. W^l and b^l are the weight matrix and bias of the $l-th$ layer of the neural network respectively. The auto-encoder does not require additional supervisory information, it is trained by continuously minimizing the reconstruction error between input and output, corresponding to the text task of minimizing the loss of the reconstruction aggregation matrix M, defined as:

$$\zeta_a = \sum_{i=1}^{n} ||\hat{M}_i - M_i||_2^2. \tag{5}$$

Where, $\hat{M}_i$ is the output after M_i is reconstructed by the autoencoder. A low-dimensional representation of the hidden layer is obtained by minimizing ζ_a, that is, the difference between the reconstructed data $\hat{M}_i$ and the original data M_i. According to the training experience of the autoencoder, in order to make the model perform better and prevent over-fitting, the regularization of L2-norm is applied to the parameters W^l and $\hat{W}^l$ of the neural network. In the fitting process, the model usually tends to make the weight, that is, the parameters of the neural network as small as possible, and finally construct a model with all parameters relatively small. Because the model with small parameter values will not have a large impact on the result due to a small amount of data offset, so as to improve the anti-interference ability of the model. The L2 norm as a regular term will make the model parameters as small as possible, but will not be 0, and try to make every eigenvalue contribute to the predicted value,

Table 1: Statistics of the data set

Dataset	node number	edge number	Attribute number	Tag number
Citeseer	3312	4714	3703	6
Cora	2708	5429	1433	7

resulting in dense solutions. So Formula (5) is rewritten as:

$$\zeta_a = \sum_{i=1}^{|V|} ||\hat{M_i} - M_i||_2^2 + \frac{\beta}{2}\sum_{l=1}^{L} ||W^l||_2^2. \tag{6}$$

$|| \cdot ||_F^2$ represents the L2-norm, and β is the regular term coefficient.

3.2 Learning of Local Structural Information

In the previous section, the global structure information and attribute information of the network can be synthesized into the low-dimensional vector space by auto-encoder, but the local structure information needs to be further strengthened. skip-gram model has been widely used in the representation learning of network structures. The basic assumption of network representation learning based on skip-gram model is that if nodes in the network have the same or similar context, their network representations should be similar, so its basic idea is to learn the vector representation of nodes through the co-occurrence relationship between nodes in the network. In this paper, the low-dimensional representation which combines the global structure information and attribute information extracted by the auto-encoder is used to make the representation vectors of the nodes with co-occurrence in the network more similar by skip-gram model. The skip-gram model models the probability of the node pairs in the local window and minimizes the log-likelihood probability represented by Formula (7):

$$\zeta_s = -\sum_{i=1}^{|V|} \sum_{v_j \in C_i} \ln Pr(v_j|M_i). \tag{7}$$

Where $C_i = v_{i-w}, \cdots, v_{i+w}$ refers to the context where w is the window of the central node v_i of the random walk sequence. Conditional probability $Pr(v_j|M_i)$ refers to the possibility that the central node v_i co-appears with the context node after combining the global structure information and attribute information, which is defined as:

$$Pr(v_j|M_i) = \frac{exp(v_j^T \cdot f(M_i))}{\sum_{s=1}^{n} exp(v_s^T \cdot f(M_i))}. \tag{8}$$

$f(M_i)$ is the low-dimensional representation of the global structure information and attribute information of the central node v_i obtained by the auto-encoder, and is the low-dimensional representation of the context of node v_j. Notice in the denominator of Formula (8) that each iteration requires traversing all nodes in the network to complete the calculation, which is quite expensive for larger networks.

4 Experiment and Result Analysis

Through experiments on two real network datasets (Cora and Citeseer), the superiority of the proposed method is verified by comparing with the traditional network representation learning algorithm and the algorithm with node attribute information. Experimental environment: Intel Core i7-7700 CPU 3.60GHz, GeForceGTX 1060Ti; Python 3.7.3, PyTorch 1.3.1.

The statistics for the two datasets covered in this article are shown in Table 1 and are available at https://snap.stanford.edu /data/. Citeseer, and Cora [17] belong to the citation network, where the network nodes represent the papers, the margins represent the citation relationships between papers, the node labels are the research topics of papers, that is, the classification results, and the node attributes represent the attribute features of each paper, such as keywords, publication year, research keywords, etc.

In this paper, the proposed algorithm is compared with five representative network representation learning algorithms, including DeepWalk, node2vec, DANE, AANE, and VGAE. Take 10% of the data set as the test set, 10% as the verification set, and the remaining 80% as the training set. The hyperparameter $\alpha = 0.5$, and each time the threshold is updated, the training sample is adaptively reselected. The parameter Settings for different data sets are shown in Table 2. Where t represents the number of layers of the Laplacian filter and lr represents the learning rate.

Table 2: Parameter Settings for different data sets

Data	t	lr
Citeseer	3	3×10^{-3}
Cora	8	1×10^{-3}

4.1 Node Clustering Task and Result Analysis

In this section, the performance of DENRL algorithm is evaluated by node clustering task, and the experimental results are shown in Table 3. Node clustering is an unsupervised method in which nodes are divided into multiple clusters, Accuracy (ACC) is calculated by comparing the label results with the real labels [17]. Normalized Mutual Information (NMI) is used to measure the similarity of the clustering results [12], the value range is [0,1]. The higher the precision and NMI value, the better the clustering result.

Table 3: Typical states of SEIR model

Data	Cora	Cora	Citeseer	Citeseer
Index	ACC	NMI	ACC	NMI
DeepWalk	0.493	0.339	0.337	0.099
node2vec	0.658	0.367	0.462	0.112
DANE	0.713	0.641	0.490	0.422
AANE	0.456	0.172	0.458	0.154
VGAE	0.565	0.418	0.388	0.292
Proposed	0.786	0.706	0.716	0.469

On the two data sets, the proposed method achieves satisfactory results in both ACC and NMI. Table 4 shows the experimental results of different k values. Compared with $k = 1$ and $k = 0.8$, the best result is obtained at $k = 2/3$, which indicates that the low-pass filter designed in this paper has certain improvement in representation learning compared with the traditional convolution operation.

Table 4: NMI value with different k

Data	$k = 1$	$k = 0.8$	$k = 2/3$
Cora	0.589	0.695	0.567
Citeseer	0.449	0.458	0.448

Table 5 compares the average running time of the proposed algorithm with other algorithms. The average running time of DeepWalk algorithm in one epoch of Cora dataset is 0.4602s, while the running time of node2vec algorithm is 0.8546s, and the running time of the algorithm in this paper is shortened by 46.2%. For the relatively large data set Citeseer, the running time of the algorithm in this paper is 17.4906s, which is less than the other four algorithms. While Cora data set is small and has a lot of attribute information, DeepWalk does not consider node attributes, so the running time is the shortest, which is 1.4997s. The above results show that the proposed algorithm is efficient.

Table 5: Run time comparison/s

Method	Cora	Citeseer
DeepWaik	0.6298	1.2638
node2vec	0.8546	1.6376
DANE	0.5554	0.9978
Proposed	0.4602	0.8547

4.2 Ablation Experiment

In order to verify the effectiveness of the algorithm proposed in this paper, a variant model is set to complete the node clustering task. The parameter Settings are as shown in Table 2. The experimental results on Cora and Citesser datasets are shown in Table 6. Where, attention-only means only the attention mechanism, GAN-only means only the generative adversarial network model, and attention+GAN means the model proposed in this paper.

Table 6: Comparison of ablation results (ACC)

Model	Cora	Citeseer
attention-only	0.667	0.699
GAN-only	0.754	0.6393
attention+GAN	0.776	0.706

For Citeseer data set, the clustering accuracy of the proposed model was 0.706, and the clustering accuracy of the variant model was 0.699 and 0.693, respectively, which increased the clustering accuracy of the proposed model by nearly 1 percentage point. For Cora data sets, the clustering accuracy of attention-only is much worse than that of the model in this paper, but the clustering accuracy of GAN-only is better than that of another variant model, indicating that the network structure and attribute information are of different importance to different data sets. The results on Citeseer data set are not much different, and on Cora data set, the model in this paper shows obvious advantages. The reason is that the number of nodes in Citeseer data set is relatively small, but the number of attributes and edges is large, and there are problems of hidden geometric structure and attribute information confrontation in learning.

5 Conclusions

In this paper, we propose an attentional mechanism and an attribute network representation learning method for generating adversarial networks. The attention mechanism is used to capture the higher-order neighborhood information of nodes, retain the local and global information of the network, and consider the nodes with similar geometric distances in the outer Euclidean space of the embedded space. The low-pass filter is designed by means of generative adversarial network, and the attribute information of neighbors in the neighboring domain is aggregated. Finally, adaptive interactive learning is carried out to obtain the final representation vector of nodes. The experimental results show that the performance of node vectors learned by the proposed method is improved in node clustering and link prediction tasks. In the future, we plan to study the hidden geometry (such as hierarchical structure) and attribute information antagonism in attribute networks.

References

[1] S. Cao, W. Lu, Q. Xu. "Grarep: Learning graph representations with global structural information," *Proceedings of the 24th ACM international on conference on information and knowledge management*, pp. 891-900, 2015.

[2] Y. Cen, X. Zou, J. Zhang, H. Yang, J. Zhou, J. Tang, "Representation learning for attributed multiplex heterogeneous network," *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*, pp. 1358-1368, 2019.

[3] T. Choudhary, V. Mishra, A. Goswami, J. Sarangapani, "A comprehensive survey on model compression and acceleration," *Artificial Intelligence Review*, vol. 53, pp. 5113-5155, 2020.

[4] K. Church, "Word2Vec," *Natural Language Engineering*, vol. 23, no. 1, pp. 155-162, 2017.

[5] O. Eguasa, E. Edionwe, J. Mbegbu, "Local Linear regression and the problem of dimensionality: A remedial strategy via a new locally adaptive bandwidths selector," *Journal of Applied Statistics*, vol. 50, no. 6, pp. 1283-1309, 2023.

[6] A. Grover, J. Leskovec, "node2vec: Scalable feature learning for networks," *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 855-864, 2016.

[7] Y. Jiang, S. Yin, "Heterogenous-view occluded expression data recognition based on cycle-consistent adversarial network and K-SVD dictionary learning under intelligent cooperative robot environment," *Computer Science and Information Systems*, vol. 20, no. 3, 2023.

[8] D. Jin, Y. Gong, Z. Wang, Z. Yu, D. He, Y. Huang, W. Wang, "Graph neural network for higher-order dependency networks," *Proceedings of the ACM Web Conference 2022*, pp. 1622-1630, 2022.

[9] K. Li, Y. Wang, J. Zhang, P. Gao,?G. Song,?Y. Liu,?H. Li,?Y. Qiao, "Uniformer: Unifying convolution and self-attention for visual recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 10, pp. 12581-12600, 2023.

[10] X. Li, X. Fu, G. Xu, Y. Yang, J. Wang, L. Jin, Q. Liu, T. Xiang, "Enhancing BERT representation with context-aware embedding for aspect-based sentiment analysis," *IEEE Access*, vol. 8, pp. 46868-46876, 2020.

[11] W. Liu, M. Gong, Z. Tang, "ETINE: Enhanced textual information network embedding," *Knowledge-Based Systems*, vol. 220, pp. 106917, 2021.

[12] Z. Liu, C. Chen, L. Li, J. Zhou, X. Li, L. Song, Y. Qi, "Geniepath: Graph neural networks with adaptive receptive paths," In *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 1, pp. 4424-4431. 2019.

[13] Z. Liu, D. Jiang, C. Zhang, H. Zhao, Q. Zhao, B. Zhang, "A novel fireworks algorithm for the protein-ligand docking on the autodock," *Mobile Networks and Applications*, vol. 26, pp. 657-668, 2021.

[14] M. Minardi, T. Vu, L. Lei, C. Politis, S. Chatzinotas, "Virtual network embedding for NGSO systems: Algorithmic solution and SDN-testbed validation," *IEEE Transactions on Network and Service Management*, vol. 20, no. 3, pp. 3523-3535, Sept. 2023.

[15] A. Mohan, K. Pramod, "Representation learning for temporal networks using temporal random walk and deep autoencoder," *Discrete Applied Mathematics*, vol. 319, pp. 595-605, 2022.

[16] B. Perozzi, R. Al-Rfou, S. Skiena, "Deepwalk: Online learning of social representations," *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 701-710, 2014.

[17] H. Poon, P. Domingos, "Joint inference in information extraction," *AAAI*, vol. 7, pp. 913-918, 2007.

[18] M. Radmanesh, A. Rezaei, A. Khafaf, M. Jalili, "Topological deep network embedding," *2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC). IEEE*, pp. 476-481, 2020.

[19] L. Teng, Y. Qiao, M. Shafiq, G. Srivastava, A. Javed, T. Gadekallu, S. Yin, "FLPK-BiSeNet: Federated learning based on priori knowledge and bilateral segmentation network for image edge extraction," *IEEE Transactions on Network and Service Management*, vol. 20, no. 2, pp. 1529-1542, 2023.

[20] H. Wang, E. Chen, Q. Liu, T. Xu, D. Du, W. Su, X. Zhang, "A united approach to learning sparse attributed network embedding," In *2018 IEEE International Conference on Data Mining (ICDM). IEEE*, pp. 557-566, 2018.

[21] J. Wang, Y. Fan, H. Li, S. Yin, "WeChat mini program for wheat diseases recognition based on VGG-16 convolutional neural network," *International Journal of Applied Science and Engineering*, vol. 20, no. 3, 2023.

[22] S. Xu, E. Shijia, Y. Xiang, "Enhanced attentive convolutional neural networks for sentence pair modeling," *Expert Systems with Applications*, vol. 151, 2020.

[23] W. Xu, L. Zikatanov, "Adaptive aggregation on graphs," *Journal of Computational and Applied Mathematics*, vol. 340, pp. 718-730, 2018.

[24] D. Zhang, J. Yin, X. Zhu, C. Zhang, "Network representation learning: A survey," *IEEE transactions on Big Data*, vol. 6, no. 1, pp. 3-28, 2018.

[25] H. Zhu, K. Sun, P. Koniusz, "Contrastive laplacian eigenmaps," *Advances in Neural Information Processing Systems*, vol. 34, pp. 5682-5695, 2021.

Biography

Shanshan Li biography. Shanshan Li is with Shenyang Institute of Technology. Interests include data analysis, information processing.

Meiling Tang biography. Meiling Tang is with Shenyang Institute of Technology. Interests include data analysis, information processing.

Yingnan Dong biography. Yingnan Dong is with Shenyang Institute of Technology. Interests include data analysis, information processing.

Research on Security Deduplication of Civil Engineering Data Based on Integration Learning and Sequence Attention Mechanism Fusion

Xiaoli Zhao and Xiaoyan Zheng
(Corresponding author: Xiaoyan Zheng)

College of Civil and Architectural Engineering & Zhengzhou University of Science and Technology
No. 1 Xueyuan Road, Mashai Industrial Park, Erqi District, Zhengzhou, 450064 China
Email: ancrum@qq.com

(Received Oct. 10, 2023; Revised and Accepted Dec. 1, 2023; First Online Dec. 24, 2023)
The Special Issue on Data Fusion, Deep Learning and Optimization Algorithms for Data Privacy Protection
Special Editor: Prof. Shoulin Yin (Harbin Institute of Technology)

Abstract

With the advent of big data, cloud storage platforms face the challenge of storing massive user data. More and more cloud service providers are using cloud data deduplication technology to avoid storing redundant data and save only one copy of user data. However, cross-user deduplication technology saves storage costs but simultaneously makes cloud data privacy face the security risk of being stolen by side-channel attacks. Therefore, we propose a security deduplication model of civil engineering data based on the fusion of integration learning and sequence attention mechanism. A cloud data deduplication model is established based on ensemble learning and fusion neural network algorithms. The sequential attention mechanism is used to obtain more data features. The safety analysis and experimental results show that the proposed method significantly improves safety at the cost of a small increase in overhead compared with the state-of-the-art methods.

Keywords: Cloud Storage; Cross-User De-Duplication Technology; Integration Learning; Sequence Attention Mechanism Fusion

1 Introduction

Cloud storage provides cloud computing services with data storage and management as the core, with the explosion of Internet data, more and more users choose to outsource data to cloud Service providers (CSPS) to reduce local pressure [2, 11, 12]. Because cloud storage has the characteristics of large amount of data and wide range of audiences, cloud service providers are faced with the problem of redundant storage [6–8]. Studies have shown that nearly 75% of data stored in the cloud exists at least one copy [19]. In addition, the repeated upload of large amounts of data also introduces heavy communication overhead for users with limited resources. In view of this, cloud service providers such as Dropbox, Mozy, Mega, Bitcasa111, Baidu Cloud, Alibaba Cloud, etc., have begun to adopt de-duplication technology to solve the problem of data redundancy. According to the different locations, the de-duplication technology can be divided into target de-duplication (also known as cloud de-duplication) and source de-duplication [28] (also known as client deduplication). In target deduplication 3, the client does not need to interact with the cloud server to perform deduplication, which can effectively prevent side channel attacks. However, users must upload complete files before the cloud determines whether to perform deduplication. The resulting increase in communication overhead and storage overhead. Before uploading data, the source uploads its hash fingerprint to the server. If the data fingerprint exists in the server index, the corresponding file will not be uploaded. This technology can detect the same data object in the data stream at the file level, block level and byte level, only transmit and store the unique data object, and use the pointer to the unique data object to replace other duplicate copies, so as to achieve the purpose of reducing massive data quickly [16]. Specifically, before outsourcing the data, the user calculates the hash value of the target file as a query tag and sends it to the cloud service provider, who searches locally to confirm whether the same file is already stored, and if so, returns a response to prevent further uploading by the user. While this approach improves storage efficiency and bandwidth utilization, the deterministic response returned by the cloud provider creates a side channel for the attacker. Once no further upload is required, the existential privacy of the object file is immediately disclosed. Especially for files containing low minimum entropy sensitive information, including emails, enterprise contracts, medical records, electronic payroll, tender slips, etc., attackers can fully guess the contents of the files and launch side-channel attacks such as file confirmation, learning residual information and additional block attacks to steal legitimate users' identity, occupation, sensitive files and other private information, seriously harming user data security [22, 24].

In order to resist side channel attacks, the existing researches are mainly divided into the following categories: First, add trusted gateway. A third-party trusted storage gateway is configured between the client and the cloud server. The client uploads data to the gateway for storage, and then the gateway uploads data to the cloud server. However, the deployment cost of the trusted gateway in the real scenario seriously hinders the practical application. Set the trigger threshold to 10. Only after the number of cloud copies of the requested files exceeds a set threshold will the deduplication mechanism be triggered, so that non-popular files with a high degree of privacy are better protected. However, the cloud needs to store multiple copies of the same file, which inevitably introduces a lot of overhead. And once the number of file copies exceeds the threshold, the deduplication mechanism loses its protective effect on the privacy of file existence. The third is to confuse the response value [14, 25]. That is, the response ambiguity strategy is introduced, so that regardless of whether the target sensitive block in the request file exists in the cloud, the returned deresponse is difficult to distinguish for the attacker. This improves the safety of the deduplication scheme to some extent. However, considering random block generation attack 1, a more complex side channel attack, the privacy leakage probability of the existing scheme will increase dramatically. Specifically, an attacker can mix randomly generated blocks with target blocks containing sensitive information to generate a de-replay request and send it to the cloud service provider. Since the probability of randomly generated blocks existing in the cloud is extremely low, it is regarded as a missed block. Once the response returns the lower boundary value, that is, the number of blocks or linear combinations that the user is required to upload equals the number of randomly generated blocks, the cloud presence of the target sensitive block will be revealed to the attacker. In addition, existing schemes do not pay attention to the intrinsic relationship between the location of the request block and the response. For low-entropy files, attackers can construct specific permutation of de-duplicate requests, and combine learning residual information attacks [17] and random block generation attacks [18] to improve the probability of successful privacy theft.

In this context, in order to achieve security against side channel attacks, we propose a security de-duplication model of civil engineering data based on integration learning and sequence attention mechanism fusion. To be specific, after receiving the deduplication request uploaded by the user, the cloud service provider adds a certain number of blocks with unknown status to the end of the request to blur the existence state of the original request block, and changes the position relationship between the original request blocks through out-of-order processing, and expands the value range of the response value with the help of the newly proposed response table, thus reducing the probability of the return of the lower boundary value.

2 Related Works

Cross-user cloud data deduplication technology is widely used to eliminate redundant data in the cloud and improve storage efficiency. However, attackers can create side channels to steal the existential privacy of cloud data by de-duplicating the results. In order to resist side channel attacks, Waters *et al.* [26] put forward a method based on random Threshold–RTS (Randomized Threshold Solution). First, it sets a storage threshold for each file in the cloud. The value is confidential to cloud users and deduplication is performed only when the number of files stored in the cloud exceeds the threshold. In this way, even if the removal result received by the detector indicates that the cloud requires the upload of the detected file, it does not mean that the file does not exist in the cloud, so as to realize the non-existent privacy protection of cloud data. However, once the cloud service provider blocks the user's upload of the detected file after detecting the label information, it means that the number of the file stored in the cloud reaches the threshold, and the existence privacy of the file will be exposed.

As an improvement, Fan *et al.* [4] put forward a Randomized Redundant Chunk Scheme (RRCS) based on response ambiguity. This method first required the cloud to correctly judge the true existence of the detected file, and then added a random number of hit blocks to the original response for the matched file and the unmatched file respectively to ensure that the number of data blocks contained in the response was within the same range, so that it was difficult for the attacker to judge the true existence of the detected file through the response. To achieve this, the number of additional random blocks is selected in different ranges in both cases. Specifically, the method assumes that all the sensitive information of the file to be detected is contained in one data block, and the remaining blocks are public blocks. For a detection file, the number of missed blocks detected by the cloud can only be 0 or 1, corresponding to the two cases of detected hit and missed. Obviously, for missed files, the cloud response must contain detected missed blocks, then the number of additional hit blocks in the response is randomly selected in $[0, \lambda N]$, λ is a scale factor used to balance security and efficiency, and N is the number of blocks detected. For hit files, the number of additional random blocks in the response is randomly selected in $[1, \lambda N + 1]$. Therefore, in both cases, the number of data blocks contained in the deduplication response is in the range of $[1, \lambda N + 1]$, and the attacker cannot judge the existence of the file by the response. However, once the attacker adds a random number of unhit blocks to the detection file, the number of unhit blocks detected by the cloud will be greater than 1 for both the hit file and the hit file. At this time, the cloud service provider cannot confirm whether the number of additional data blocks in the response is selected in $[0, \lambda N]$ or $[1, \lambda N + 1]$. According to the definition in the scenario of no additional block attack, Tang *et al.* [20] stipulated that the number of additional data blocks was randomly selected in $[1, \lambda N + 1]$ in both cases, so the number of data blocks contained in the response of hit file and non-fatal file would inevitably exist in different interval ranges. This method had the risk of leaking file privacy in the scenario of additional block attack. Later, Karri *et al.* [10] studied a dual data block simultaneous detection method from the perspective of data block detection, and adopted XOR technology to blur the cloud's de-duplication response to realize resistance to side channel attacks. However, their approach still does not achieve the security of file existential privacy in the case of an additional block attack. Kamal *et al.* [9] had improved this approach, but still had this flaw in nature.

For ciphertext cloud data, most of the existing work generates ciphertext based on the Convergent Encryption (CE) encryption technology, which uses the hash value of the plaintext as the key, so that multiple owners of the same file can generate the same ciphertext. The cloud service provider only stores ciphertext, but cannot know the plaintext hash value, so it is difficult to decrypt the plaintext. Based on this technology, Zhu *et al.* [29] used a third-party trusted server to generate random numbers and introduced them into the key generation process based on interactive blind signature technology to resist side channel attacks. However, this technology cannot prevent attackers from obtaining random information by forging identities, that is, attackers can forge normal users to execute protocols, thus generating ciphertext containing random information and then uploading it to remove the duplicate. On the basis of Zhu's work, Lu *et al.* [13] introduced bilinear pair technology to further improve security. Eckert *et al.* [3] proposed a hardware-based random number and key generation method. However, these subsequent works inherit the limitations of Zhu's work. In this category, the side channel attack faced by ciphertext cloud data in the process of deduplication is equivalent to that of plaintext cloud data. If plaintext cloud data resists side channel attack as described above, sufficient security cannot be obtained in the scenario of additional block attack.

3 System Model

The network model in this article consists of two entities, the cloud service provider and the user. A cloud service provider has the powerful computing power to provide cloud storage, cross-user deduplication, and download services for multiple users at the same time. Specifically, after receiving the deduplication request, the cloud service provider stores it locally for comparison, and generates a response according to the deduplication protocol to return to the user, preventing the same file from being uploaded later. Users, on the other hand, reduce the pressure on local storage by outsourcing files. Specifically, the file f that the user is going to upload is divided into blocks $C_i, i \in 1, 2, \cdots, n$ with a fixed length φ. It calculates their Hash values $H_i = Hash(C_i)$ and sends them to the cloud as a query tag, where $Hash()$ is the hash summary function (e.g. SHA-256). It calculates and uploads the corresponding number of linear combinations according to the response value R returned by the cloud, so that the cloud can decode the missed blocks.

Side channel attack is the main threat of cross-user source de-duplication. An attacker disguised as an ordinary user steals the existence state of sensitive block C_i by analyzing the cloud response value of the block, that is, trying to know whether a copy of C_i already exists in the cloud storage. In the traffic obturating policy, the minimum value of the response value changes according to whether the sensitive block exists. Therefore, the attacker can determine the existence status of the sensitive block by observing whether the de-response value is equal to the number of known unhit blocks. If the cloud response value R is equal to the number of known missed blocks, the attacker learns that the sensitive block C_i is already stored in the cloud.

This scheme mainly considers two complex side channel attacks including random chunks generation attack (RCGA) and learn the remaining information attack (LRIA).

RCGA. The attacker constructs a random block whose content is a random bit stream and whose length is a block length φ. Since the probability of a random block being retrieved by the cloud is negligible, it is considered as a missed block. An attacker can construct an upload request that includes a block with an unknown state of private information, t random blocks and s hit blocks, and the existence state of all blocks is known except for sensitive blocks. When the sensitive block exists in the cloud, the minimum response value of the cloud is t. When the sensitive block is not present, the minimum response value is $t + 1$. Therefore, once the response value received by the attacker is t, the attacker can determine that the sensitive block is matched.

LRIA is brute Force dictionary attack. For a low-entropy template file, an attacker can generate and request that all possible versions of the privacy information be uploaded to obtain their respective response values. Combined with random block generation attack, once the generated sensitive block hits and returns the lower boundary value of the response, the content of the privacy information is immediately leaked.

This scheme does not consider "revocation request attack", "Sybil attack "and ProofofOwnership (Pow). Since the dirty block list mechanism of ZEUS is used in this scheme, a blacklist mechanism for file blocks is designed, which can record illegal file blocks as dirty blocks when the user denies and the uploaded content fails to pass the integrity check, and prevent attackers from obtaining more information through "rerequest attack" and "Sybil attack ". In addition, under this model, an attacker can construct a deduplication request, then all blocks of the request are owned by the attacker and will inevitably pass the ownership authentication. Therefore, the ownership verification threats involved in fuzzy de-duplication and ciphertext de-duplication are not considered.

4 The Proposed Data Security Deduplication Method

In this section, we introduce the security de-duplication model of civil engineering data based on integration learning and sequence attention mechanism fusion, named ILSAMF. There are two main reasons why the existing response fuzzy deduplication strategy cannot resist the side channel attack:

1) Defects of multiple sensitive blocks. Existing schemes rely on upload requests for multiple sensitive blocks for security. According to CIDER [23], when multiple sensitive blocks are scattered in l groups, the privacy disclosure probability of each group is 50% if and only if the response values of each group are equal to the number of unmatched blocks in the group. Therefore, the upload request containing l sensitive blocks has a privacy disclosure probability of 2^{-l} for all sensitive blocks. However, an attacker can construct multiple independent upload requests, each containing only a single sensitive block, and each request has a 50% probability of compromising the privacy of a single sensitive block. While the probability of an attacker stealing all the privacy is low, he can steal half of the sensitive blocks with a higher probability.

2) Disclosure of positional relationships. The existing scheme does not fuzzy the relationship between block and block position. In the existing scheme, the sensitive block is fixed at the location of the upload request, so the attacker can construct a single packet containing the sensitive block. The response value of the packet is determined by the number of hit blocks and non-fatal blocks in the group, and the response value interval is small and the fuzziness is poor. When the location relation fuzzy processing is adopted, the sensitive block is distributed in any group with equal probability, the attacker can only steal privacy by checking the response value of the entire upload request, and the response value interval is expanded. Taking probability calculation as an example, if the sequence of privacy disclosure is event A, and the response table privacy disclosure is event B, then the probability of privacy disclosure of the existing scheme is $P(B)$. In this scheme, the sequence of privacy disclosure is event $A \cap B$, and the probability is $P(AB)$, which is less than or equal to $P(B)$. Therefore, the absence of location ambiguity is one of the reasons for the high probability of privacy disclosure in the existing scheme.

ILSAMF's process usually begins with the user sending a deduplication request. The user blocks the file and sends the hash value of each block to the cloud as a query tag. Next, the response value is obtained through the six submodules of the response value processing module.

- The cloud queries the fingerprint uploaded by the user from the database, checks whether the hash value of each requested weight removal block exists, and obtains the existence status array of the block to be uploaded. Query the list of dirty blocks to check whether each block is marked as a dirty block. If so, modify the dirty block status identifier of the corresponding data block in the presence status array.

- The add-on block generator appraises k state random labels to the presence state array.

- Perform out-of-order processing.

- Generate a response value for the upload request according to the response table.

- Modify the response value according to the lower boundary value of the response value, so that its minimum is 1.

- Return the response value.

Users upload a specified number of linear combinations based on the response value returned by the cloud service provider. The improperly uploaded blocks are added to the dirty block list. The cloud decodes the linear combination uploaded by the user to restore the corresponding missed blocks and saves them in the local database. In addition, the cloud service provider maintains a block list of stored files, and the block list stores the physical address of the blocks contained in the file in the database.

Uploading blocks out of order. Out-of-order is the possible dispersion of sensitive blocks into any group, so as to expand the value range of the response value, so as to achieve confusion to the attacker. In order to achieve out-of-order, we usually use a library function to implement scrambling, input ordered sequence, get equal length out-of-order sequence. From the implementation point of view, the out-of-order can be expressed in the form of a meta-algorithm as follows:

$$D_{shuffle} \leftarrow Shuffle(D', w). \quad (1)$$

Where w is the scrambled key, which is randomly generated by $ILSAMF$. The algorithm takes the existing status array D' of the file block that has been blurred by the cloud as the input, and uses the array $D_{shuffle}$ which is out of order as the output.

Integration Learning.

K-nearest neighbor (KNN) classifier [1] represents a simple and general classification method whose relative performance is closely related to the choice of distance metric or similarity metric. Its core working principle is to calculate the distance between the sample to be classified and K neighbor samples, find the K samples that are closest to the sample to be classified, and determine the category of the sample to be classified by the category with the largest number of samples, using the principle of minority obedience to majority. KNN classifiers are one of the laziest learning methods because they have low data distribution requirements and a certain tolerance for noise.

RF is an integrated algorithm based on decision tree [21], and randomness is an important attribute that is different from decision tree. This randomness is reflected in Bagging technology, which can realize the re-sampling of samples, which is conducive to reducing the correlation between decision tree models and increasing the accuracy of RF. RF has many advantages, such as simplicity, small computation, and reducing the over-fitting of decision trees.

SGD is an optimization form based on the total gradient descent method (GD). GD needs to use all the training set samples in each iteration to update the model parameters, while SGD only needs to randomly extract a sample in the training set to calculate the gradient for each update training. In comparison, GD has higher accuracy, but it also has the problems of excessive calculation, obvious time consuming, and easy to fall into local extreme points. Although the calculation amount of SGD is small and the speed is fast, the accuracy is slightly lower than GDP.

XGBoost also uses DT as a base learner to reduce the deviation and achieve the optimal classification effect through multi-round iterative calculation of the weak classifier. Generally, the classification regression tree is used as the weak classifier, and the weak classifier obtained by each training needs to be weighted and summed to generate the total classifier. XGBoost mainly uses second-order Taylor loss function to calculate and uses regular term to get the optimal solution, which can make full use of the advantages of multi-core CPU parallel computing and realize faster model exploration.

Fusion is the process of organic integration of different model prediction results of the same data to characterize the results. In general, fusion can be divided into pre-fusion and post-fusion, and post-fusion is the fusion at the result level. The fusion after stacking probabilities is a classic collection algorithm. In this study, the probabilistic fusion mode of the back end of the stacking is used to predict the probabilistic output of the trained machine learning classifier, and then input the original image features and deep learning features into the BP back-propagation neural network for fusion. Among them, four traditional machine learning models are used as primary learners, while BP neural networks are meta-learners. The deep learning network of the meta-classifier is realized by BP artificial neural network algorithm, which consists of three parts: input layer, hidden layer and output layer. The working principle is as follows: First, the input signal to the input layer neurons, and then the signal layer by layer forward until the output layer results are generated; Then the error of the output layer is calculated, and the error is diffused to the hidden layer neurons. Finally, the connection weight and threshold are adjusted based on the error of hidden layer neurons until the stop condition is satisfied, and finally a model with good generalization ability is obtained. The whole study is completed by torch module, which uses cross entropy loss function to classify the features. Optimizer selects the Adam optimizer. The classification result is set to 5 and the hidden layer is set to 100.

Sequence Attention Mechanism.

The multi-head self-attention mechanism gives the model the ability to compute attention in multiple semantic Spaces, which enables the model to more accurately capture the diverse semantic information of the input sequence [5]. In this paper, in the process of updating sentence representation by using sequential self-attention mechanism: First, the Query matrix $Query$, Key matrix Key, and Value matrix $Value$ are computed through the full connection layer. Then, the Query matrix $Query$ is multiplied with Key matrix Key and normalized using Softmax activation function. Finally, the normalized result is multiplied with Value matrix $Value$ to get the final output.

$$h_i = [h'_1; h'_2; \cdots, h'_{num}]. \quad (2)$$

$$head_j = softmax(\frac{[W_Q, h'_j][W_K, h'_j]^T}{\sqrt{num}})[W_V, h'_j]. \quad (3)$$

$$t_i = W_O Concat(head_1, head_2, \cdots, head_{num}). \quad (4)$$

Where $h_i \in R^d$ is the $i-th$ initial character feature. $h'_j \in R^{length}$ represents the feature under the $j-th$ head after the initial character feature is equally divided according to the number num of attention heads. $W_V \in R^{length\times length}$, $W_Q \in R^{length\times length}$, $W_K \in R^{length\times length}$, $W_O \in R^{length\times length}$ represent trainable parameter matrices respectively. $length$ indicates d/num. d represents the dimension of the initial character node feature. $Concat$ means to stitch together features under different heads of attention.

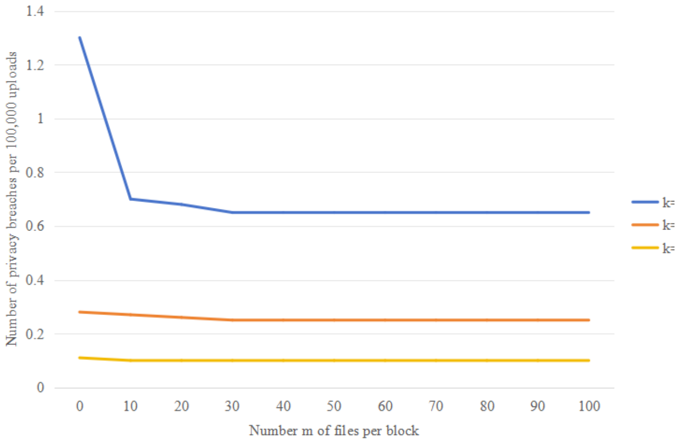


Figure 1: ILSAMF privacy disclosure times graph (where the number of missed blocks $t = 1$, the lower limit of response value has been set)

5 Experiment and Result Analysis

In this section, experiments are conducted to test the security of this scheme against side channel attacks and the performance of deduplication. The comparison objects are SAED and BEDC. SAED indicates cross-user source deduplication without any fuzzy processing. BEDC is the CIDER deweighting model in which four blocks are uploaded each time into a group. The experiment adopts Python to implement the cloud data security deduplication method using integration learning and sequence attention mechanism fusion. In this system, the hash summary function adopts SHA-256 algorithm. In this method, Amazon EC2 is selected to deploy the cloud program, and a group of servers configured with Intel(R)Core(TM) i7-8565U CPU 1.80 GHz. 8GB RAM memory and 512GB capacity 7200 RPM hard disk are selected to deploy the client program. Civil engineering data are used in the safety verification experiment. Part of the civil engineering data, the size of which is 2.0GB, is used in the weight removal performance experiment.

5.1 Model Parameter Setting

In this model, the block size φ (B) is the experimental variable, $\varphi \in 64, 128, 256, 512$. The probability α that the attached block is a public block is set to 0.5. The number of additional blocks k is the experimental variable, $k \in 2, 4, 6$. The default threshold T for the number of blocks contained in a file is 30, because as shown in Figure 1 to Figure 3, the security is gradually enhanced with the increase of the number of blocks. When the number of blocks is greater than or equal to 30, the security tends to a constant value, so $T = 30$ is selected.

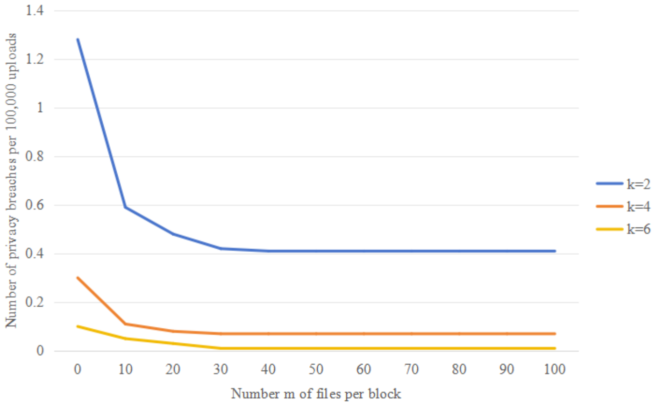


Figure 2: ILSAMF privacy disclosure times graph (where the number of missed blocks $t = 2$, the lower limit of response value has been set)

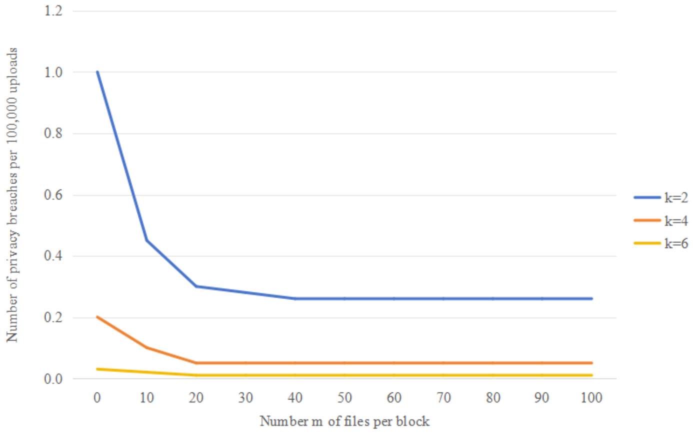


Figure 3: ILSAMF privacy disclosure times graph (where the number of missed blocks $t = 4$, the lower limit of response value has been set)

5.2 Comparison with Other Methods

The number of privacy leaks of ILSAMF and the comparison object under the same input parameters is shown in Table 1. In the 100000 attacks, the number of privacy breaches in the control scheme SAED is 100000 times, and the number of privacy breaches of scheme BEDC is 50012 times. Under different parameter settings of ILSAMF, the number of privacy leaks is 25007, 6265 and 1398 times respectively, which is significantly lower than that of the comparison scheme, and users can set different additional block number k according to their own needs to achieve dynamic adjustment of security. Therefore, ILSAMF has higher security in the face of the hybrid attack of "random block generation attack" and "learning residual information attack". The reason for this phenomenon is that in the control group A, the cloud service provider does not make any ambiguity on the response value, and the response value is equal to the number of missed blocks, so the probability of exposing privacy is 100%. In scheme B, the blocks requesting deweighting are checked in pairs, and the attacker constructs a match between the target sensitive block and the unmatched block. When the target block hits, the response value is fixed as 1, and the user is required to upload the XOR value of the two blocks; otherwise, the response value is fixed as 2, and the user is required to upload the two blocks themselves. Thus, an attacker can determine the existence of the target block based on a response value of 1.

Table 1: Safety performance comparison

Deduplication method	Number of privacy breaches
SAED	100000
BEDC	50012
ILSAMF (k=2)	25007
ILSAMF (k=4)	6265
ILSAMF (k=6)	1398

5.3 Deduplication Efficiency

In this experiment, the optimal block length is explored by comparing the transmission cost of each model under different block lengths. The smaller the block length, the higher the intensity of weight removal. However, the smaller the block length, the amount of metadata will increase correspondingly, the number of client fingerprints will increase, and the fingerprint transmission cost will be higher. To this end, this section designs experiments to investigate the effect of block length on transmission overhead.

The experimental Settings are as follows: The client data test set to be de-duplicated is the VE1 folder in the civil engineering data set, which contains 1600 text data files with a total size of 5.84MB. The cloud data is all the folder data in this dataset, excluding the test set, and contains 510000 files totaling about 1.5GB of text data. Scenario: Most employees of a construction company already have their email data stored in the cloud, and the employee may be the last to upload her email. Due to the very similar form between different emails, it is possible for May to save more transmission overhead by using source-side replays than by using target-side replays.

The relationship between block length and total transmission overhead is shown in table 2. When the block length is 128B, all schemes have the best deduplication performance and the total transmission overhead is maintained at the lowest level. When the block length is 64B or 256B, the deduplication transmission cost of each scheme increases, and when the block length reaches 512B, a large amount of additional data needs to be transmitted. Therefore, 128B is the best block length for each scheme.

Table 2: The relationship between the transmission cost of each scheme and the block length

block size	SAED	BEDC	ILSAMF
64	2.3	2.42	2.52
128	2.21	2.31	2.39
256	2.41	2.46	2.53
512	2.69	2.78	2.75

6 Conclusions

In this paper, a security de-duplication model of civil engineering data based on integration learning and sequence attention mechanism fusion is proposed to realize the security against random block generation attacks and learning residual information attacks. Specifically, the strategy uses additional random blocks, out-of-order file blocks, response table construction and other technologies to hide the storage status and location information of file blocks, expand the value range of response values, and greatly reduce the probability of lower boundary value of response values with a small amount of overhead, thus reducing the security risk to the acceptable probability range. The security analysis and experimental results show that compared with the current popular techniques, the proposed method can improve the security against side-channel attacks at the cost of increasing a small amount of overhead, and has a wide application prospect in the field of low-entropy file resetting, including E-mail, electronic payroll, enterprise contract, medical record, tender sheet, etc.

Acknowledgments

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

References

[1] H. Alfeilat, A. Hassanat, O. Lasassmeh, A. S. Tarawneh, M. Alhasanat, "Effects of distance measure choice on k-nearest neighbor classifier performance: a review," *Big data*, vol. 7, no. 4, pp. 221-248, 2019.

[2] P. S. Chung, C. W. Liu, M. S. Hwang, "A study of attribute-based proxy re-encryption scheme in cloud environments", *International Journal of Network Security*, vol. 16, no. 1, pp. 1-13, 2014.

[3] C. Eckert, F. Tehranipoor and J. A. Chandy, "DRNG: DRAM-based random number generation using its startup value behavior," in *2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS), Boston, MA, USA*, pp. 1260-1263, 2020, doi: 10.1109/MWSCAS.2017.8053159.

[4] J. Fan, C. Guan, K. Ren and C. Qiao, "Middlebox-based packet-level redundancy elimination over encrypted network traffic," *IEEE/ACM Transactions on Networking*, vol. 26, no. 4, pp. 1742-1753, 2018.

[5] S. Hao, D. Lee, D. Zhao, "Sequence to sequence learning with attention mechanism for short-term passenger flow prediction in large-scale metro system," *Transportation Research Part C: Emerging Technologies*, vol. 107, pp. 287-300, 2019.

[6] M. S. Hwang, C. C. Lee, T. H. Sun, "Data error locations reported by public auditing in cloud storage service," *Automated Software Engineering*, vol. 21, no. 3, pp. 373–390, Sep. 2014.

[7] M. S. Hwang, T. H. Sun, "Using smart card to achieve a single sign-on for multiple cloud services," *IETE Technical Review*, vol. 30, no. 5, pp. 410–416, 2013.

[8] M. S. Hwang, T. H. Sun, C. C. Lee, "Achieving dynamic data guarantee and data confidentiality of public auditing in cloud storage service," *Journal of Circuits, Systems, and Computers*, vol. 26, no. 5, 2017.

[9] M. Kamal, S. Amin, F. Ferooz, M. Awan, "Privacy-aware genetic algorithm based data security framework for distributed cloud storage," *Microprocessors and Microsystems*, vol. 94, 2022.

[10] R. Karri, K. Wu, P. Mishra and Yongkook Kim, "Concurrent error detection schemes for fault-based side-channel cryptanalysis of symmetric block ciphers," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 21, no. 12, pp. 1509-1517, 2002.

[11] C. W. Liu, W. F. Hsien, C. C. Yang, M. S. Hwang, "A survey of attribute-based access control with user revocation in cloud data storage", *International Journal of Network Security*, vol. 18, no. 5, pp. 900-916, 2016.

[12] C. W. Liu, W. F. Hsien, C. C. Yang, and M. S. Hwang, "A survey of public auditing for shared data storage with user revocation in cloud computing", *International Journal of Network Security*, vol. 18, no. 4, pp. 650-666, 2016.

[13] N. Lu, Y. Zhang, W. Shi, S. Kumari, K. Choo, "A secure and scalable data integrity auditing scheme based on hyperledger fabric," *Computers & Security*, vol. 92, 2020.

[14] B. Mahesh, K. Pavan Kumar, S. Ramasubbareddy & E. Swetha, "A review on data deduplication techniques in cloud," *Embedded Systems and Artificial Intelligence: Proceedings of ESAI 2019, Fez, Morocco*, pp. 825-833, 2020.

[15] Y. Ming, C. Wang, H. Liu, Y. Zhao, J. Feng, N. Zhang, W. Shi, "Blockchain-enabled efficient dynamic cross-domain deduplication in edge computing," *IEEE Internet of Things Journal*, vol. 9, no. 17, pp. 15639-15656, 2022.

[16] J. Periasamy, B. Latha, "An enhanced secure content de-duplication identification and prevention (ESCDIP) algorithm in cloud environment," *Neural Computing and Applications*, vol. 32, no. 2, pp. 485-494, 2020.

[17] K. Raja, G. Gupta, S. Venkatesh, R. Ramachandra, C. Busch, "Towards generalized morphing attack detection by learning residuals," *Image and Vision Computing*, vol. 126, 2022.

[18] O. Sanda, M. Pavlidis, S. Seraj, N. Polatidis, "Long-range attack detection on permissionless blockchains using deep learning," *Expert Systems with Applications*, vol. 218, pp. 119606, 2023.

[19] W. Shi, Q. Tang, "Cost-optimized data placement strategy for social network with security awareness in edge-cloud computing environment," *Journal of Combinatorial Optimization*, vol. 45, 2023. DOI:10.1007/s10878-022-00934-2.

[20] X. Tang, Y. Zhang, L. Zhou, D. Liu, B. Hu, "Request merging based cross-user deduplication for cloud storage with resistance against appending chunks attack," *Chinese Journal of Electronics*, vol. 30, no. 2, pp. 199-209, 2021.

[21] L. Teng, H. Li, S. Yin, S. Karim, Y. Sun, "An active contour model based on hybrid energy and fisher criterion for image segmentation," *International Journal of Image and Data Fusion*, vol. 11, no. 1, pp. 97-112. 2020.

[22] L. Teng, H. Li, S. Yin and Y. Sun, "A modified advanced encryption standard for data security," *International Journal of Network Security*, vol. 22, no. 1, pp. 112-117, 2020.

[23] R. Vestergaard, Q. Zhang and D. E. Lucani, "CIDER: A low overhead approach to privacy aware client-side deduplication," In *GLOBECOM 2020 - 2020 IEEE Global Communications Conference, Taipei, Taiwan*, pp. 1-6, 2020, doi: 10.1109/GLOBECOM42002.2020.9348272.

[24] X. Wang, S. Yin, H. Li, "A network intrusion detection method based on deep multi-scale convolutional neural network," *International Journal of Wireless Information Networks*, vol. 27, no. 4, pp. 503-517, 2020.

[25] Z. Wang, Y. Lu, G. Sun, ”A policy-based deduplication mechanism for securing cloud storage,” *International Journal of Electronics and Information Engineering*, vol. 2, no. 2, pp. 70-79, 2015.

[26] P. Waters, B. Rucker, M. Love, M. Vassar, ”Lowering the statistical significance threshold of randomized controlled trials in three major general anesthesiology journals,” *Canadian Journal of Anesthesia/Journal canadien d’anesthesie*, vol. 70, pp. 1441-1448, 2023.

[27] Q. Xie, C. Zhang and X. Jia, ”Security-aware and efficient data deduplication for edge-assisted cloud storage systems,” *IEEE Transactions on Services Computing*, vol. 16, no. 3, pp. 2191-2202, 2023.

[28] G. Zhu, X. Zhang, L. Wang, Y. Zhu and X. Dong, ”An intelligent data de-duplication based backup system,” in *2012 15th International Conference on Network-Based Information Systems, Melbourne, VIC, Australia*, pp. 771-776, 2012. doi: 10.1109/NBiS.2012.150.

[29] X. Zhu, Y. Su, M. Gao and Y. Huang, ”Privacy-preserving friendship establishment based on blind signature and bloom filter in mobile social networks,” in *2015 IEEE/CIC International Conference on Communications in China (ICCC), Shenzhen, China*, pp. 1-6, 2015, doi: 10.1109/ICCChina.2015.7448735.

Biography

Xiaoli Zhao biography. Mengya Wei is with the College of Civil and Architectural Engineering & Zhengzhou University of Science and Technology. Her research interests include Data security analysis and Civil engineering data analysis.

Xiaoyan Zheng biography. Xiaoyan Zheng is with the College of Civil and Architectural Engineering & Zhengzhou University of Science and Technology. Her research interests include Data security analysis and Civil engineering data analysis.

Research on Network Security Intrusion Detection Method Based on Optimization Algorithm and Neural Network

Jie Li and Jing Li
(Corresponding author: Jing Li)

Yinchuan University of Energy
Wangtaibao, Yongning County, Yinchuan, Ningxia 750105, China
Email: ljjing81@outlook.com

(Received Oct. 17, 2022; Revised and Accepted Dec. 1, 2023; First Online Oct. 24, 2023)

Abstract

While the Internet brings convenience, it also brings the challenge of information security. In this paper, a convolutional neural network (CNN) is briefly introduced, and it was used in network intrusion detection. At the same time, the genetic algorithm (GA) was introduced in the training process to adjust the parameters of CNN. After that, the performance of the GA-CNN algorithm was tested in the simulation experiment and compared with the support vector machine (SVM) and traditional back-propagation neural network (BPNN) algorithms. The results showed that the GA-CNN algorithm converged to stability faster, and the error was smaller when it was stable. The GA-CNN algorithm exhibited the best performance in identifying abnormal traffic and consumed the least average time in identification.

Keywords: Convolutional Neural Network; Genetic Algorithm; Intrusion Detection; Network Security

1 Introduction

Network intrusion detection is one of the important research directions in the field of network security [1]. It mainly finds behaviors that violate security policies or potential attacks by collecting and analyzing data such as network traffic, system logs, and application logs, and takes corresponding measures to prevent and respond. The ultimate goal is to protect the security of the network and system [14] and prevent or mitigate the loss caused by potential attacks. The methods of network intrusion detection are mainly based on characteristics, statistics, time series, etc. When the network intrusion detection system detects abnormal data or attack behavior [3], it will send an alarm and record the relevant intrusion information when the connection is blocked.

The big data of the Internet is not only large in quantity, but also complex in law. It is difficult for traditional detection methods to mine hidden rules in the big data. Neural networks, with their strong characteristics of self-learning, self-organization, and adaptability, can effectively handle hidden patterns in big data to accurately identify intrusion attack behaviors [4, 9, 17]. Li *et al.* [8] developed a cooperative intrusion detection network in order to improve the detection accuracy of independent intrusion detectors, and verified the robustness of this method through experiments. Wu *et al.* [18] proposed a distributed intrusion detection model for big data based on lossless segmentation and balanced allocation and verified its effectiveness. Li *et al.* [7] proposed a hybrid learning model based on k-nearest neighbor (kNN), and introduced density peaks into kNN. The experimental results showed that this method effectively detected intrusion attacks.

In this paper, the convolutional neural network (CNN) is briefly introduced, and it was used for network intrusion detection. At the same time, the genetic algorithm (GA) was introduced to adjust the parameters of CNN in the training process. In the subsequent simulation experiments, the performance of the GA-CNN algorithm was tested and compared with the support vector machine (SVM) and traditional back-propagation neural network (BPNN) algorithms.

2 Network Intrusion Detection Algorithm Based on Neural Network

2.1 Convolutional Neural Network

With the development of computer technology, neural network technology has been applied to more and more fields. In the face of big data on the Internet, neural networks can use the characteristics of self-learning, self-organization, and self-adaptation to mine hidden rules [10]. When neural network is applied to network intrusion detection, it will determine whether the data is abnormal according to the relevant characteristics of the data to be detected. In this paper, CNN is used to detect network intrusion data [13]. Figure 1 shows the basic structure of a CNN, where the input layer and output layer are used for the input of data and the output of judgment results, respectively. In this paper, when detecting the network data, the features of the data will be extracted initially, and a two-dimensional matrix of "data number × feature dimension" can be obtained after arranging the data [5].

The two-dimensional matrix can be regarded as an image, and the convolution pooling operation is performed on it after input to CNN. The convolution kernel is used in the convolution layer to perform convolution calculations on the two-dimensional matrix, and the corresponding formula is:

$$y_j^l = f(\sum_{i=1}^{N_j^{l-1}} w_{i,j} \otimes x_i^{l-1} + b_j^l), j = 1, 2, \cdots, m, \quad (1)$$

where l represents the current layer number, w represents the convolution kernel weight matrix, x_i^{l-1} represents the input data or the output feature map matrix of the previous layer, f represents the activation function, $\otimes$ represents the convolution operation [2], b_j^l represents the bias of the j^{th} feature map in the l^{th} layer. The pooling layer compresses the extracted convolutional features. During the compression process, the pooling box slides on the feature map according to a certain step size, and each slide will compress the data in the box, which can be taken as the average or the maximum. The number of convolution, pooling and other operations on data will be determined according to the actual needs. Finally, the processed convolution features are calculated in the fully connected layer to obtain the classification results [11].

2.2 CNN-based Network Intrusion Detection Algorithm Optimized by a Genetic Algorithm

The traditional adjustment method is to adjust the parameters by back propagation according to the error between the results obtained by forward calculation and the actual results, but this method is easy to fall into local optimal solutions and overfitting under a fixed learning rate [15]. In order to solve this defect, this paper introduces a GA to adjust the parameters in the process of training CNN, and the improved CNN training process is illustrated in Figure 2.

1) The traffic packet data to be detected is read. PCAP is a common traffic data format in the transmission traffic of the Internet. This paper takes this type of traffic data as the object.

2) Features are extracted from the traffic packet data. On the one hand, the amount of traffic data transmitted in the Internet is large, and it will consume a lot of time to directly detect all the data. On the other hand, in order to ensure security, traffic data is usually encrypted, so it is necessary to extract features from it to facilitate detection [12]. When extracting traffic features, a bidirectional flow is constructed first, and the packets in the PCAP file are put into the flow one by one until the packet interval expires, the packet indicates an end, or the PCAP file is read. Then, the traffic features of the bidirectional flow are counted to generate a CSV format file, and a new bidirectional flow is created. There are a total of 19 traffic features in the bidirectional flow, as shown in Table 1. These features can be directly observed from the packets of bidirectional flow. They do not refer to the specific content of the packets, nor are they the encrypted objects of the packets. Therefore, the above features can be counted normally even for encrypted packets, which is conducive to the identification of disguised malicious traffic.

3) A chromosome population with a certain size is generated randomly. Each chromosome represents a parameter scheme of CNN, and each gene position represents a parameter to be optimized. The gene length of the chromosome depends on the number of parameters to be optimized.

4) The chromosome is substituted into the CNN algorithm after decoding.

5) The traffic features in the CSV format file are input into the CNN algorithm for forward calculation. The convolution formula used in the convolution layer and the convolution feature compression method in the pooling layer are as described above.

6) Whether the CNN training is terminated is judged. If terminated, the training is over, and the result is output; if it does not terminate, genetic operation is carried out on the chromosome population. "Crossover" refers to randomly selecting two chromosomes according to the crossover probability to exchange the values on the same gene loci. "Mutation" refers to selecting a gene locus in a single chromosome according to the mutation probability and randomly changing the value on the gene locus [16]. Return to Step 4 after the genetic operation is completed.

The termination conditions of the algorithm training include: the error between the forward calculated results of the CNN algorithm and the actual results converges to a stable level or the number of training reaches a threshold. The error between the result of forward calculation and the actual result is also the fitness value in the process of population genetic operation. The smaller the error is, the better the scheme represented by the chromosome is.

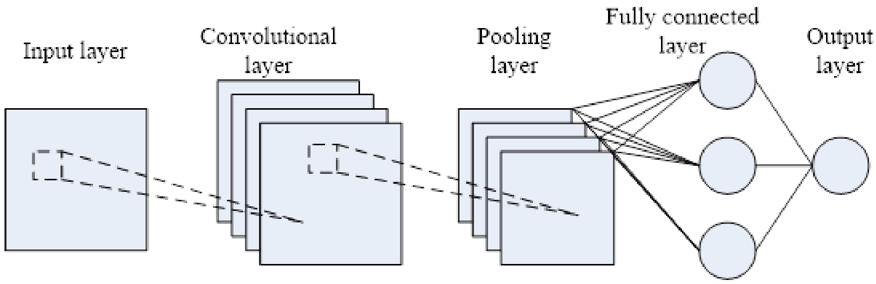


Figure 1: Structure of a convolutional neural network

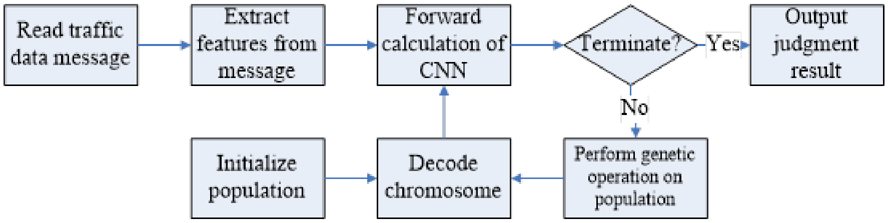

Read traffic data message
Extract features from message
Forward calculation of CNN
Terminate?
Yes
Output judgment result
No
Perform genetic operation on population
Decode chromosome
Initialize population

Figure 2: Training process of the CNN-based network intrusion detection algorithm optimized by a genetic algorithm

Table 1: Traffic features of bidirectional data flows

Traffic feature	Note
Mean packet inter-arrival time	According to the inter-arrival time recorded in the message, the inter-arrival time of adjacent messages can be calculated. Its mean and standard deviation can be calculated.
Standard deviation of message inter-arrival time	
The total number of bytes of the forward message header	The number of bytes in the forward and backward header
The total number of bytes of the backward message header	
Forward message transmission efficiency	Forward (backward) packet transmission time interval under the number of forward (backward) messages per unit.
Backward message transmission efficiency	
Minimum packet length	The length of bytes in the message, from which we can know the shortest message, the longest message, and the average message length. Then, the variance and standard deviation of the message length are calculated.
Maximum packet length	
Average packet length of a packet	
Variance of packet length	
Standard deviation of packet length	
Number of FIN flags in the message	The number of in the messages with FIN flag bit 1
The number of SYN flags in the message	The number of in the messages with SYN flag bit 1
The number of RST flags in the message	The number of in the messages with RST flag bit 1
The number of PSH flags in the message	Number of in the messages with PSH flag bit 1
Number of ACK flags in the message	The number of in the messages with ACK flag bit 1
The number of URG flags in the message	Number of in the messages with URG flag bit 1
Number of CWE flags in the message	The number of in the messages with CWE flag bit 1
Number of ECE flags in the message	The number of in the messages with ECE flag bit 1

3 Simulation Experiments

3.1 Experimental Data

The simulation experiment was carried out in the server of a laboratory, and MATLAB software [6] was used to simulate and analyze the algorithm. The data used in the experiment came from the Canadian Institute for Cybersecurity, which has a large number of different types of available traffic data sets. CIC IoT Dataset 2022 was finally selected as the experimental data set, and the data was collected in the experiment of Internet of things devices. There were six types of data: traffic when power is turned on, traffic when the device is idle, traffic when the device is interactive, traffic when the device is running in different scenarios, traffic when the device is active, and traffic when the device is attacked.

3.2 Experimental Setup

Table 2 presents the relevant parameters of the GA-CNN-based network intrusion detection algorithm. In addition to the GA-CNN algorithm, two other network intrusion detection algorithms were also tested for comparison. They were the SVM-based intrusion detection algorithm and the BPNN-based intrusion detection algorithm. The SVM-based algorithm also used CSV files as features, and the related parameters of the algorithm were set as follows: sigmoid function was used as the kernel function, and the penalty parameter was set to 1. When the SVM algorithm identifies the traffic, it uses the kernel function to map the sample features into a high-dimensional space, calculates in the space to obtain the support vector (hyperplane), and divides the space into two parts, which is suitable to identify whether the traffic is malicious in this paper.

Table 2: Related parameter settings for the GA-CNN algorithm

Parameter name	Value
Population size	20
Mutation probability	0.02
Pooling layer 1	Mean-pooling in a size of 1×2
Pooling layer 2	Mean-pooling in a size of 1×2
Crossover probability	0.3
Convolutional layer 1	32 convolution kernels with a size of 1×3, sigmoid activation function
Convolutional layer 2	32 convolution kernels with a size of 1×3, sigmoid activation function
Maximum number of training	300

The related parameters of the traditional BPNN algorithm were set as follows: the number of nodes in the input layer was 19, the number of hidden layer was set as 2, the number of nodes in each layer was set as 64, the hidden layer used sigmoid activation function, the training used stochastic gradient descent method, the learning rate was set as 0.1, and the maximum training time was 300.

3.3 Evaluation Criteria

$$\begin{cases} P = \frac{TP}{TP+FN} \\ R = \frac{TP}{TP+FP} \\ F = \frac{(\lambda^2+1)\cdot P\cdot R}{\lambda^2\cdot P+R} \end{cases} \tag{2}$$

In the above equations, P is the precision, R is the recall rate, F is the combined value of the precision and recall rate, and λ reflects the importance of the precision and recall rate during evaluation. If both are equally important, the value is 1; if the recall rate is more important, the value is greater than 1; if the precision is more important, the value is less than 1.

3.4 Experimental Results

Among the three network intrusion detection algorithms for comparison, the SVM-based detection algorithm used training samples to fit a "hyperplane" for classification during training, which was different from the gradual iterative convergence of the other twoalgorithms. Therefore, only the training convergence curves of the traditional BPNN algorithm and the GA-CNN algorithm are shown (Figure 3). It can be seen that both algorithms converged during the training process and eventually tended to be stable. In the process of convergence, the GA-CNN method converged to a smaller training error faster, and the training error tended to be stable after about 150 times. The convergence of the traditional BPNN algorithm was relatively slow and tended to be stable after about 250 iterations. In addition, the training error of the GA-CNN algorithm was smaller when the convergence was stable.

The recognition performance of the three network intrusion detection algorithms is shown in Figure 4. It can be seen that the SVM algorithm had the worst recognition performance for abnormal traffic, the traditional BPNN algorithm had better recognition performance than the SVM algorithm, and the network intrusion detection algorithm based on the GA-CNN algorithm had the best recognition performance. The reasons are shown below. When the SVM algorithm identifies traffic, it will first map the characteristics of traffic to a high-dimensional space and then use the support vector (hyperplane) obtained by training to divide the feature space into two

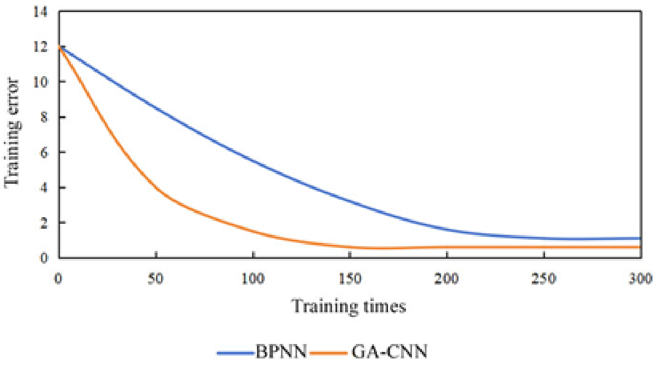


Figure 3: Training convergence curves of the traditional BPNN and GA-CNN algorithms

parts, dividing normal traffic and malicious traffic. Although the principle of this method is simple, the hyperplane itself is a linear division law obtained by mapping nonlinear features to linear features through the kernel function and fitting. It is inevitable to lose some effective information in the mapping process. The activation function used by the BPNN algorithm can effectively fit the nonlinear law, so its recognition performance was better than that of the SVM algorithm. However, when the activation function in the BPNN algorithm faces the input with a large span, the gradient change of the output value is small, which is not conducive to the reverse adjustment during training and easy to fall into the local optimal solution. The improved CNN algorithm not only has the advantage of the CNN algorithm taking into account both local features and global features, but also uses PSO to adjust the parameters in the training process, which avoids the decrease of the output value gradient in the later stage of training.

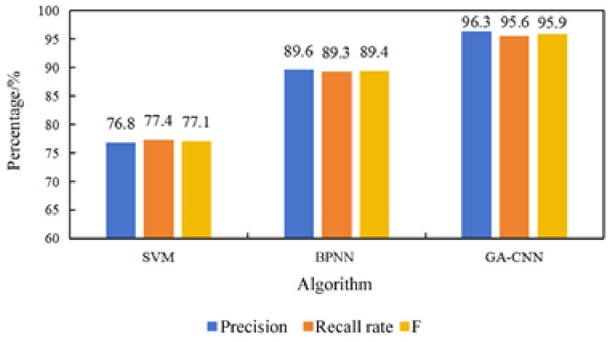


Figure 4: Recognition performance of three network intrusion detection algorithms

The average recognition time of the three network intrusion detection algorithms is illustrated in Figure 5. It can be seen that the average time consumption of the SVM algorithm was the most, the traditional BPNN algorithm was the second, and the GA-CNN algorithm was the least. The reason is that when the GA-CNN algorithm detects the traffic, it combines the features of multiple traffic into a two-dimensional matrix. That is to say, the algorithm identifies multiple traffic at the same time in parallel, so it took the least time.

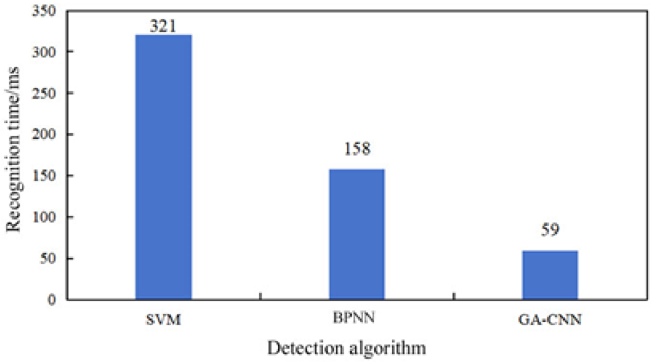


Figure 5: Recognition time consumption of three network intrusion detection algorithms

4 Conclusion

In this paper, the CNN is briefly introduced, and it was used in network intrusion detection. At the same time, a GA was introduced in the training process to adjust the parameters of the CNN algorithm. After that, the performance of the GA-CNN algorithm was tested in the simulation experiment and compared with the SVM and traditional BPNN algorithms. The main results are as follows. (1) The traditional BPNN algorithm converged to stability after about 250 iterations; the GA-CNN algorithm converged to stability after about 150 iterations, and its error was smaller after stability. (2) The SVM algorithm had the worst recognition performance for abnormal traffic, the BPNN algorithm had better performance than the SVM algorithm, and the GA-CNN algorithm had the best performance. (3) The average time consumption of the SVM algorithm was the most, the traditional BPNN algorithm was the second, and the GA-CNN algorithm was the least.

Acknowledgments

This study was supported by First Class Undergraduate Course in Ningxia (2021YLKC246), Scientific Research Project of Ningxia (2022-KY-132) and Research Project of Yinchuan Energy College (2023-KY-Z-3).

References

[1] G. Abdiyeva-Aliyeva, M. Hematyar, "Statistic approached dynamically detecting security threats and

updating a signature-based intrusion detection system's database in NGN," *Journal of Advances in Information Technology*, vol. 13, no. 5, pp. 524-529, 2022.

[2] H. Damania, A. Jagtap, A. Jain, C. Chavan, S. Khonde, "MAIDEn: A machine learning approach for intrusion detection using ensemble technique," *International Journal of Computer Applications*, vol. 179, no. 13, pp. 34-36, 2018.

[3] J. H. Huh, "Implementation of lightweight intrusion detection model for security of smart green house and vertical farm," *International Journal of Distributed Sensor Networks*, vol. 14, no. 4, pp. 1-11, 2018.

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

[5] S. Jose, D. Malathi, B. Reddy, D. Jayaseeli, "A survey on anomaly based host intrusion detection system," *Journal of Physics: Conference Series*, vol. 1000, pp. 1-10, 2018.

[6] A. Karami, "An anomaly-based intrusion detection system in presence of benign outliers with visualization capabilities," *Expert Systems with Applications*, vol. 108, no. OCT., pp. 36-60, 2018.

[7] L. Li, H. Zhang, H. Peng, and Y. Yang, "Nearest neighbors based density peaks approach to intrusion detection," *Chaos Solitons & Fractals*, vol. 110, pp. 33-40, 2018.

[8] W. Li, W. Meng, L. F. Kwok, "Investigating the influence of special on–off attacks on challenge-based collaborative intrusion detection networks," *Future Internet*, vol. 10, no. 1, pp. 1-16, 2018.

[9] I. C. Lin, H. H. Ou, M. S. Hwang, "A user authentication system using back-propagation network", *Neural Computing & Applications*, vol. 14, pp. 243-249, 2005.

[10] N. Lu, Y. Sun, H. Liu, S. Li, "Intrusion detection system based on evolving rules for wireless sensor networks," *Journal of Sensors*, vol. 2018, pp. 1-8, 2018.

[11] H. M. Rais, T. Mehmood, "Dynamic ant colony system with three level update feature selection for intrusion detection," *International Journal of Network Security*, vol. 20, no. 1, pp. 184-192, 2018.

[12] S. Sharma, A. Kaul, "A survey on intrusion detection systems and honeypot based proactive security mechanisms in VANETs and VANET Cloud," *Vehicular Communications*, vol. 12, no. APR., pp. 138-164, 2018.

[13] I. Studnia, E. Alata, V. Nicomette, M. Kaaniche, Y. Laarouchi, "A language-based intrusion detection approach for automotive embedded networks," *International Journal of Embedded Systems*, vol. 10, no. 1, pp. 1-12, 2018.

[14] T. Yang, "A time series data mining based on ARMA and MLFNN model for intrusion detection," *Communications and Computers*, vol. 003, no. 007, pp. 16-30, 2019.

[15] R. Vijayanand, D. Devaraj, B. Kannapiran, "Intrusion detection system for wireless mesh network using multiple support vector machine classifiers with genetic-algorithm-based feature selection," *Computers & Security*, vol. 77, no. AUG., pp. 304-314, 2018.

[16] C. R. Wang, R. F. Xu, S. J. Lee, and C. H. Lee, "Network intrusion detection using equality constrained-optimization-based extreme learning machines," *Knowledge-Based Systems*, vol. 147, no. MAY1, pp. 68-80, 2018.

[17] H. J. Wu, Y. H. Chang, M. S. Hwang, I. C. Lin, "Flexible RFID location system based on artificial neural networks for medical care facilities", *ACM SIGBED Review*, vol. 6, no. 2, pp. 1-8, 2009.

[18] X. Wu, C. Zhang, R. Zhang, Y. Wang, J. Cui, "A distributed intrusion detection model via nondestructive partitioning and balanced allocation for big data," *Computers, Materials and Continua*, vol. 2018, no. 7, pp. 61-72, 2018.

Biography

Jie Li, born in Shaanxi in 1982, graduated from China University of Petroleum (East China) in 2014 and is an associate professor of Yinchuan University of Energy. His research interests include computer applications, data analysis and mining, and data and information visualization.

Jing Li, born in Hebei in 1981, graduated from Xi'39;an Jiaotong University in 2017 and is an associate professor of Yinchuan University of Energy. Her research interests include power system, electrical automation, and computer network.

Privacy Preserving Scheme in Mobile Edge Crowdsensing Based on Federated Learning

Jian Chen[1], Linming Gong[1,2], and Jingyu Chen[1]
(Corresponding author: Jian Chen)

School of Computer Science, Xi'an Polytechnic University[1]
Xian 710048, China
The Shaanxi Key Laboratory of Clothing Intelligence, School of Computer Science, Xi'an Polytechinic University[2]
Xian 710048, China
Email: jianchenxpu@163.com

(Received July 2, 2023; Revised and Accepted Mov. 22, 2023; First Online Oct. 24, 2023)

The Special Issue on Advanced Security System for Network in Fog and Edge Computing
Special Editor: Dr. Salim El Khediri (Qassim University), Dr. Virginia Pilloni (University of Cagliari),
Dr. Mohamed Lahby (Hassan II University ENS Casablanca)

Abstract

As an emerging data collection model, mobile crowdsensing can collect large amounts of data efficiently and quickly. By applying distributed machine learning to crowdsensing systems, we can fully exploit the value of the sensing data and make more accurate predictions. Studies have shown that attackers can infer privacy information from the trained model. However, existing solutions do not consider protecting the positive and negative signs of the model parameters. Therefore, this paper proposes a Mobile Edge Crowdsensing scheme based on Federated Learning (FL-MECS). The scheme aims to provide efficient sensing data analysis and effective privacy protection methods for crowdsensing systems. The approach utilizes local training and initial aggregation of the model at the edge node to avoid the risk of privacy leakage caused by the direct transmission of sensing data to the aggregation node. The evaluation shows that FL-MECS can withstand collusion attempts and effectively preserves the model's local positive and negative sign parameter information. Experimental results on the MNIST dataset show that FL-MECS outperforms other alternatives in terms of model correctness.

Keywords: Edge Computing; Federated Learning; Mobile Crowdsensing; Privacy-Preserving; Zero-sum Noise

1 Introduction

Mobile Crowdsensing (MCS) is a new data collection model that combines mobile sensing and crowdsourcing [19], which can collect large amounts of data efficiently and quickly. MCS has been widely used in various fields such as intelligent transportation, environmental monitoring, air quality detection, and healthcare. However, there may be potential security risks in the collection, transmission, and aggregation of data, especially in the presence of untrusted sensing platforms and attackers. Without appropriate privacy measures, sensitive information such as personal data, location trajectories, and identity information may be exposed [2,25]. In addition, the central server of the MCS faces enormous computational and communication demands due to the widespread engagement of sensing users and the continuous expansion of application scenarios.

Applications of Edge Computing (EC) [9] technology in MCS can increase system efficiency. For example, it can relieve the pressure on the server under the traditional central computing mode [4] and provide faster response times. Although edge nodes provide a new way of data processing and transmission, they can be semi-honest [1]. Because edge nodes run on edge devices, edge devices have certain read and write permissions to the code and algorithms on the nodes. This semi-honesty poses some risks, as attackers can compromise the security of the system by tampering with algorithms or stealing privacy data, leading to sensing data leakage and personal security threats.

To address this problem, Federated Learning (FL) technology has been applied to the MCS field [8, 11, 12, 26].The basic concept of FL is to minimize the privacy risks associated with direct data transfer to edge nodes by first using aggregation algorithms [3] to obtain the final global model from data that numerous sensing workers have locally trained to create local models.

Despite the privacy benefits of FL technology, several researchers have demonstrated [6, 13] that attackers can still use shared local models to infer sensitive information from the dataset. If the models of the sensing workers are trained from small and shared datasets, attackers can even reconstruct the privacy information of the original data by stitching these models together. This type of attack is called a "model inversion attack," where the key is that attackers can use shared local models to infer sensitive information from the original dataset. This indicates that while FL can significantly reduce the risk of data leakage, further strengthening of its security protections is still needed.

To further improve the privacy strength of FL technology, researchers have proposed various solutions. Among them, methods based on differential privacy (DP), anonymity mechanisms, and homomorphic encryption (HE) are common user privacy protection technologies [17, 21].DP technology ensures that privacy is protected by adding appropriate noise processing, which reduces the risk of potential data breaches. The anonymity mechanism uses relabeling, generalization, or complete data deletion to depersonalize data and prevent data breaches. However, the use of DP technology could lead to accuracy errors when updating model parameters, and the anonymity mechanism could be exposed to background knowledge attacks. On the other hand, HE technology doesn't require parameter adjustments and is better suited for high-level privacy protection. It can group and share models in FL without much impact on model accuracy and security, making it a widely used technology for privacy protection in data grouping [14, 24, 26, 27]. However, when dealing with large-scale FL, there are numerous model parameters. Using the Paillier encryption scheme directly to group models would take a lot of time.

In addition, FL [7] requires that both the positive and negative signs of local model parameters be taken into account to protect the privacy of MCS employees. Especially in linear regression tasks, the polarity of the model corresponds to the positive or negative correlation between a variable and the dependent variable [16]. To protect the positive and negative signs of the local model privacy, DP technology can be used. Specifically, using the noise amplitude of DP to change the polarity of the model to prevent the aggregation node from directly obtaining the polarity value of the model. However, in traditional DP, in order to change the positive and negative signs of the local model, the amplitude of the noise must be increased, but this will inevitably worsen the impact on the accuracy of the model. Therefore, the application of FL technology to mobile edge crowdsensing also requires the exploration of more efficient privacy protection techniques.

Protecting the privacy and security of crowdsensing workers while facilitating improved crowdsensing data mining has emerged as a critical challenge. In this paper, we present a privacy protection scheme for mobile edge crowdsensing using federated learning (FL-MECS). This builds on the proposal described in [15]. The system uses edge nodes to perform initial aggregation on the local models of sensing workers. Then, after local training, it introduces zero-sum noise into the model parameters to protect the private models of sensing workers. In particular, the polarity of the model parameters is not exposed, thus protecting the local sensing data of the sensing workers. The main contributions of this paper are listed below:

1) The proposed FL-MECS efficiently trains and combines user data in sensing, effectively lessening the burden on the sensing platform. It is important to note that this approach does not necessitate the transfer of sensing data to the sensing platform.

2) To safeguard the privacy of crowdsensing workers' local models, we have opted for the method of adding zero-sum noise to the model parameters. This approach prevents the sensing platform from directly inferring information from the local models and resolves the issue of exposing the polarity information of the model parameters of sensing workers during the aggregation process. Not only does this method preserve the privacy of crowdsensing workers, but it also ensures the efficacy of the data. To the best of our knowledge, FL-MECS is the initial program to ensure that the model's polarity remains encrypted from the aggregator without encrypting the model itself.

3) In addition, we simulated the data collection process of FL-MECS, specifically deploying a convolutional neural network on the MNIST dataset to simulate the collection and mining of sensing data using FL. We also measured the accuracy of the model and the time consumption of each module.

The subsequent sections are structured as follows: Firstly, the technical overview section introduces the pertinent technologies involved in this paper. Subsequently, the FL-MECS section offers a comprehensive insight into the proposed scheme's design and implementation. The ensuing scheme analysis section scrutinizes the correctness and security of FL-MECS, the performance evaluation segment assesses the privacy protection strength and availability of FL-MECS through simulation experiments. Finally, the concluding section provides a comprehensive summary of the scheme, highlights its limitations, and suggests future research directions.

2 Technical Overviews

This section introduces the idea of zero-sum noise technology and gives an overview of basic concepts and Federated Averaging (FedAvg) learning.

2.1 Zero-sum Noise

Compared to Differential Privacy methods, zero-sum noise [20] technology defends data privacy by adding a distinct type of chance variation to private data. This technique creates separate random variations for collective data, ensuring the total adds up to zero. After joint work, the variation is combined with private data and sent to the aggregator. By taking out the collected zero-variation during aggregation, employees' privacy remains secure. This coordinated approach boosts privacy protection.

The basic process for using zero-sum noise technology is straightforward. Suppose participant N^i has private data M_i (where i ranges from 1 to m). Then aggregator N^0 can calculate $\alpha = \sum_{i=1}^{m} M_i$ without learning any of the individual participant's private information. For added security, zero-sum noise technology generates random matrices and collaboratively creates zero-sum matrices to add noise to the private data. Each person, labeled N^i, creates a total of m random matrices called p_i^k. These matrices must add up to zero, and they should be the same size as their own private matrices M_i. N^i chooses one random matrix to keep while sending the remaining $m-1$ matrices to other participants. By repeatedly going through this procedure, every participant gets randomized matrices from other participants and creates a new random matrix P_i with their saved random matrix. P_i meets the requirement of $\sum_{i=1}^{m} P_i = 0$. Participants submit the processed confidential data $\alpha_i = M_i + P_i$ to the aggregator, who computes the aggregate outcome without awareness of the initial confidential data of each participant, as shown in Equation (1).

$$\alpha = \sum_{i=1}^{m} \alpha_i = \sum_{i=1}^{m} M_i + P_i = \sum_{i=1}^{m} M_i. \quad (1)$$

2.2 Federal Average Algorithm

The FedAvg is a model training algorithm for FL [15]. It allows models to be averaged using weights across several data centers, eliminating the need to transfer data to a trusted third-party server. In FedAvg, if there are K parties participating with their own local datasets $D^i = D^1, D^2, \ldots D^K$, each party trains the model using their local dataset D^i and the initial model ω_t. Once the updates are ready, they are shared with the server. During the aggregation process, the aggregator calculates the mean of local models ω_{t+1}^i from all participating parties and updates the global model ω_{t+1}. Then, the new weights of the overall model are redistributed to each participating party as their starting weights for the next round of iteration. Using this method, the FedAvg algorithm can train the model while preserving data privacy.

3 FL-MECS

In this section, we explained how FL-MECS works and outlined the processes for protecting data privacy, updating the model, and performing secure aggregation using privacy-preserving algorithms.

3.1 System Model

We assume that the ENs and Crowdsensing server are entities that are semi-honest and follow the protocol in relation to the FL-MECS framework [5, 22]. However, these entities are also interested in the private models of Sensing workers, specifically the sensing data, and try to access their training models to infer the private sensing data. The system model of the FL-MECS framework consists of five modules.

Task publisher (TP). TP publishes tasks that require workers to collect sensing data for the purpose of training a machine learning model that represents the sensing data. The workers then outsource the model to the sensing platform to be effectively aggregated. TP also provides an initial model to the sensing platform when publishing the task.

Crowdsensing server (CS). CS receives data requests from TP and hires sensing workers. During data collection, CS works with SWs to train the model and delivers aggregated model data to TP.

Sensing workers (SWs). SWs collect data using mobile devices and analyze it locally using a trained model to extract valuable insights. To make sure that sensitive data is not leaked, we will process the local model for privacy before uploading it. After privacy processing, we will send the resulting model to an edge node for initial aggregation.

Edge nodes (ENs). ENs are crucial components of the FL-MECS framework, as they have efficient storage and computing capabilities and primarily handle edge model training and preliminary model aggregation.

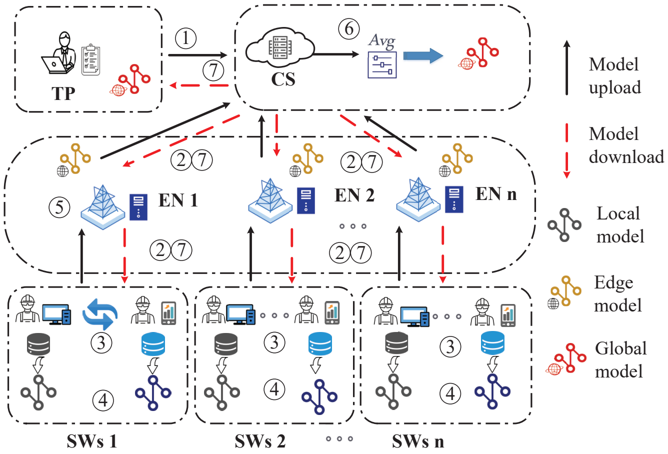


Figure 1: FL-MECS Privacy Processing and Model Update Workflow Diagram

3.2 System Workflow

Figure 1 shows how privacy processing and model updates work in FL-MECS. TP requests data, then CS sends the task to ENs' SW. Workers train sensing data and add zero-sum noise to local model parameters for secure data collection and protection during FL. The model first gathers data at the ENs, and then merges it into a worldwide model at the CS. The finished model is then handed over to the TP. Here are the specific steps:

Step 1: Task publishing. The TP sends a request to the CS with task details, the number of SWs needed, and the deadline. Additionally, the TP must include an initial model so that software workers can use it for local model training.

Step 2: Worker recruitment. After registration, SWs will communicate with the CS through locally managed edge servers. When CS gets a task request, it will send the task to appropriate users to take part in the sensing task.

Step 3: Collaborative generation of zero-sum noise. SWs under the same ENs jurisdiction collaborate to generate a matrix that sums up to zero, referred to as a zero-sum noise matrix.

Step 4: Local training and privacy preservation. SWs obtain data and conduct model training locally to yield a local model after receiving a task. After adding zero-sum noise to the model parameters for privacy protection, the privacy-preserving model is uploaded to the ENs.

Step 5: Preliminary model aggregation. ENs collect all privacy-preserving models from SWs for preliminary model aggregation and updating.

Step 6: Global model aggregation. Following the same idea as ENs, CS gathers all initial global models sent by ENs. CS assigns weight values based on the differing number of SWs linked to each EN, and then applies the federated averaging algorithm to update the global model. Lastly, the updated global model is sent to the corresponding SWs.

Step 7: Iterative Training and Data Delivery. SWs have the option to further iterate on the global model to gradually improve its accuracy and usability until the predetermined number of iterations is reached. Once the global model is updated, CS delivers the task results to the sensing TP.

3.3 Zero-sum Noise Generation and Model Noising

The process of collaborative generation of zero-sum noise and addition of noise to model parameters is as follows: Assuming that SWs update their local models $\omega^{ig}(g = 1, \ldots, m)$ based on private data D^i and initial models. The model parameters m refer to the weights and biases of the model. The aggregation node N^0 needs to obtain the average sum $N^i(\varpi{=}\frac{1}{K}\sum_{i=1}^{K}\omega^i)$ without knowing the private data of the SWs.

First, the participating party N^i randomly generates a noise matrix $n^i(i = 1, \ldots, K)$ with i rows and g columns. The number of rows in the matrix n^i is the same as the number of participating parties K, and the number of columns is the same as the number of parameters m in ω^i. Each column of the matrix n^i sums up to zero.

Before uploading the model, the participating party N^i keeps one row of the noise matrix n^i and sends the remaining rows one by one to other SWs. Upon receiving noise vectors from other participating parties, N^i generates a new noise matrix P^i by combining the received noise vectors with its own retained noise vector. Since P^i itself is a matrix composed of multiple zero-sum noise matrices, it satisfies $\sum_i^K P^i = 0$, where the summation represents the addition of all elements in the same column of the noise matrices P^i.

Finally, N^i combines the new model with zero-sum noise with its own model to generate a new model $\varpi^i = \omega^i + P^i$. The aggregation node only needs to integrate the new models from each SW after privacy processing into a global model $\varpi = \frac{1}{K}\sum_{i=1}^{K}\varpi^i$.Figure 2 illustrates an example of the model generated by three SWs within the same ENs jurisdiction after collaborative generation and addition of zero-sum noise. Algorithm 1 shows the specific process of the model noise addition algorithm.

3.4 Model Updating and Aggregation Algorithm

After users are selected as SWs, they download tasks and the initial model ω_t from ENs and collect relevant data using their own mobile devices. SWs then replace the initial model ω_t with the local model ω_t^i. Next, SWs invoke the Stochastic Gradient Descent (SGD) algorithm. In other words, for each sample in the training set, the gradient of its cost function is calculated, and the model parameters are updated in the direction of the gradient, and the optimal solution of the model parameters is found by minimizing the loss function.

ENs are located near the SWs and their main task is to collect the local models from the SWs for preliminary aggregation and upload them to the CS. They also download the global model and broadcast it to the sensing SWs within their jurisdiction. In this paper, ENs are considered as semi-honest entities, i.e., they are curious about the data from the sensing SWs and try to infer the sensing data from the local models of the SWs. Therefore, noise must be added to the preliminary aggregated model before it is uploaded to ensure the privacy and security of SWs. CS, as the core component of the system, is responsible for aggregating the preliminary models uploaded by ENs into a global model and finally delivering the fitted or maximum-iteration global model to TP. Algorithm **??** demonstrates the aggregation process, where sensing data

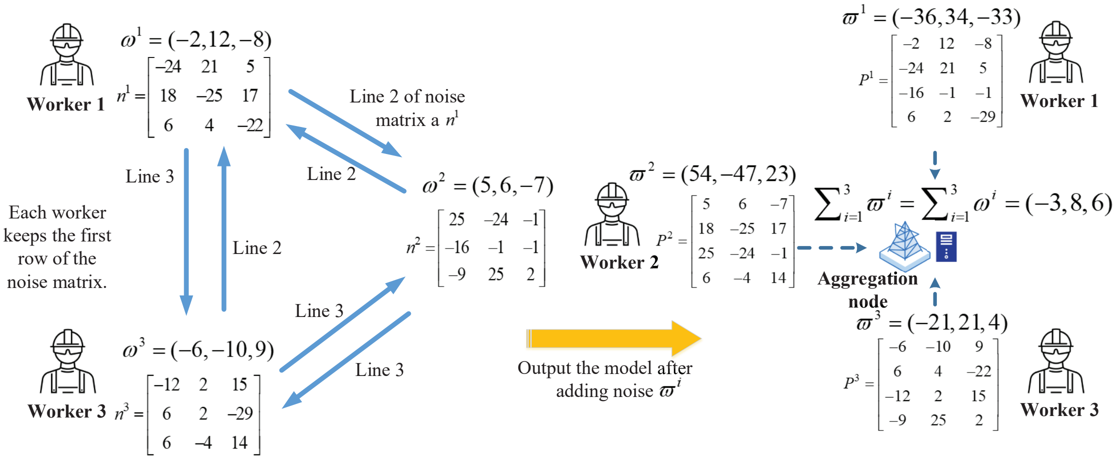

Worker 1
$\omega^1 = (-2,12,-8)$
$n^1 = \begin{bmatrix} -24 & 21 & 5 \\ 18 & -25 & 17 \\ 6 & 4 & -22 \end{bmatrix}$
Line 2 of noise matrix a n^1
Line 3
Line 2
Line 2
Each worker keeps the first row of the noise matrix.
$\omega^2 = (5,6,-7)$
$n^2 = \begin{bmatrix} 25 & -24 & -1 \\ -16 & -1 & -1 \\ -9 & 25 & 2 \end{bmatrix}$
Line 3
Line 3
Worker 3
$\omega^3 = (-6,-10,9)$
$n^3 = \begin{bmatrix} -12 & 2 & 15 \\ 6 & 2 & -29 \\ 6 & -4 & 14 \end{bmatrix}$
Output the model after adding noise ϖ^i
$\varpi^1 = (-36,34,-33)$
$P^1 = \begin{bmatrix} -2 & 12 & -8 \\ -24 & 21 & 5 \\ -16 & -1 & -1 \\ 6 & 2 & -29 \end{bmatrix}$
Worker 1
$\varpi^2 = (54,-47,23)$
Worker 2
$P^2 = \begin{bmatrix} 5 & 6 & -7 \\ 18 & -25 & 17 \\ 25 & -24 & -1 \\ 6 & -4 & 14 \end{bmatrix}$
$\sum_{i=1}^{3} \varpi^i = \sum_{i=1}^{3} \omega^i = (-3,8,6)$
Aggregation node
$\varpi^3 = (-21,21,4)$
$P^3 = \begin{bmatrix} -6 & -10 & 9 \\ 6 & 4 & -22 \\ -12 & 2 & 15 \\ -9 & 25 & 2 \end{bmatrix}$
Worker 3

Figure 2: Collaborative Generation of Zero-Sum Noise and Model Addition Example

is trained locally and privacy-preserved before being sent to the corresponding ENs. The local models are preaggregated at ENs and sent to CS for global model aggregation. Finally, CS uses an averaging algorithm to obtain the global model.

Algorithm 1 Model noise addition algorithm

```
Input: Local model parameter $\omega^i$ of worker $N^i$.
Output: Model $\varpi^i$ with zero-sum noise added.
worker $N^i(i = 1, \dots, k)$is randomly distributed in $n$ edge nodes and forms a perceived worker group $E^j(j = 1, \dots, n)$.
for $j = 1, \dots, n$ do
    $N_j{}' = \{N_j^1, \dots, N_j^e\}$ is selected from $E^j$.
    The number of rows of noise $\{P_j^1, P_j^2, \dots, P_j^e\}//P_j$ generated by SWs is the number of workers in the $j$-th edge node, and it is listed as the total number of model parameters.
    for $i = 1, \dots, e$ do
        $\varpi_j^i \leftarrow \omega_j^i + P_j^i.//\varpi_j^i$ represents the local model of the $i$ worker under the jurisdiction of the $j$ edge node after noise processing.
    end for
end for
End
```

4 Scheme Analysis

The main focus of this section is to analyze the correctness of FL-MECS and evaluate its security. Additionally, we also analyze the ability of individual sensing workers under the same edge node in FL-MECS to resist collusion attacks.

Algorithm 2 Aggregation algorithm

```
Input: Private data $D = D^1, D^2, \dots D^K$ of worker $N^i$.
Output: Global model $\omega_t$.
Local training of sensing workers: local training will be conducted for perceptual worker $N^i$ according to the collected perceptual data and the previous model (or initial model $\omega_0$) to generate local model $\omega_t^i$, and the local model and the trained sample size $|D^i|$ will be sent to the edge node.
Initial aggregation of edge nodes: edge node $E^j(j = 1, \dots, n)$ sets initial values $D^{edge} \leftarrow 0$ and $Sum^{edge} \leftarrow 0$.
for $i = 1, \dots, K$ do
    $Sum^{edge} \leftarrow Sum^{edge} + |D^i| \varpi_t^i$,$D^{edge} \leftarrow D^{edge} + |D^i|$.
end for
Crowdsensing server aggregation global model: set initial values $D^{global} \leftarrow 0$ and $Sum^{global} \leftarrow 0$.
for $j = 1, \dots, n$ do
    $Sum^{global} \leftarrow Sum^{global} + Sum_j^{edge}$,$D^{global} \leftarrow D^{global} + D_j^{edge}$.
end for
if $t \neq E$ then
    Send $D^{global}$ and $Sum^{global}$ to each sensing workers.//When the maximum number of training rounds is not reached.
end if
Sensing workers computing global model $\omega_t \leftarrow \frac{Sum}{D^{global}}$.
End
```

4.1 Correctness Analysis

The correctness of FL-MECS means that the locally processed models can be correctly aggregated by the aggregation entity. Since the local models are sent to the ENs after undergoing local privacy processing, and the process of sending them to the CS by the ENs does not require privacy processing operations, we only discuss in this paper the ability of the ENs to correctly aggregate the model parameters after adding noise, ensuring that the aggregated model parameters do not contain additional noise. We use theoretical proofs to verify that the aggregated model result at aggregation node N^0 is equal to the sum of the models from individual sensing workers N^i.

Lemma 1. *The model parameters in the FL-MECS scheme can be correctly aggregated by the ENs.*

Proof. Suppose there are K sensing workers, denoted as $N^i(i = 1, \ldots, K)$, each having a private model vector denoted as $\omega^{ig}(i = 1, \ldots, K, g = 1, \ldots, m)$. After undergoing local privacy processing at the SWs, these vectors are sent to the ENs. We want to prove the following Equation (2):

$$\delta = Sum^{edge}. \tag{2}$$

Where δ represents the sum of local models from participating nodes $N^i(i = 1, \ldots, K)$ under the jurisdiction of ENs, and Sum^{edge} represents the result of the aggregated model by ENs.We can easily obtain $\delta = \sum_i^K \omega^i$, which represents the locally processed models of the participating nodes sent to ENs for aggregation, denoted as $Sum^{edge} = \sum_i^K \omega^i + P^i$. Furthermore, since it was shown in Section 3.3 that $\sum_i^K P^i = 0$ satisfies the collaborative design zero-sum noise, we can derive the following Equation (3):

$$Sum^{edge} = \sum_{i=1}^{K} \omega^i + P^i = \sum_{i=1}^{K} \omega^i = \delta. \tag{3}$$

It can be concluded that the local model parameters at the SWs can be correctly aggregated at the ENs after undergoing local privacy processing. The CS only needs to collect the preliminary aggregated models from each EN to aggregate the global model. Therefore, FL-MECS is correct. □

4.2 Security Analysis

In this section, we evaluate the security of FL-MECS, and the results show that the local models and sensing data of the SWs in FL-MECS are not leaked to the ENs, CS, and TP. In addition, we also analyze the ability of FL-MECS to withstand collusion attacks.

Lemma 2. *The algorithm 2 is a secure algorithm for aggregating training models, meaning that no entity in the system can infer the private model of any individual worker.*

Proof. After collecting the sensing data, the SWs must train their own models locally before uploading them to the ENs. Before uploading, SWs introduce zero-sum noise into their local models. Eliminating this noise requires the participation of all SWs under the same ENs. In addition, the ENs do not have access to any information other than receiving the models from each SW after the zero-sum noise has been added. Thus, the ENs cannot obtain or infer the private models of individual workers from the locally trained models with added noise. After collecting the sensing data, the SWs must locally train their own models before uploading them to the ENs. Before uploading, SWs introduce zero-sum noise into their local models. Eliminating this noise requires the participation of all SWs under the same ENs. In addition, the ENs do not have access to any information other than receiving the models from each SW after the zero-sum noise has been added. Thus, the ENs cannot obtain or infer the private models of individual workers from the locally trained models with added noise. Therefore, Lemma 2 holds. □

Lemma 3. *FL-MECS can resist collusion attacks under certain conditions.*

Proof. We assume that the ENs and CS do not collude with each other, which is a common assumption in other schemes [22]. Therefore, we only discuss the maximum number of SWs under the jurisdiction of the same ENs that can resist collusion attacks. Let's assume there are $n(n \geq 3)$ workers, and the i-th worker receives the local parameters ω^i after local training. The worker also generates a random noise matrix n^i, while the matrix n^i is generated independently by each worker, with each element being random. However, the matrix n^i satisfies the condition that the sum of each column is zero. After collaborative aggregation, the noise-added matrix of the i-th worker is P^i, and the parameters after adding noise are ϖ^i. We can express this relationship as follows Equation (4):

$$\varpi^i = \omega^i + P^i. \tag{4}$$

To illustrate this, let's assume there are only three workers, and their noise-added matrices are P^1, P^2, and P^3, respectively. In this case, the noise-added parameters for the first worker can be represented as Equation (5):

$$\varpi^1 = \omega^1 + P^1. \tag{5}$$

The other two workers added noise to the local model parameters, resulting in ϖ^2 and ϖ^3.The aggregator receives three noise-added model parameters and combines them by weighted averaging to obtain the final global model ω. This process can be represented as Equation (6):

$$\omega = \frac{1}{n}\sum_{i=1}^{n} \varpi^i = \frac{1}{n}\sum_{i=1}^{n} (\omega^i{+}P^i). \tag{6}$$

Due to the fact that the sum of each column is equal to zero, we have the Equation (7):

$$\sum_{i=1}^{n} P_j^i = 0, j = 1, 2, \ldots d. \tag{7}$$

Now, let's assume that there are k colluding workers with their noise matrices $P^1, P^2, \ldots, P^k$, and k workers aggregating the noise as $\sum_{i=1}^{k} P_j^i = 0$ (only when $k = n$ holds true). Let's consider the first worker as the target of the attack. When $k = n - 1$ is true, we have the Equation (8):

$$\omega^1 + P^1 = n\omega - \sum_{i=2}^{n} (\omega^1 + P^i). \qquad (8)$$

In this case, P^1 can be inferred by other colluding workers, leading to the inference of ω^1 as well. When $k = n - 2$ is true, we have the Equation (9):

$$\omega^1 + \omega^2 + P^1 + P^2 = n\omega - \sum_{i=3}^{n} (\omega^1 + P^i). \qquad (9)$$

Since there are two workers who are not involved in collusion, other workers cannot know the random numbers exchanged between the first and second workers. As a result, P^1 and P^2 cannot be inferred, and the model ω^1 of the attacked worker is not leaked. Therefore, it can be concluded that our approach still preserves the privacy of the local model parameters when there are $k \leq n - 2$ colluding workers. □

4.3 Scheme Comparison

To further demonstrate the effectiveness of our proposed FL-MECS scheme, we compare it with other similar schemes [15, 18, 23, 27] in recent years, focusing on model confidentiality, noise addition, edge nodes, and positive and negative sign protection of model parameters. Table 1 shows the comparison of these relevant schemes with FL-MECS in terms of privacy features. (Note: 'Y' indicates the existence of the method or feature, 'N' indicates the non-existence, '-' indicates that HE technology is used to protect the model, so the privacy protection of positive and negative sign model parameters is not discussed in this scheme).

Table 1: Features provided by different schemes

Schemes	[15]	[18]	[23]	[27]	Ours
Model protection	N	Y	Y	Y	Y
Noise addition	N	Y	N	N	Y
Edge nodes	N	N	Y	N	Y
sign protection	N	-	N	-	Y

5 Performance Evaluation

In this section, we first analyzed the impact of the positive/negative nature of the model parameters by using the ratio of the noise value range to the model parameter value range to evaluate the privacy protection strength of FL-MECS. In addition, we simulated the horizontal federated learning task of FL-MECS on a real dataset and statistically analyzed the accuracy and running time of the model, which fully demonstrated the superiority and practicality of our scheme.

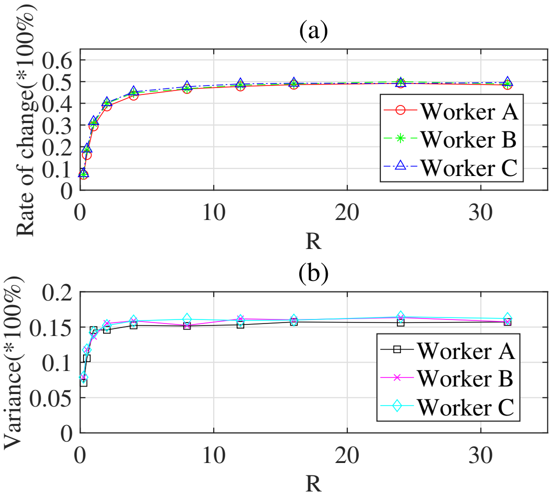


Figure 3: The effect of the range of zero-sum noise on the sign of model parameters

5.1 Analysis of Privacy Protection Strength

We conducted a large number of experiments in a simulation environment to observe the effect of zero-sum noise on the positive/negative signs of the model parameters, and to find the most appropriate noise value distribution to achieve the best privacy protection, thus analyzing the privacy protection strength of the FL-MECS scheme.

We initially think that adding zero-sum noise will change the positive and negative of local model parameters. When the probability that the sign of the local model parameter changes is closer to 50%, the probability that the attacker judges whether the sign is positive or negative is the same, as shown in Equation (10),

$$\Pr[- \leftarrow \omega_{signs}] = \frac{1}{2}, \Pr[+ \leftarrow \omega_{signs}] = \frac{1}{2}. \qquad (10)$$

This makes the sign change of the model appear semantically safe, thus reducing the attacker's ability to guess these signs and improving the level of protection. To investigate this, we ran a simulation with three Sensing Workers (SWs) who collectively added noise and aggregated the models. We observed the proportion of sign changes in the model parameters after incorporating zero noise. For testing purposes, we set the positive and negative signs of local model parameters to uniform distribution, and monitor the change rate by adjusting the range of noise matrix elements.

We denote the ratio of the range of noise values to the range of model parameter values as R. Figure 3(a) shows

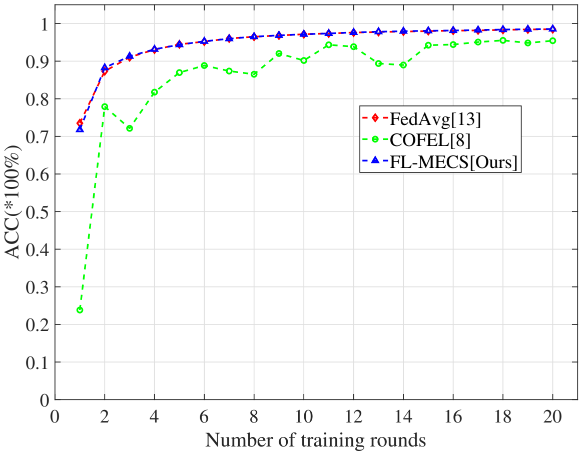


Figure 4: Model accuracy comparison on IID for different approaches

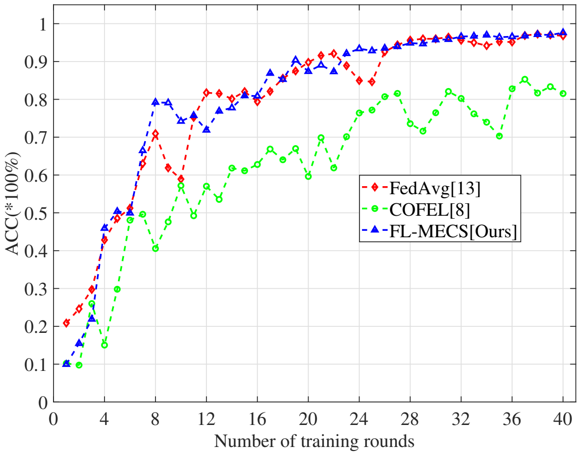


Figure 5: Model accuracy comparison on non-IID for different approaches

the rate of change of the sign of the model parameters for different R values among the SWs. Figure 3(b) shows the variance of the rate of change of model parameters across different R values. To ensure the reliability of the data, we performed 1000 replicate experiments for each data point and calculated their averages. The results show that when R is greater than 16, the model parameter change rate stabilizes around 50% and the variance remains stable, resulting in optimal privacy protection.

5.2 Experiment and Efficiency Evaluation

We evaluated the effectiveness of the FL-MECS approach on the MNIST dataset. Specifically, we simulated a mobile edge crowdsensing network environment with evenly distributed SWs as clients on each EN and a fixed number of 5 ENs, with the CS as the final global aggregation node. Since the data collected by the sensing devices in the MCS are mostly non-independent and identically distributed (non-IID), we divided the data set into two types: independent and identically distributed (IID) and non-independent and identically distributed, and uniformly assigned them to each SW. In addition, we used a standard convolutional neural network (CNN) as the global model. The model consists of three convolutional layers, two pooling layers, and two fully connected layers, with ReLU as the activation function. MNIST is one of the most commonly used datasets in machine learning. It contains a large number of handwritten digit images with a size of 28*28 pixels, grayscale images, and a value range of [0,255].

The experimental setup and configuration were as follows: The CPU used was an Intel(R) Core(TM) i7-8550U @1.80GHz, with 8GB of RAM, running a 64-bit Windows 10 operating system. The software environment included Python 3.11.0, Torch 2.0.0+CPU. The parameters for the FL were configured as follows: the learning rate was set to 0.001, the local training epoch was set to 1, and the participation rate of the SWs in each training round was set to 100%. The number of client-server communications, i.e., global aggregation iterations, was set to 20, with a batch size of 100.

According to Figure 4 and Figure 5, the performance of FedAvg [15], COFEL [10], and FL-MECS on IID and non-IID datasets are compared based on model accuracy (Acc). From Figure 4, it can be observed that the model accuracy of our approach on the IID dataset is similar to that of the conventional FedAvg approach. After the 10th iteration, the model accuracy of both approaches is higher than 95%, and the accuracy of each round is significantly higher than that of the DP approach. This is because the zero-sum noise added in our approach can cancel each other out during aggregation without affecting the accuracy of the model. On the other hand, the personalized privacy of the DP approach provides a way to balance privacy protection and data utilization to protect individual privacy based on the privacy budget of the approach. However, when the privacy budget is small, the interference with the data is significant, which will inevitably affect the accuracy of the model. From Figure 4, it can be seen that the model accuracy of our approach on the non-IID dataset is still consistent with that of the conventional FedAvg approach. After the 40th iteration, the models of both approaches tend to fit the data, while the DP approach requires more training rounds to achieve a good fit.

Since no existing approach has been found to achieve the same privacy function as our proposed approach, and since model training, model aggregation, and zero-sum noise adding are the main time consuming operations in FL-MECS, they directly affect the performance and efficiency of the whole system. Therefore, we only considered the model training time, model aggregation time,

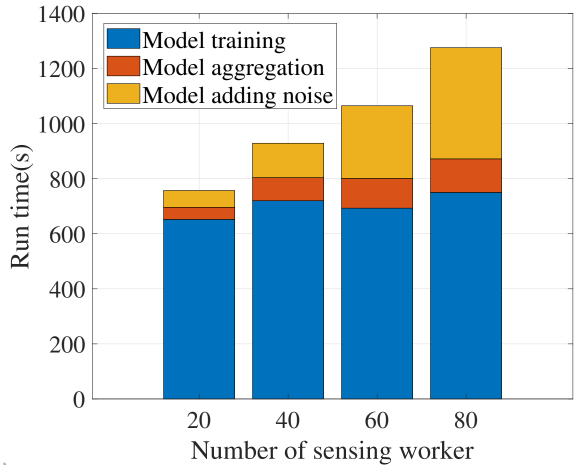


Figure 6: The main time consumption of the FL-MECS approach

and model noise adding time of FL-MECS with 20, 40, 60, and 80 participating SWs. According to Figure 6, we can see that the model training time did not change much when the number of workers increased from 20 to 80, averaging about 700 seconds. The time required to add zero-sum noise increased as the number of SWs increased, because the total number of workers increased, and the number of SWs in the same ENs also increased, resulting in more frequent generation and exchange of noise matrices. However, overall, the time consumption of the system meets the requirements of practical applications. In summary, our proposed approach meets the expected design requirements in terms of model accuracy and time consumption.

6 Conclusions

This paper proposes a novel FL-MECS approach based on federated learning for mobile edge crowdsensing, which provides a feasible method for privacy protection in the research field of mobile crowdsensing. Our method introduces zero-sum noise into the local model, which not only protects the privacy of the sensing data, but also does not affect the utility of the data. In addition, we investigated the effect of the amplitude of the zero-sum noise on the privacy protection strength of the approach through a large number of simulation experiments. We also conducted simulation experiments on the MNIST dataset to verify the effectiveness of the FL-MECS approach. The results show that the accuracy of FL-MECS is comparable to traditional federated learning and, in most cases, higher than the accuracy of differential privacy-based approaches. However, our approach still has some limitations. For example, in the model aggregation process, the aggregator needs to collect all the models sent by sensing workers in order to perform aggregation. However, in mobile edge crowdsensing, due to the computational power of devices and network latency issues, some sensing workers may not be able to upload local models in a timely manner, which increases the risk of inaccurate or failed model aggregation. Therefore, in future work, we will conduct deeper research to improve the system's dynamic entry and exit mechanisms.

Acknowledgments

This work is supported by Natural Science Basic Research Program of Shaanxi (No.2021JM-453) and Key projects of Education Department of Shaanxi Province (23JS025), and is partly supported by The National Natural Science Foundation of China (No.61972225), Natural Science Basic Research Program of Shaanxi (No.2023JC-YB-826), and Shaanxi Province Key Industrial In novation Chain (Cluster) (No.2020ZDLGY07-05). The authors gratefully acknowledge the anonymous reviewers for their valuable comments.

References

[1] J. An, S. Wu, X. Gui, X. He, and X. Zhang, "A blockchain-based framework for data quality in edge-computing-enabled crowdsensing," *Frontiers of Computer Science*, vol. 17, no. 4, p. 174503, 2023.

[2] Y. C. Chen, W. L. Wang, M. S. Hwang, "RFID authentication protocol for anti-counterfeiting and privacy protection", in *The 9th International Conference on Advanced Communication Technology*, pp. 255-259, 2007.

[3] W. Du, M. Li, L. Wu, Y. Han, T. Zhou, and X. Yang, "A efficient and robust privacy-preserving framework for cross-device federated learning," *Complex & Intelligent Systems*, pp. 1–15, 2023.

[4] R. Ganjavi and A. R. Sharafat, "Edge-assisted public key homomorphic encryption for preserving privacy in mobile crowdsensing," *IEEE Transactions on Services Computing*, 2022.

[5] L. Gong, S. Li, L. Shao, T. Xue, and D. Wang, "Protocols for secure test on relationship on number axis," *Journal of Software*, vol. 31, no. 12, pp. 3950–3967, 2020.

[6] W. P. Hu, C. B. Lin, J. T. Wu, C. Y. Yang, M. S, Hwang, "Research on privacy and security of federated learning in intelligent plant factory systems", *International Journal of Network Security*, vol. 25, no. 2, pp. 377-384, 2023.

[7] P. Kairouz, H. B. McMahan, B. Avent, A. Bellet, M. Bennis, A. N. Bhagoji, K. Bonawitz, Z. Charles, G. Cormode, R. Cummings *et al.*, "Advances and open problems in federated learning," *Foundations and Trends® in Machine Learning*, vol. 14, no. 1–2, pp. 1–210, 2021.

[8] S. Kielienyu, B. Kantarci, D. Turgut, and S. Khan, "Bridging predictive analytics and mobile crowdsensing for future risk maps of communities against covid-19," in *Proceedings of the 18th ACM Symposium on Mobility Management and Wireless Access*, 2020, pp. 37–45.

[9] H. Li, K. Ota, and M. Dong, "Learning iot in edge: Deep learning for the internet of things with edge computing," *IEEE network*, vol. 32, no. 1, pp. 96–101, 2018.

[10] Z. Lian, W. Wang, and C. Su, "Cofel: Communication-efficient and optimized federated learning with local differential privacy," in *ICC 2021-IEEE International Conference on Communications*. IEEE, 2021, pp. 1–6.

[11] C. H. Liu, C. Piao, and J. Tang, "Energy-efficient uav crowdsensing with multiple charging stations by deep learning," in *IEEE INFOCOM 2020-IEEE Conference on Computer Communications*. IEEE, 2020, pp. 199–208.

[12] Y. Liu, Z. Ma, X. Liu, S. Ma, S. Nepal, R. H. Deng, and K. Ren, "Boosting privately: Federated extreme gradient boosting for mobile crowdsensing," in *2020 IEEE 40th international conference on distributed computing systems (ICDCS)*. IEEE, 2020, pp. 1–11.

[13] L. Lyu, H. Yu, and Q. Yang, "Threats to federated learning: A survey," *arXiv preprint arXiv:2003.02133*, 2020.

[14] J. Ma, S.-A. Naas, S. Sigg, and X. Lyu, "Privacy-preserving federated learning based on multi-key homomorphic encryption," *International Journal of Intelligent Systems*, vol. 37, no. 9, pp. 5880–5901, 2022.

[15] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Artificial intelligence and statistics*. PMLR, 2017, pp. 1273–1282.

[16] D. A. Minaam and E. Amer, "Survey on machine learning techniques: Concepts and algorithms," *International Journal of Electronics and Information Engineering*, vol. 10, no. 1, pp. 34–44, 2019.

[17] P. Sun, Z. Wang, L. Wu, Y. Feng, X. Pang, H. Qi, and Z. Wang, "Towards personalized privacy-preserving incentive for truth discovery in mobile crowdsensing systems," *IEEE Transactions on Mobile Computing*, vol. 21, no. 1, pp. 352–365, 2020.

[18] S. Truex, N. Baracaldo, A. Anwar, T. Steinke, H. Ludwig, R. Zhang, and Y. Zhou, "A hybrid approach to privacy-preserving federated learning," in *Proceedings of the 12th ACM workshop on artificial intelligence and security*, 2019, pp. 1–11.

[19] J. Tu, P. Cheng, and L. Chen, "Quality-assured synchronized task assignment in crowdsourcing," *IEEE Transactions on Knowledge and Data Engineering*, vol. 33, no. 3, pp. 1156–1168, 2019.

[20] G. Wang, J. He, X. Shi, J. Pan, and S. Shen, "Analyzing and evaluating efficient privacy-preserving localization for pervasive computing," *IEEE Internet of Things Journal*, vol. 5, no. 4, pp. 2993–3007, 2017.

[21] H. Wu, L. Wang, and G. Xue, "Privacy-aware task allocation and data aggregation in fog-assisted spatial crowdsourcing," *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 1, pp. 589–602, 2019.

[22] X. Yan, B. Zeng, and X. Zhang, "Privacy-preserving and customization-supported data aggregation in mobile crowdsensing," *IEEE Internet of Things Journal*, vol. 9, no. 20, pp. 19 868–19 880, 2022.

[23] L. Yin, J. Feng, H. Xun, Z. Sun, and X. Cheng, "A privacy-preserving federated learning for multiparty data sharing in social iots," *IEEE Transactions on Network Science and Engineering*, vol. 8, no. 3, pp. 2706–2718, 2021.

[24] C. Zhang, S. Li, J. Xia, W. Wang, F. Yan, and Y. Liu, "Batchcrypt: Efficient homomorphic encryption for cross-silo federated learning," in *Proceedings of the 2020 USENIX Annual Technical Conference (USENIX ATC 2020)*, 2020.

[25] J. Zhang and X. Zhang, "Multi-task allocation in mobile crowd sensing with mobility prediction," *IEEE Transactions on Mobile Computing*, 2021.

[26] B. Zhao, X. Liu, and W.-N. Chen, "When crowdsensing meets federated learning: Privacy-preserving mobile crowdsensing system," *arXiv preprint arXiv:2102.10109*, 2021.

[27] B. Zhao, X. Liu, W.-N. Chen, and R. Deng, "Crowdfl: privacy-preserving mobile crowdsensing system via federated learning," *IEEE Transactions on Mobile Computing*, 2022.

Biography

Jian Chen. Jian Chen is a master student of Xi'an Polytechnic University. His main research interests are information security and privacy protection technology.

Linming Gong. Linming Gong is a lecturer in the School of Computer Science, Xi'an Polytechnic University, China, and a mem ber of International Association for Computing Machin ery (ACM). He received the Ph.D. degree from the School of Computer Science, Shaanxi Normal University, Xi'an, China, in 2017. His current research interests include ap plied cryptography, secure multiparty computation, computer and network security, mobile and wireless communication security, and privacy-preserving data mining.

Jingyu Chen. Jingyu Chen is a master student of the School of Computer Science, Xi'an Polytechnic University, China. His main research interests are information security and privacy protection technology.

Tag Ownership Transfer Protocol Based on Parity Check Patching Operation

Zhen-Hui Li
(Corresponding author: Zhen-hui Li)

School of Data Science and Computer Science, Guangdong Peizheng College
Guangzhou 510830, China
Email:2602441@peizheng.edu.cn

(Received Apr. 27, 2023; Revised and Accepted Nov. 23, 2023; First Online Dec. 24, 2023)

Abstract

In the RFID system, the owner of the electronic label may change during its life. In order to ensure the security of the private information stored in the label by other owners when the owner of the electronic label changes, this paper proposes an ultra-lightweight ownership transfer protocol to ensure the security of users' private information. The proposed protocol is based on parity check patching operation to encrypt the plaintext and then carry out data interaction. Parity check patching operation can be realized by bitwise operation. Combined with the unique hamming weight of encryption parameters, different ways of checking and patching can be carried out according to the different size relations. In the case of not adding new parameters and reducing the storage space, it can also increase the difficulty of cracking. One end of the label introduces a parameter of the attribution status bit, which marks the owning status of the current label. Multiple protocols are analyzed based on GNY logical formal analysis and calculation amount, starting from various specific attack types. The proposed protocol has high-security performance and is superior to other comparison protocols in the calculation amount.

Keywords: Electronic Tag; Internet of Things; Parity Check Patching Operation; Transfer Protocol

1 Introduction

Radio frequency identification technology is a technology that can read the data stored in a specific item without contact with it [1, 8]. This technology is widely used in RFID system. A classic RFID system has at least three physical devices, label, reader and server [6, 19, 20, 27], among which RFID technology is widely used due to many advantages, such as low production cost [2,5,18], long service life [16], convenient portability and deployment [24].

In the process of label use, the ownership of the label is often changed. When the ownership of the label is changed, the previous user has no right to access the important data information stored in the label [9, 12]. A classic case is described as follows: User A works and lives in city H, and buys a metrocard to facilitate subway travel. Due to job changes and other factors, user A wants to leave city H, and user A returns the metrocard purchased to the subway company. Some time later, metro sold the card back to user B. When user B gets the metrocard, user B shall have no right to access the information stored in the metrocard by user A, and user A shall have no right to access the data stored in the card by user B [3, 15, 23].

In order to protect the security and privacy data of the above users A and B, as well as each subsequent user, different experts and scholars proposed different types of label ownership transfer protocols to protect the data security of users. Based on the analysis of many other classical protocols, this paper proposes an ultra-lightweight label ownership transfer protocol. This protocol abandons the traditional encryption algorithm to encrypt private data, but adopts the innovative design of ultra-lightweight encryption algorithm, namely parity check patching operation. The implementation principle of this operation is based on the implementation of bitwise operation, which can reduce the calculation amount to the level of ultra-lightweight. At the same time, this operation fully combines the hamming-weight of the encryption parameter itself, compares the size of two hamming-weights, and carries out different modes of odd parity check or even parity check according to the different size relations. After the verification is complete, the patching operation in different ways is carried out again, and the encryption result is finally obtained. This calculation not only reduces the storage space, but also increases the difficulty of cracking without introducing new parameters..

2 Related Research Works

In order to ensure the security of the private information stored in labels by different users, experts and scholars at home and abroad have proposed many ownership transfer protocols.

In literature [21], a lightweight ownership transfer protocol is designed by using the classical hash function for encryption. The encryption algorithm selected by this protocol is relatively classical and can be used in many labels. However, the protocol has the following security defects: some information is sent in plaintext, which can be obtained by third-party eavesdropping. When combined with other obtained messages, the third party can extract the correct values of some private data.

In literature [10], the classical elliptic ECC algorithm is used to encrypt the transmitted data. In view of the mathematical security of ECC algorithm itself, the protocol has super high security attributes. At the same time, because the ECC algorithm requires a lot of computation, it can't be used in low-cost labels with limited computation, so the protocol can only be applied in some special situations.

In literature [17], a label ownership transfer protocol is proposed by using physical unclonable function. Due to the unclonable advantage of the physical unclonable function, the attacker in this protocol can't launch impersonation attacks and has certain security requirements. The in-depth analysis of the protocol shows that some messages sent from one end of the label are not encrypted, and some encrypted messages are not added with random numbers. As a result, the message values of the two rounds of sessions are completely consistent, and the attacker can locate the specific location of the label based on this, resulting in the disclosure of the private location of the label.

In literature [22], an ownership transfer protocol is also designed by using hash function. From the perspective of computation, it is suitable for existing low-cost labels, but the design of the protocol has certain security defects. For example, after receiving a message from the original owner, the label doesn't verify the message immediately. As a result, the attacker can pretend to be the original owner and launch an impersonation attack. Even if other session entities discover the impersonation attack in subsequent communication steps, the protocol has already carried out many steps, resulting in a certain loss of resources.

In literature [13], a protocol based on modular operation is proposed. In terms of the amount of computation, the extension scope of the protocol has been limited because of the large amount of computation. In terms of the security, the protocol can't resist exhaustive attacks. By eavesdropping on multiple complete sessions, an attacker can obtain all message values. Some important privacy data in these messages is sent in plaintext, while only one parameter in some encrypted messages is unknown to the attacker. In this case, the attacker can select a high-performance computer to obtain the correct value of privacy parameter by exhaustive method.

In reference [4], a protocol is proposed using hash functions and the Chinese remainder theorem. This protocol is better than other protocols in terms of security, but it can't be popularized from the perspective of calculation. Because the hash function and the Chinese remainder theorem are far beyond the scope of low-cost label computing capacity. In addition, the protocol still has some security deficiencies. For example, some messages are encrypted without random numbers, so that the attacker can reverse derive the current message value obtained by eavesdropping or analyze the privacy data used in previous sessions. In other words, the protocol can't provide backward-secure. In view of space constraints and other factors, more ownership transfer protocols can be found in literature [11, 14, 25, 26].

Considering that most of the existing classical protocols have some problems, such as large computation, defects in protocol design, or security issues, this paper proposes a ultra-lightweight ownership transfer protocol on the basis of many protocols.

3 Tag Ownership Transfer Protocol Design

This section first gives the meaning of each symbol in the protocol, and then describes the specific steps of the protocol in details.

3.1 Protocol Symbol Definition

O_{new} represents the new owner of the tag;

O_{old} represents the original owner of the tag;

T represents tag;

K_n represents the shared key between O_{new} and T;

K_O represents the shared key between O_{old} and T;

ID represents the tag unique identifier;

ID_L represents the left half of the tag unique identifier;

ID_R represents the right half of the tag unique identifier;

y represents the random number generated by T;

x represents the random number generated by O_{old};

$STATUS$ represents the state position of tag ownership. The $STATUS$ value is 0, representing that the current ownership is owned by O_{old}; the The $STATUS$ value is 1, representing that the current ownership is O_{new} owned.

$\oplus$ represents bitwise XOR operation;

$\&$ represents bitwise and operation;

$Pcpo(x, y)$ represents parity check patching operation [7].

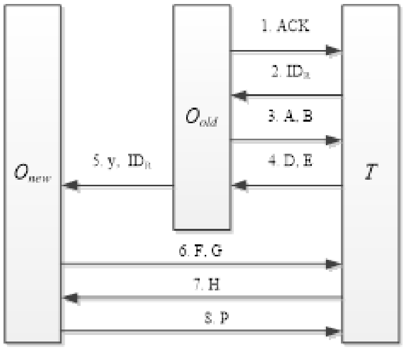


Figure 1: Ownership Transfer Agreement Diagram

3.2 Protocol Step Description

The ownership transfer protocol in this paper involves three communication entities, namely O_{new}, O_{old} and T. Information is exchanged between O_{old} and T through an insecure channel, between O_{new} and T through an insecure channel, and between O_{new} and O_{old} through a secure channel.

The ownership transfer protocol diagram designed in this paper is shown in Figure 1.

In combination with the diagram in Figure 1, the specific steps of the protocol can be described as follows:

Step 1. O_{old} sends ACK to T to initiate the ownership transfer request.

Step 2. T receives the message and T sends ID_R to O_{old} to indicate the response.

Step 3. O_{old} receives information, and O_{old} checks whether there is data in the database equal to the ID_R received.

Not found, indicating that the source is forged, the protocol terminated.

It is found, indicating that T is true and can be followed up. O_{old} generates a random number x, and then calculates messages A and B successively, and finally sends messages A and B to T.

$$A = x \oplus ID_L, B = Pcpo(x\&K_O, ID_L).$$

Step 4. After receiving the information, T first deforms A to obtain $x' = A \oplus ID_L$. Then, x' obtained by deformations is combined with other parameters to calculate $B' = Pcpo(x'\&K_O, ID_L)$ according to the same operation rules, and the sizes of B' and B are compared.

If the two values are different, O_{old} fails to pass the authentication of T, and the protocol is terminated.

If the two values are same, O_{old} passes the authentication of T, and the protocol can continue. Check the value of $STATUS$, when $STATUS = 0$, it indicates that the attribution of T currently belongs to O_{new}, O_{old} has no right to initiate the ownership transfer request, and the protocol is terminated. T generates a random number y, and then calculates messages D and E in turn. Finally, messages D and E are sent to O_{old}.

$$D = y \oplus (x\&ID_L), E = Pcpo(x, y\&K_O).$$

Step 5. O_{old} receives the information, processes the message D to get $y' = D \oplus (x\&ID_L)$, and then calculates $E' = Pcpo(x, y'\&K_O)$ with the same rules, and the sizes of E' and E are compared.

If the two values are different, T fails to pass the authentication of O_{old}, and the protocol is terminated.

If the two values are same, T passes the authentication of O_{old}, and the protocol can continue. O_{old} will send y and ID_R to O_{new} through a secure channel.

Step 6. O_{new} receives the message and saves y and ID_R. O_{new} generates a random number z, and then calculates messages F and G in turn. Finally, messages F and G are sent to T.

$$F = z \oplus (y\&ID_L), G = Pcpo(z, z\&y\&K_n).$$

Step 7. T receives the information, processes the message F to get $z' = F \oplus (y\&ID_L)$, and then calculates $G' = Pcpo(z', z'\&y\&K_n)$ with the same rules, and the sizes of G and G' are compared.

If the two values are different, O_{new} passes the authentication of T, and the protocol can continue. T calculates message H and finally sends message H to O_{new}.

$$H = Pcpo(z\&K_n, y).$$

Step 8. O_{new} receives the information, calculates $H' = Pcpo(z\&K_n, y)$ with the same rules, and the sizes of H and H' are compared.

If the two values are different, T fails to pass the authentication of O_{new}, and the protocol is terminated.

If the two values are same, T passes the authentication of O_{new}, and the protocol can continue. O_{new} calculates message P and finally sends message P to T.

$$P = Pcpo(z \oplus y, y\&ID_L\&K_n).$$

Step 9. T receives the information, calculates $P' = Pcpo(z' \oplus y, y\&ID_L\&K_n)$ with the same rules, and the sizes of P and P' are compared.

If the two values are different, O_{new} fails to pass the authentication of T, and the protocol is terminated.

If the two values are same, O_{new} passes the authentication of T, and the protocol can continue. T modifies the value of $STATUS$, let $STATUS = 1$, indicating that the ownership of T changes, and the ownership is no longer attributed to O_{old}, but to O_{new}. The ownership transfer protocol has now been concluded.

4 Analysis of Formal Protocol Based on GNY Logic

Here, we choose to formalize the protocol with GNY logical formalization as follows:

1) Protocol Formal Description

$MSG1 : O_{old} \rightarrow T : Ack$

$MSG2 : T \rightarrow O_{old} : ID_R$

$MSG3 : O_{old} \rightarrow T : A, B$

$MSG4 : T \rightarrow O_{old} : D, E$

$MSG5 : O_{old} \rightarrow O_{new} : y, ID_R$

$MSG6 : O_{new} \rightarrow T : F, G$

$MSG7 : T \rightarrow O_{new} : H$

$MSG8 : O_{new} \rightarrow T : P$

The above protocol is specified in GNY logical language as follows:

$MSG1 : T < *\{Ack\}$

$MSG2 : O_{old} < *\{ID_R\}$

$MSG3 : T < *\{A, B\}$

$MSG4 : O_{old} < *\{D, E\}$

$MSG5 : O_{new} < *\{y, ID_R\}$

$MSG6 : T < *\{F, G\}$

$MSG7 : O_{new} < *\{H\}$

$MSG8 : T < *\{P\}$

2) Protocol Initialization Assumption

$SUP1 : (K_O, ID, K_n) \in T$

$SUP2 : (ID, K_O) \in O_{old}$

$SUP3 : (ID, K_n) \in O_{new}$

$SUP4 : O_{old}| \equiv \#(x, y, z)$

$SUP5 : T| \equiv \#(x, y, z)$

$SUP6 : O_{new}| \equiv \#(x, y, z)$

$SUP7 : T| \equiv O_{old} \xleftrightarrow{K_O} T$

$SUP8 : O_{old}| \equiv T \xleftrightarrow{K_O} O_{old}$

$SUP9 : T| \equiv O_{old} \xleftrightarrow{ID} T$

$SUP10 : O_{old}| \equiv T \xleftrightarrow{ID} O_{old}$

$SUP11 : T| \equiv O_{new} \xleftrightarrow{K_n} T$

$SUP12 : O_{new}| \equiv \xleftrightarrow{K_n} O_{new}$

$SUP13 : T| \equiv O_{new} \xleftrightarrow{ID} T$

$SUP14 : O_{new}| \equiv T \xleftrightarrow{ID} O_{new}$

3) Agreement Proof Objectives

$GOAL1 : T| \equiv O_{old}| \sim \#\{A, B\}$

$GOAL2 : O_{old}| \equiv T| \sim \#\{D, E\}$

$GOAL3 : T| \equiv O_{new}| \sim \#\{F, G\}$

$GOAL4 : O_{new}| \equiv T| \sim \#\{H\}$

$GOAL5 : T| \equiv O_{new}| \sim \#\{P\}$

4) Agreement Certification Process

Due to the limited length of the text, only the first target of proof $GOAL1 : T| \equiv O_{old}| \sim \#\{A, B\}$ is chosen as an example to prove.

$\because MSG3 : O_{old} \rightarrow T : A, B$ and rules $P1 : \frac{P<X}{X \in P}$

$\because SUP4 : O_{old}| \equiv \#(x, y, z)$ and rules $F1 : \frac{P|\equiv(X)}{P|\equiv(x,y), P|\equiv\#F(X)}$

$\therefore T = \#\{A, B\}$

$\because$ Rules $P2 : \frac{X \in P, Y \in P}{(X,Y) \in P, F(X,Y) \in P}$, $SUP1 : (K_O, ID, K_n) \in T$, $SUP2 : (ID, K_O) \in O_{old}$

$\therefore \{A, B\}\# \in T$

$\because$ Rules $F10 : \frac{P|\equiv(X), X \in P}{P|\equiv\#(H(X))}$ and derived $T = \#\{A, B\}, \{A, B\}\# \in T$

$\therefore T| = \#\{A, B\}$

$\because$ Rules $I3 : \frac{P<H(X,<S>)>, (X,S) \in P, P|\equiv P \leftrightarrow Q, P|\equiv\#(X,S)}{P|\equiv Q|\sim(X,S), P|\equiv Q \sim H(X,<S>)}$

$\because SUP7 : T| \equiv O_{old} \xleftrightarrow{K_O} T, SUP8 : O_{old}| \equiv T \xleftrightarrow{K_O} O_{old}, SUP9 : T| \equiv O_{old} \xleftrightarrow{ID} T$ and $MSG3 : O_{old} \rightarrow T : A, B$

$\therefore T| = \{A, B\}$

$\because$ Freshness definition and derived $T| = \{A, B\}, T| = O_{old} \sim \{A, B\}$

$\therefore GOAL1 : T| \equiv O_{old}| \sim \#\{A, B\}$ freshness definition and derived

5 Comparative Analysis of Protocol Performance

This chapter analyzes the performance of multiple protocols from the perspective of the amount of calculation, storage, random number, and overall traffic of a complete session at one end of the tag. The comparative analysis results are shown in Table 1.

The meanings of the symbols in the above table are described as follows: $PUF()$ represents the amount of calculation of physically unclonable functions; represents the amount of calculation by bit or operation; xor represents the amount of calculation in XOR operation by bit; $Hash()$ represents the amount of calculation of the hash function; and represents the amount of calculation by bit and operation; $MOD()$ represents the amount of calculation in modular operation; $Pcpo()$ represents the amount of calculation in the parity check patching operation; L is the number of bits of length. In the above different types of operations, $xor, or, and, Pcpo()$ belong to ultra-lightweight operations, others belong to lightweight or heavyweight operations.

From the perspective of amount of calculation at one end of the label: According to the description of protocol

Table 1: Performance Comparison Analysis Between Multiple Protocols

Reference	**Calculations at One End of Label**	**Storage Capacity**	**Number of Random Numbers**	**One Full Session Traffic**
Ref [17]	$8PUF() + 2or + 1xor$	$5L$	2	$9L + 2bit$
Ref [22]	$6Hash() + 5and + 7xor$	$2L$	1	$13L + 1bit$
Ref [13]	$4MOD() + 2and + 6xor$	$4L$	3	$10L + 2bit$
Ref [4]	$4MOD() + 2Hash() + 4xor$	$3L$	2	$15L + 2bit$
This Protocol	$5Pcpo() + 2and + 3xor$	$3L + 1bit$	1	$11L + 1bit$

steps in this paper, when we compute x', we first use xor. When we compute B', we first use $Pcpo()$; When we compute D, we first use and, and second use xor; when we compute E, we second use $Pcpo()$; When we compute z' we first use xor, and second use and; When we compute G', we third use $Pcpo()$; When we compute H, we forth use $Pcpo()$; When we compute P', we fifth use $Pcpo()$; Based on the above, the amount of calculation of the protocol in this paper at one end of the label is $5Pcpo()+2and+3xor$. Meanwhile, based on the analysis of Table 1, it can be found that the protocol in this paper is superior to other comparison protocols in terms of the amount of calculation at one end of the label, because only the calculation used by the protocol in this paper is ultra-lightweight, while other protocols are used for lightweight or heavyweight operations.

From the perspective of storage capacity of one end of the label: The protocol in this paper needs to store parameters ID, K_n, and K_O, and also needs to store $STATUS$, but parameter $STATUS$ only needs one byte space, so the storage of one end of the protocol label in this paper is $3L+1bit$. Based on Table 1, it can be found that there is little difference in the storage capacity of each comparison protocol label.

From the perspective of number of random numbers generated at one end of the label: The protocol in this paper only needs to generate a random number, which is somewhat reduced compared with other protocols.

From the perspective of overall traffic of a complete session: The comparison protocols are roughly equivalent in terms of a full session overall traffic.

Based on the comparative analysis of various aspects, the main advantage of the protocol in this paper lies in the amount of calculation at one end of the label, and the total amount of calculation is much less than other comparison protocols. At the same time, in terms of security, it can make up for the security defects of other protocols, so that the protocol in this paper has the value of practical application.

6 Comparative Analysis of Protocol Security

In this chapter, protocol security will be analyzed from several perspectives, such as impersonation attack and exclusivity of ownership.

1) Exclusivity of Ownership

   Exclusivity of ownership means that the original owner no longer has access to the label after the label ownership is changed. The protocol in this paper sets the state variable $STATUS$ of the owner at one end of the label. According to the value of $STATUS$, we can clearly know who owns the current label.

   According to the protocol steps described in this article, to change the value of $STATUS$ on one end of the label, several steps must be passed to authenticate the change. The value of $STATUS$ is changed to 1 at one end of the label if and only if the preceding steps are taken. At this time, the attribution will be changed, and other illegal personnel can't modify the value of variable $STATUS$ through other ways. Therefore, the protocol in the text ensures the exclusivity of ownership.

2) Target Transfer Tag

   Target transfer label refers to the labels to be transferred that need to be determined, and not other labels.

   In order to ensure the correctness of the label, the original owner and the label will verify each other before the ownership transfer. Only when the authentication of each other is passed, it indicates that the label to be transferred is the label we need. Other illegal members can't be authenticated by either of the two session entities without important data.

3) Impersonation Attack

   The session entities involved in the ownership transfer protocol in this paper include the original owner, the new owner and the label. Theoretically speaking, other illegal personnel can impersonate any of them and launch impersonation attacks.

   For example, other illegal people pretend to be labels for conversation. The illegal person can obtain the ID_R of a label by eavesdropping. In later sessions, the illegal person sends the ID_R to the original owner. Although it can be verified by the original owner temporarily, when the illegal person receives the message from the original owner again, the illegal person can't calculate the random number generated by the original owner due to the lack of necessary data. The message obtained by the illegal personnel is sent to the original owner, and the original owner can identify the authenticity by simple calculation. The message sent by the illegal person after impersonating the label failed and didn't obtain any useful private data.

4) Replay Attack

The illegal personnel first quietly eavesdrop on multiple complete communication processes to obtain multiple session messages, and then in a subsequent round of communication, illegal personnel send one or some messages obtained by eavesdropping again in an attempt to obtain privacy data through authentication by the receiver. But for illegal personnel, illegal personnel can't succeed. Firstly, the message is encrypted before transmission mechanism, that is, the message obtained by illegal personnel is ciphertext. Second, the message encryption mixed with random numbers, can maintain the freshness of the message; Third, different random numbers are used in each round, and the message values of the two rounds before and after are also different. When an unauthorized person replays a previous message, the random number used in the current session has changed, and the receiver receives the message. A simple calculation can verify the authenticity of the sender.

5) Localization Attack

The messages sent by labels are encrypted. Data captured by illegal personnel is ciphertext and must be decrypted first. But the illegals lack important data, so that the decryption can't succeed, locating the attack failed, this is one of the reasons. Second, encrypted messages contain random numbers, not only the number of random numbers, and random number producers may be different, so that illegal personnel can't predict; The encryption of random numbers makes the message values calculated by the two rounds of adjacent sessions different and irrelevant. The effect presented to illegal personnel is that the label is in constant change, so illegal personnel can't locate the specific location of the label.

6) Brute-force Attack

The illegal personnel can use the high performance computer to extract the correct value of private data directly from the message obtained by eavesdropping, but the illegal personnel can't succeed. Here, messages A and B are taken as examples.

The illegal personnel can obtain information A and B in various ways. They can process A first and get $x' = A \oplus ID_L$. Here, illegal personnel do not know ID_L, so illegal personnel can only temporarily choose a random number to participate in the calculation. And then illegal personnel combine x' with B to get $B' = Pcpo(x' \& K_O, ID_L)$, that is $B' = Pcpo((A \oplus ID_L) \& K_O, ID_L)$. In the above final deformation formula, it seems that illegal personnel only do not know ID_L and K_O, but in fact, illegal personnel also have other unknown data.

According to the definition of parity check patching operation in this paper, illegal personnel don't know ID_L and the corresponding Hamming-weight value. According to this calculation, there are at least four variables that illegal personnel can't know, and illegal personnel can't exhaust the correct value of private data through exhaustive method. Even if ID_R is transmitted in a civilized way in this paper, illegal personnel can obtain it, but there is no relationship between ID_R and ID_L, that is, ID_R is known, and the correct ID_L value can't be derived at all. Based on the above, illegal personnel brute-force attacking failed.

7) Backward-secure

The illegal personnel tried to analyze the private data used in the previous conversation from the messages obtained by eavesdropping to carry out other illegal activities, but the illegal personnel can't succeed. The message is encrypted and sent to ensure that the data obtained by illegal personnel is ciphertext. In the case of missing important data, the difficulty of cracking the message increases instantaneously. The random number is generated randomly by the random number generator, which guarantees the advantages of randomness, irrelevance and unpredictability between the random numbers before and after. Adding the random number into the information encryption process will make the message values of the previous round and the message values of the subsequent round also have randomness, irrelevance, unpredictability and unreversible, which can ensure the security of the private data in the previous session.

Table 2 shows the results of security comparison between other protocols and the protocol in this paper.

7 Conclusion

In this paper, an ultra-lightweight ownership transfer protocol is proposed. The protocol doesn't use the traditional encryption algorithms, such as hash function, ECC, physical unclonable function, but uses an innovative design of ultra-lightweight encryption algorithm, namely parity check patching operation; The algorithm can be realized by the bitwise operation principle, which can greatly reduce the total operation amount. At the same time, the

Table 2: Security Comparison of Other Protocols with Those in the Article

Attack Type	**Ref [17]**	**Ref [22]**	**Ref [13]**	**Ref [4]**	**This Protocol**
Exclusivity of Ownership	√	√	√	√	√
Target Transfer Tag	√	√	√	√	√
Impersonation Attack	√	×	√	√	√
Replay Attack	√	√	√	√	√
Localization Attack	×	√	√	√	√
Brute-force Attack	√	√	×	√	√
Backward-secure	√	√	√	×	√

Explain: × indicates irresistibility; √ indicates resistance.

operation makes full use of the Hamming-weight information carried in the parameters, carries out the size comparison, and carries out the corresponding odd check or even check. After the check is completed, the patching operation in different ways is carried out, and finally gets the operation result. This algorithm can increase the difficulty for illegal personnel to crack without adding new parameters, which can reduce the storage. Based on the analysis of multiple attack types, the protocol can resist common attack types. GNY logic is used to formalize the protocol and demonstrate its rigor. The calculation of one label end shows that the calculation of the protocol is better than that of other comparison protocols.

Acknowledgments

This paper is supported by the research project of Department of Education of Guangdong Province (2021KTSCX147).

References

[1] Y. C. Chen, W. L. Wang, M. S. Hwang, "RFID authentication protocol for anti-counterfeiting and privacy protection", in *The 9th International Conference on Advanced Communication Technology*, pp. 255-259, 2007.

[2] Y. C. Chen, W. L. Wang, M. S. Hwang, "Low-cost RFID authentication protocol for anti-counterfeiting and privacy protection", Asian Journal of Health and Information Sciences, vol. 1, no. 2, pp. 189-203, 2006.

[3] A. Dewanje and K. A. Kumar, "A new malware detection model using emerging machine learning algorithms," *International Journal of Electronics and Information Engineering*, vol. 13, no. 1, pp. 24–32, 2021.

[4] A. Dewanje and K. A. Kumar, "A new malware detection model using emerging machine learning algorithms," *International Journal of Electronics and Information Engineering*, vol. 13, no. 1, pp. 24–32, 2021.

[5] M. Hosseinzadeh, O. H. Ahmed, S. H. Ahmed, and et al., "An enhanced authentication protocol for RFID systems," *IEEE Access*, vol. 8, pp. 126977–126987, 2020.

[6] M. S. Hwang, C. H. Wei, C. Y. Lee, "Privacy and security requirements for RFID applications", *Journal of Computers*, vol. 20, no. 3, pp. 55-60, Oct. 2009.

[7] Z. H. Li, "RFID mobile authentication protocol based on parity check patching operation," *International Journal of Network Security*, vol. 25, no. 5, 2023.

[8] D. W. Liu and J. Ling, "An improved RFID authentication protocol with backward privacy," *Computer Science*, vol. 43, no. 8, pp. 128–130, 2016.

[9] Q. Ma, X. Li, G. Li, and et al., "Mrliht: mobile RFID-based localization for indoor human tracking," *Sensors*, vol. 20, no. 6, pp. 1–19, 2020.

[10] S. Maheshwari, "Detection of amplitude shift keying signals using current mode scheme," *nternational Journal of Electronics and Information Engineering*, vol. 11, no. 2, pp. 73–80, 2019.

[11] M. M. Nabi and F. Nabi, "Cybersecurity mechanism and user authentication security methods," *International Journal of Electronics and Information Engineering*, vol. 14, no. 1, pp. 1–9, 2022.

[12] H. Rezk, H. El-Bakry, and M. El-Mikkawy, "Adaptive intelligent decision support system for enhancing higher education quality assurance," *International Journal of Electronics and Information Engineering*, vol. 13, no. 4, pp. 180–195, 2021.

[13] Z. Sun and S. Li, "Lightweight authentication protocol for location privacy using PUF in mobile RFID system," *Journal of Frontiers of Computer Science and Technology*, vol. 13, no. 3, pp. 418–428, 2019.

[14] F. Tang and D. Huang, "A BLS signature scheme from multilinear maps," *International Journal of Network Security*, vol. 22, no. 5, pp. 728–735, 2020.

[15] K. Veena and K. Meena, "Identification of cyber criminal by analysing the users profile," *International Journal of Network Security*, vol. 20, no. 4, pp. 738–745, 2018.

[16] J. Q. Wang, Y. F. Zhang, and D. W. Liu, "Provable secure for the ultra-lightweight RFID tag ownership transfer protocol in the context of IoT commerce," *International Journal of Network Security*, vol. 22, no. 1, pp. 12–23, 2020.

[17] L. Wang, E. Li, Y. Ji, and et al., "PUF-based anti-physical cloning RFID security authentication protocol," *Netinfo Security*, vol. 20, no. 8, pp. 89–97, 2020.

[18] C. H. Wei, M. S. Hwang, Augustin Y. H. Chin, "A secure privacy and authentication protocol for passive RFID tags," *International Journal of Mobile Communications*, vol. 15, no. 3, pp. 266–277, 2017.

[19] C. H. Wei, M. S. Hwang, Augustin Y. H. Chin, "An improved authentication protocol for mobile agent device in RFID," *International Journal of Mobile Communications*, vol. 10, no. 5, pp. 508–520, 2012.

[20] C. H. Wei, M. S. Hwang, Augustin Y. H. Chin, "Security analysis of an enhanced mobile agent device for RFID privacy protection," *IETE Technical Review*, vol. 32, no. 3, pp. 183–187, 2015.

[21] Y. Wei and J. Chen, "Tripartite authentication protocol RFID/NFC based on ECC," *International Journal of Network Security*, vol. 22, no. 4, pp. 664–671, 2020.

[22] R. Xie, B. Y. Jian, and D. W. Liu, "An improved ownership transfer for RFID protocol," *International Journal of Network Security*, vol. 20, no. 1, pp. 149–156, 2018.

[23] R. Xie, J. Ling, and D. W. Liu, "A wireless key generation algorithm for RFID system based on bit operation," *International Journal of Network Security*, vol. 20, no. 5, pp. 938–950, 2018.

[24] Y. Yao and J. Su, "An efficient identification algorithm to identify mobile," *Wireless Communications and Mobile Computing*, vol. 1, pp. 1–8, 2021.

[25] S. H. Zhan and C. Q. Yu, "Mobile rfid authentication protocol based on permutation cross synthesis for anti counterfeit attack," *International Journal of Network Security*, vol. 22, no. 4, pp. 305–313, 2022.

[26] F. Zhu, P. Li, H. Xu, and et al., "A lightweight RFID mutual authentication protocol with PUF," *Sensors*, vol. 19, no. 13, pp. 2957–2978, 2019.

[27] C. Zuo, "Defense of computer network viruses based on data mining technology," *International Journal of Network Security*, vol. 20, no. 4, pp. 805–810, 2018.

Biography

Zhen-Hui Li received a master's degree in School of computer science and engineering from SUN YAT-SEN University (China) in 2007. His current research interest fields include information security and artificial intelligence.

Research on Traffic Monitoring for Abnormal Intrusion in Campus Network under OpenFlow Network Structure

Yanlei Zhang

(Corresponding author: Yanlei Zhang)

Academic Affairs Office, Hubei University of Arts and Sciences
No. 296, Longzhong Road, Xiangyang, Hubei 441053, China
Email: zylz74@outlook.com

(Received Oct. 20, 2022; Revised and Accepted Dec. 1, 2023; First Online Dec. 24, 2023)

Abstract

Campus network contains a lot of important information, so its information security protection is very important. This paper briefly introduces the OpenFlow network structure that can be used to construct the campus network. Then, it describes the algorithm, which is composed of a convolutional neural network (CNN) and long short-term memory (LSTM) for distributed denial of service (DDoS) detection in the campus network. Then, simulation experiments were performed, and the designed algorithm was compared with the entropy method and the K-means algorithm. The results showed that the abnormal traffic algorithm based on CNN-LSTM had higher detection accuracy, higher abnormal traffic interception rate, and lower detection time consumption.

Keywords: Campus Network; Distributed Denial of Service; OpenFlow Network; Traffic Detection

1 Introduction

Campus network refers to the computer network covering the campus, which connects all kinds of computer equipment in the campus, thus constructing a huge information exchange and data sharing network system [?]. The campus network can be used for teaching, scientific research, management, and communication purposes, which is essentially a campus local area network, so the computer equipment in the campus network can obtain high-speed and stable bandwidth [4]. As a local area Internet, the campus network provides convenience for campus life at the same time, but the large flow of data will increase the difficulty of data management and cause the problem of information security.

For the campus network, in order to facilitate data management, the OpenFlow network structure is used to realize the control and transmission of data, and the abnormal traffic in the network will be judged by using the flow table entries in the OpenFlow network structure [9]. Jing *et al.* [10] proposed a network intrusion detection method based on correlation deep learning. The simulation results showed that this approach had high average detection and false detection rates for unknown intrusions and attacks. Saied *et al.* [17] proposed an artificial neural network (ANN) algorithm based on a specific feature (pattern) to detect distributed denial of service (DDoS) attacks. The feature separated DDoS attack traffic from real traffic [8, 18].

Sahayet *et al.* [16] proposed an autonomous DDoS defense framework called ArOMA. The experimental results showed that ArOMA could effectively maintain the performance of video streaming at a satisfactory level in the face of DDoS flooding attacks. This article briefly introduces the OpenFlow network architecture which can be used to construct the campus network and then describes the algorithm which is composed of convolutional neural network (CNN) [7, 13, 20] and long short-term memory (LSTM) for the detection of DDoS traffic in the campus network. Simulation experiments were carried out, and the designed algorithm was also compared with the entropy method and the K-means algorithm.

2 Abnormal Traffic Detection of Campus Network Based on OpenFlow Network

2.1 Campus Network and OpenFlow Network Structure

With the rapid development of information technology, the campus network refers to all kinds of computer networks covering the campus. These networks connect all kinds of computer equipment, servers, terminals, and sensors in the campus, forming a huge information exchange and data sharing network system. This network has become an important platform for college teachers and students to learn, exchange, and live. Campus network traffic monitoring can effectively prevent network congestion, malicious attacks, and the spread of viruses and other issues to ensure the stability and security of the campus network. Traditional traffic monitoring is performed independently by each router, which not only complicates the detection of abnormal traffic but also hinders scalability and information transmission in campus networks due to the individual communication protocols of routers. In order to improve the management efficiency of the campus network, the OpenFlow network structure is introduced.

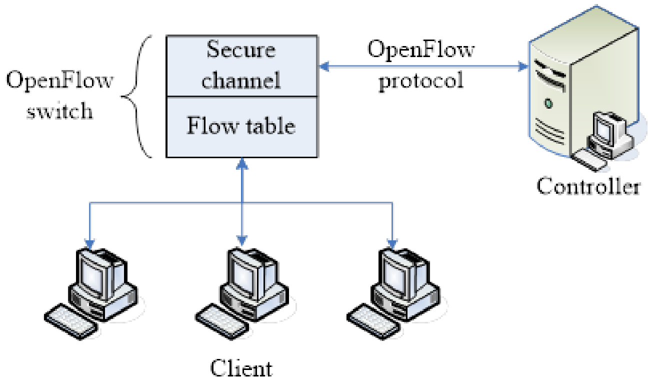


Figure 1: The basic architecture of OpenFlow network

The OpenFlow network is a network architecture based on the OpenFlow protocol, as shown in Figure 1. The structure separates the control plane and the data plane of the network equipment, realizing the flexible control of the network traffic [6]. In the process of data forwarding, the controller centrally manages and controls the flow table in the switch through the OF protocol, and the data packet sent by the client is forwarded by the switch according to the rules set by the flow table. Only when the data packet does not meet the flow table rules [1], the data packet will be sent to the controller. The controller will supplement the flow table rules and send them to the switch.

2.2 Detection of Abnormal Traffic in the Campus Network

In the campus network built by OpenFlow network structure, the switch is responsible for the data transmission between each computing device, and the controller manages the flow table of the switch [14]. However, malicious attacks in the Internet will not be reduced because of the change of network structure. DDoS attack is a common intrusion attack on OpenFlow network structure. This attack method generates a large number of Packet_In requests and flow entries through a large number of fake internet protocol (IP) data, consumes the computing resources of controllers and switches, and realizes the purpose of paralysis of the network. Therefore, it is necessary to prevent this kind of attack [19].

Because the goal of DDoS attack is to paralyze the controller with a large number of useless requests, even if the source IP of DDoS attack is relatively scattered, its target IP is often relatively concentrated. Therefore, the number of target IPs can be counted, and then the information entropy can be calculated. The more concentrated the target IP is, the smaller its information entropy is, and the more likely it is DDoS. The information entropy [3] is calculated:

$$\begin{cases} H_\alpha(x) = \frac{\log_2(\sum_{i=1}^{n} p_i^\alpha)}{1-\alpha} \\ p_i = \frac{n_i}{S} \\ S = \sum_{i=1}^{N} n_i, \end{cases} \tag{1}$$

where $H_\alpha(x)$ represents the information entropy of the total packet, p_i is the occurrence probability of the target IP, α is the order of information entropy, n_i is the number of occurrences of the target IP, N is the number of different target IPs, and S is the total number of occurrences of different target IPs. A threshold is set. When the information entropy of the data flow is lower than the preset threshold, the counter of the abnormal queue increments by 1, and the information entropy of the next data flow is recorded. When the information entropy of the data flow is higher than the preset threshold, the abnormal queue counter is reset [2]. When the abnormal queue counter count exceeds a certain value, it means that abnormal data flow has been detected continuously, and it is judged to be under DDoS attack.

The method of using information entropy to detect data flow is relatively simple, but in actual situation, the strong directivity of data flow may not indicate a DDoS attack. For example, when there is hot information on the network, all kinds of data flow will be relatively concentrated in the hot IP. Or in the promotion period, shopping websites will also receive concentrated data flow [15]. The simple use of information entropy will cause misjudgment. Although the abnormal queue counter is introduced for secondary screening, the threshold set is fixed, which is not flexible enough in the face of changing networks. Therefore, this paper chooses to use neural networks to detect DDoS attacks. The monitoring process in OpenFlow network is shown in Figure 2. In this paper, a CNN algorithm and a LSTM algorithm are combined. The CNN algorithm is used to extract local features, while LSTM effectively utilizes historical data. The specific steps are shown below.

1) The controller in the OpenFlow network and the OF protocol are used to collect flow table entry information from switches at a periodic interval of T.

2) The time feature [21] is the gradient of the hit rate of flow table entries, which is calculated by:

$$\begin{cases} e_t = \frac{\sum_{f \in F_p} FC_f}{N_t \cdot \sum_{f \in F_p} D_f} \\ \frac{e_{t+T}}{e_t} = \Delta e \cdot \frac{N_{t+T}}{N_t} \end{cases} \tag{2}$$

where e_t and e_{t+T} are the average hit rate of the flow table at destination port P at time t and $t+T$, F_p is

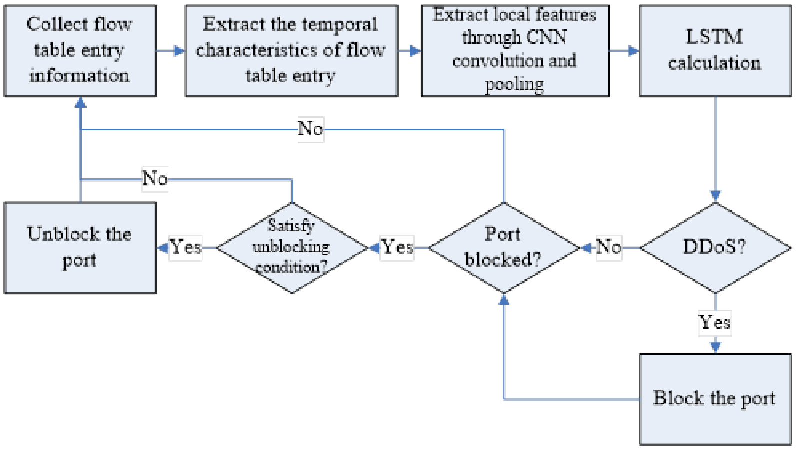

Collect flow table entry information
Extract the temporal characteristics of flow table entry
Extract local features through CNN convolution and pooling
LSTM calculation
DDoS?
Yes
Block the port
No
Port blocked?
Yes
No
Satisfy unblocking condition?
Yes
No
Unblock the port

Figure 2: DDoS monitoring process based on CNN-LSTM in OpenFlow network

the set of flow table entries of P at time t, f is a flow table entry in F_p, N_t stands for the number of flow table entries, FC_f is the number of hit packets of f, D_f is the existence time of f. Δe is the change gradient of the hit rate of the flow table entry in period T. Then, change gradients at $T' = n \cdot T(n = 2, 3, 4, \cdots)$ are arranged in chronological order to obtain change gradient vector $\overrightarrow{e} = (\Delta e_1, \Delta e_2, \Delta e_3, \cdots, \Delta e_n)$ of the flow table entry hit rate of destination port P in the period composed of continuous complex period T.

3) The input is put into the CNN algorithm, and the convolution kernel in the convolution layer is used to further extract the local features of the traffic time feature. The convolution calculation formula of the convolution kernel is:

$$x_j^l = f(\sum_{j \in M} x_i^{l-1} \cdot W_{ij}^l + b_j^l), \quad (3)$$

where x_j^l is the feature map obtained by the convolution kernel, x_i^{l-1} is the feature output after the last convolution and pooling, W_{ij} is the weight parameter, b_j^l is a bias, M is the number of convolution kernels, and $f(\cdot)$ is the activation function. After obtaining the convolutional features of the traffic, in order to reduce the amount of computation, the convolutional features are compressed in the pooling layer, a pooling box is used to slide on the convolutional features, and the average or maximum value of the numbers within the box are taken during the sliding process.

4) The convolutional features extracted by the CNN algorithm are input into the LSTM algorithm for forward calculation. The LSTM algorithm introduces a gate structure of forgetting mechanism to avoid the disappearance of gradient caused by long sequence data. The calculation formulas of the LSTM algorithm are:

$$\begin{cases} f_t = \sigma(\omega_f \cdot [h_{t-1}, x_t] + b_f) \\ i_t = \sigma(\omega_i [h_{t-1}, x_t] + b_i) \\ \widetilde{C_t} = \tan h(\omega_C [h_{t-1}, x_t] + b_C) \\ C_t = f_t \times C_{t-1} + i \times \widetilde{C_t} \\ o_t = \sigma(\omega_o [h_{t-1}, x_t] + b_o) \\ h_t = o_t \times \tan h(C_t) \end{cases} \quad (4)$$

where $\widetilde{C_t}$ and C_t are the temporary state and the updated state of the "cell" at the current time respectively, h_t is the hidden state of sequence data at the current time, x_t is the input at the current time, f_t, i_t, and o_t are the outputs of the forgetting, input, and output gating units at the current time respectively, ω_f, ω_i, and ω_o are weights in the corresponding gating unit, b_f, b_i, and b_o are biases in the corresponding gating unit.

5) If the LSTM calculation result is DDoS attack, then destination port P is blocked, and enter the next step; if not, it will go directly to the next step.

6) Whether there is a blocked port in the network is determined [12]. If there is, enter the next step; if no, return to Step 1.

7) Whether the blocked port has reached the unblocking condition is determined. If it has reached, it returns to Step 1; if not, it directly returns to Step 1. The unblocking condition is that the average hit rate of flow table entries for the blocked port is less than the preset threshold.

3 Simulation Experiment

3.1 Experimental Setup

A small OpenFlow network was built in the laboratory by using multiple servers, as shown in Figure 3. The

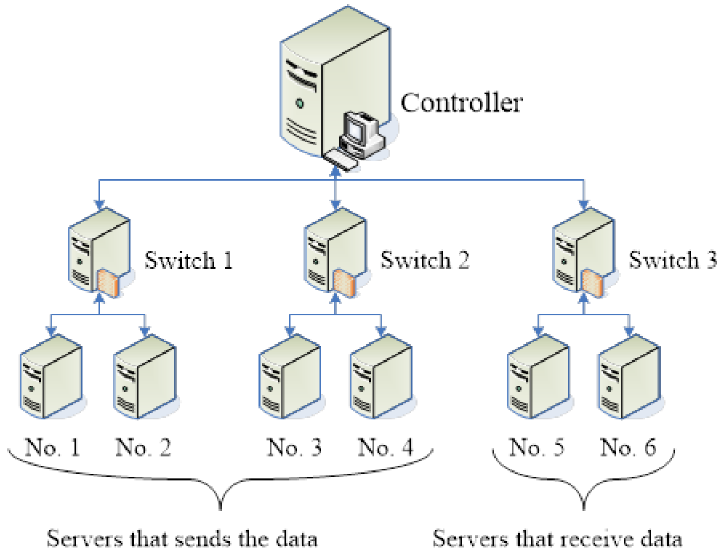


Figure 3: The basic structure of OpenFlow network in the simulation experiment

controller and the OF switch were connected by the OF protocol. Switch 3 was connected with two servers (5 and 6) that accept data packets, and the remaining two switches were connected with two servers that send data packets. The three switches were also connected.

The detection algorithm for DDoS attacks was set in the server acting as the controller. The relevant parameters of the detection algorithm based on CNN-LSTM were set as follows. The CNN part consisted of two convolution layers with 32 convolution kernels in a size of 2 × 2, one mean pooling layer containing a 3 × 3 pooling box, two convolutional layers with 16 convolution kernels in a size of 2 × 2, and one mean pooling layer with a 3 × 3 pooling box. The sigmoid activation function was used in the convolutional layer. The LSTM part had two hidden layers, and the number of nodes in each layer was 1,024. The sigmoid function was adopted, cross entropy was used during training, and the learning rate was set to 0.1.

The relevant parameters of the algorithm based on information entropy are as follows. The order of information entropy was set to 5 after orthogonal experiments, the attack threshold was set to 3.55, and the threshold of abnormal queue counter was set to 5. In the algorithm based on K-means clustering, the number of classifications was set as 2: one for low flow data and the other for non-low flow data. When the low flow data exceeded the preset threshold, the flow was determined as abnormal, and the threshold was set as 30.

3.2 Experimental Methods

During the test, server 1 initiated data transmission to the whole network, and the data packets transmitted included 3.5 GB of normal network data and 2.4 GB of DDoS attack data. In the DDoS attack monitoring process mentioned above, the processing method of abnormal traffic is port blocking. However, in this experiment, in order to observe the interception effect of the detection algorithm on DDoS more intuitively, the processing method was changed to: normal network traffic was sent to server 5, and abnormal network traffic was sent to server 6.

3.3 Experimental Results

The performance of three DDoS attack detection algorithms is shown in Figure 4. It can be seen that the DDoS detection algorithm based on CNN-LSTM was the highest, the algorithm based on K-means was the second, and the algorithm based on entropy was the lowest in terms of precision, recall rate, and F value.

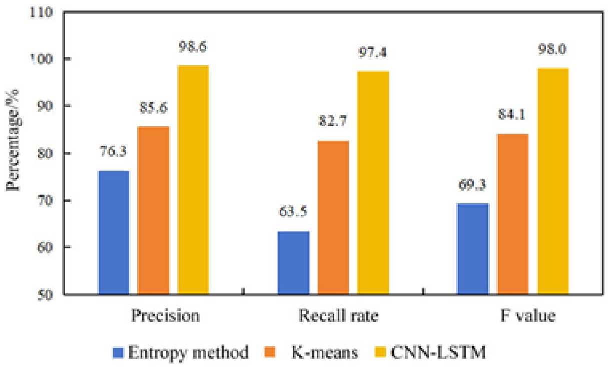


Figure 4: Recognition performance of three DDoS detection algorithms

The interception performance of three DDoS detection algorithms for abnormal traffic is presented in Table 1. It can be observed that the interception rate of the entropy method was 84.58%, the interception rate of the K-

Table 1: Interception performance of three DDoS detection algorithms against abnormal traffic

Detection algorithm	Traffic in server 5/GB	Traffic in server 6/GB	Anomalous traffic interception rate/%
Entropy method	3.87	2.03	84.58
K-means	3.65	2.25	93.75
CNN-LSTM	3.51	2.39	99.58

means method was 93.75%, and the interception rate of the CNN-LSTM algorithm was 99.58%. It was concluded that that the DDoS detection algorithm based on CNN-LSTM had the highest abnormal traffic interception rate, the algorithm based on K-means was the second, and the algorithm based on entropy was the lowest.

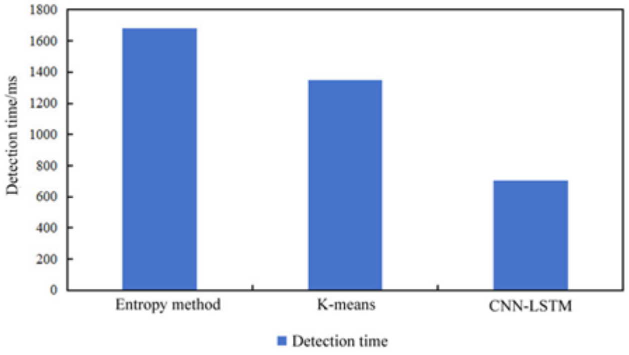


Figure 5: Detection time consumption of three DDoS detection algorithms

It can be intuitively seen from Figure 5 that the entropy-based DDoS detection algorithm took the longest time, followed by the k-means-based algorithm, while the CNN-LSTM-based algorithm took the least time. From the above experimental results, it can be seen that the DDoS detection algorithm based on CNN-LSTM was not only better in the detection accuracy of DDoS attacks, but also had the shortest detection time. The reason is that when using entropy to identify DDoS, it cannot effectively distinguish whether the large-traffic packets are caused by malicious attacks or hot spots. The K-means algorithm classifies the data packets from the perspective of the data characteristics, which improves the recognition accuracy to a certain extent, but there are still errors in the face of changing data types. In terms of detection time, both the entropy method and K-means method require iterative calculations on multiple data packets. However, the CNN-LSTM neural network offers advantages in terms of parallel processing capabilities for data analysis. Additionally, once trained, it only requires layer-by-layer computations on input data without any repetitive loops, resulting in a significant reduction in detection time.

4 Conclusion

This paper briefly introduces the OpenFlow network structure that can be used to build the campus network, and then describes the algorithm which is composed of CNN and LSTM for DDoS abnormal traffic detection in the campus network. Simulation experiment were conducted, and the CNN-LSTM algorithm was compared with the entropy method and the K-means algorithm. The following results were obtained. (1) In terms of precision, recall rate, and F value, the DDoS detection algorithm based on CNN-LSTM was the highest, the algorithm based on K-means was the second, and the algorithm based on entropy was the lowest. (2) The DDoS detection algorithm based on CNN-LSTM had the highest abnormal traffic interception rate, followed by the algorithm based on K-means, and the algorithm based on entropy had the lowest rate. (3) The entropy-based DDoS detection algorithm took the most time, the k-means-based algorithm took the second-longest time, and the CNN-LSTM-based algorithm took the least time.

References

[1] M. De Assis, A. H. Hamamoto, T. Abrão, M. L. Proença, "A Game Theoretical Based System Using Holt-Winters and Genetic Algorithm With Fuzzy Logic for DoS/DDoS Mitigation on SDN Networks," *IEEE Access*, vol. 5, pp. 9485-9496, 2017.

[2] L. Duan, F. Yu, L. Zhan, "An improved fuzzy C-means clustering algorithm," in *International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery*, pp. 44-46, 2016.

[3] D. Gugelmann, F. Gasser, B. Ager, V. Lenders, "Hviz: HTTP(S) traffic aggregation and visualization for network forensics," *Digital Investigation*, vol. 12, pp. S1-S11, 2015.

[4] K. Hong, Y. Kim, H. Choi, J. Park, "SDN-Assisted Slow HTTP DDoS Attack Defense Method," *IEEE Communications Letters*, vol. 22, no. 4, pp. 688-691, 2017.

[5] Y. Y. Hsieh, L. H. Chang, A. Y. H. Liao, C. Y. Yang, M. S. Hwang, "The System Adoption Evaluation of RFID Safety Management System on Campus", in *International Journal of Network Security*, pp. 176-180, 2022.

[6] N. Hoque, D. K. Bhattacharyya, J. K. Kalita, "Botnet in DDoS Attacks: Trends and Challenges," *IEEE*

Communications Surveys & Tutorials, vol. 17, no. 4, pp. 2242-2270, 2015.

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

[8] M. S. Hwang, S. K. Chong, , H. H. Ou, "The moderately hard DoS-resistant authentication protocol on client puzzles", *Informatica*, vol. 27, no. 1, pp. 31-48, 2016.

[9] S. Jantila, K. Chaipah, "A Security Analysis of a Hybrid Mechanism to Defend DDoS Attacks in SDN," *Procedia Computer Science*, vol. 86, pp. 437-440, 2016.

[10] L. Jing, W. Bin, "Network Intrusion Detection Method Based on Relevance Deep Learning," in *International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS'16)*, pp. 237-240, 2016.

[11] K. Kalkan, G. Gur, F. Alagoz, "Defense Mechanisms against DDoS Attacks in SDN Environment," *IEEE Communications Magazine*, vol. 55, no. 9, pp. 175-179, 2017.

[12] N. Khamphakdee, N. Benjamas, S. Saiyod, "Improving Intrusion Detection System Based on Snort Rules for Network Probe Attacks Detection with Association Rules Technique of Data Mining," *Journal of ICT Research & Applications*, vol. 8, no. 3, pp. 234-250, 2015.

[13] I. C. Lin, H. H. Ou, M. S. Hwang, "A user authentication system using back-propagation network", *Neural Computing & Applications*, vol. 14, pp. 243-249, 2005.

[14] H. Mahmood, D. Mahmood, Q. Shaheen, R. Akhtar, C. Wang, "S-DPS: An SDN-Based DDoS Protection System for Smart Grids," *Security and Communication Networks*, vol. 2021, pp. 1-19, 2021.

[15] P. Nayak, A. Devulapalli, "A Fuzzy Logic-Based Clustering Algorithm for WSN to Extend the Network Lifetime," *IEEE Sensors Journal*, vol. 16, no. 1, pp. 137-144, 2015.

[16] R. Sahay, G. Blanc, Z. Zhang, H. Debar, "ArOMA: an SDN based autonomic DDoS mitigation framework," *Computers & Security*, vol. 70, pp. 482-499, 2017.

[17] A. Saied, R. E. Overill, T. Radzik, "Detection of known and unknown DDoS attacks using Artificial Neural Networks," *Neurocomputing*, vol. 172, no. C, pp. 385-393, 2016.

[18] J. R. Sun, M. S. Hwang, "A new investigation approach for tracing source IP in DDoS attack from proxy server", in *Intelligent Systems and Applications*, pp. 850-857, 2015.

[19] R. Ujjan, Z. Pervez, K Dahal, W. A. Khan, A. M. Khattak, B. Hayat, "Entropy Based Features Distribution for Anti-DDoS Model in SDN," *Sustainability*, vol. 13, no. 3, pp. 1-27, 2021.

[20] H. J. Wu, Y. H. Chang, M. S. Hwang, I. C. Lin, "Flexible RFID location system based on artificial neural networks for medical care facilities", *ACM SIGBED Review*, vol. 6, no. 2, pp. 1-8, 2009.

[21] T. Yoshioka, S. Karita, T. Nakatani, "Far-field speech recognition using CNN-DNN-HMM with convolution in time," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 4360-4364, 2015.

Biography

Yanlei Zhang, born in March 1974, has received the master's degree from Wuhan University of Technology in 2007. He is an experimentalist (engineer) and works at Hubei University of Arts and Sciences. He is interested in computer aided education and practical teaching management.

Lightweight RFID Bidirectional Authentication Protocol Based on Improved Hash Function

Fang-Ming Cao and Xiao-Ping He
(Corresponding author: Fang-Ming Cao)

School of Data and Computer Science, Guangdong Peizheng College
Guangzhou 510800, China
Email:caofangming1983@163.com

(Received Feb. 28, 2023; Revised and Accepted Oct. 26, 2023; First Online Dec. 24, 2023)

Abstract

This paper analyzes He *et al.*'s protocol, points out that its protocol cannot resist replay and asynchronous attacks, and proposes a lightweight RFID bidirectional authentication protocol; the proposed security protocol considers that the computing capacity of the tag end is limited, abandons the elliptic ECC encryption algorithm, and adopts a self-designed deformation Hash function algorithm to encrypt necessary digital Combined with the Hamming weight comparison of its parameters, different methods of reverse or exchange operations are carried out. Combined with the Hamming weight comparison of its parameters, different methods of reverse or exchange operations are carried out to reduce the introduction of parameters and also increase the difficulty of cracking security; performance and formalization analysis of the proposed protocol shows that the protocol is superior to other comparison protocols in terms of The security, performance, and formalization analysis of the proposed protocol shows that the protocol is superior to other comparison protocols in terms of security. The total calculation of performance is less than other protocols.

Keywords: Deformation Hash Function; Lightweight Protocol; Radio Frequency Identification (RFID) Technology

1 Introduction

Radio Frequency Identification (RFID) technology is a kind of Internet of things technology that can read out the data stored in a specific product without contacting the specific product [2,14,15]. This technology was produced in the last century, but because of the last century's science and technology development limitations and other factors, which make radio frequency identification technology did not get wide development application [?]. After entering the new century, along with cloud computing [4], the Internet of things [16], big data [17] and other new technology generations, as well as the rapid development of science and technology, radio frequency identification technology again into the people's vision, and get rapid development.

Radio frequency identification technology is most widely used in radio frequency identification systems, one of the most classic radio frequency identification systems contains at least the following entities: tag, server, and reader [3,9,19,20]. The tag is used to store the user or the commodity's important privacy data; the reader is used to read/write the data stored inside the corresponding tag, and will send the data to the server; the server is used to verify the authenticity of the tag or the read-writer. In the classical radio frequency identification system, reader and server between commonly used in fiber optic and other limited link communication, which has high safety, generally, the two can be regarded as a whole; the reader and tag need to use the wireless link interaction, because the wireless link has the openness, making communication data very insecure [10,23].

In order to guarantee the security of the data interaction between the reader and the tag, the authentication protocol needs to be designed to solve the security problem. There are a variety of encryption algorithms in the commonly used protocol, but combined with the actual cost of the tag and other factors, making many encryption algorithms cannot be used, the reason is: the tag is generally divided into two [18], one is the active tag, the other is the passive tag; the former carries power, without the help of external force, itself can send information, while the latter because of the lack of power, itself cannot output information, must rely on the external force before sending data; the tag itself receives strict cost constraints, which will lead to limited computing power and limited storage space, and heavyweight encryption algorithms such as elliptic curve ECC [22] cannot be used in low-cost tags. Therefore, there is a need to design lightweight authentication protocols, which can ensure the security of important data and enable the protocols to be used in low-cost tags.

2 Related Works

A lightweight protocol based on hash functions is proposed in the paper [7], which uses too many hash function operations at the tag end, leading to increased cost overhead; at the same time, an attacker can consume all the pre-stored pseudo-identifier PID data at the tag end by replaying the identifier TID, thus making the protocol unable to perform the next round of authentication.

A lightweight protocol is designed in the paper [5] using a quadratic residual theorem and pseudo-random number generator, in terms of computational power, the proposed protocol can satisfy the use of low-cost tags, but since the tag end only uses a simple bit-loop operation and pseudo-random number generator, it makes the protocol unable to resist physical cloning attacks.

In the paper [11], an RFID protocol is proposed based on the cloud, and the protocol is designed in such a way that the reader is only used as an intermediate forwarder that the data read from the tag end is directly forwarded to the cloud server, while the data read from the cloud server is directly forwarded to the tag, which makes it possible for an attacker to impersonate as a reader and thus launch a counterfeit attack and tamper with the message content.

The protocol designed in the paper [6] can only be used in specific scenarios, i.e., mutual authentication between multiple servers and the same tag, and the use of this protocol is limited in application and cannot be used on a large scale. At the same time, a tag needs to achieve authentication with multiple servers, the tag must store more data information, which increases the pressure on the tag storage space.

The protocol proposed in the paper [12] cannot resist denial of service attacks because the protocol requires a large number of traversal search IDs at the server end, and when the server is performing this operation, the attacker can send a large number of queries to the server, resulting in the paralysis and eventual collapse of the server; at the same time, the attacker can block the messages M4 and M5 in a way that is the loss of consistency in the information shared between the tag and the server, i.e., the protocol cannot resist asynchronous attacks.

An authentication protocol is designed in the paper [13] using a dual physical unclonable function PUF, which can resist physical unclonable attacks while reducing resource consumption to some extent, but PUF deployed on cloud services may lead to leakage of private data and have some security concerns.

In the paper [8], a protocol is designed based on elliptic curve ECC and hash function, and an in-depth analysis of the protocol reveals the following inappropriate or defective points: first, the protocol needs to be able to implement both elliptic curve ECC and hash function encryption algorithm at the tag end, which is not applicable to the tag with strictly limited low-cost computation; second, in the last step of the protocol, the tag only sends accept message to the server, and the server does not verify the tag, which allows the attacker to truncate the normal communication, and the attacker continues to send accept message to the server, which makes the server continuously update the ID and S at one end, and eventually leads to the loss of consistency of the information shared between the server and the tag, i.e., the protocol cannot resist replay attacks, impersonation attacks, and asynchronous attacks; third, when the protocol is compared in terms of security and performance, the protocols selected for comparison are all relatively old protocols, which cannot reflect the advantages of the protocol in comparison.

Based on the analysis and comparison of many classical protocols, many protocols have a large computational capacity and cannot be used in low-cost tags; there are security flaws in the design of the protocol itself; the protocol design takes into account the applicable scenarios, making the protocol limited in the application, etc. The paper integrates the advantages of many protocol frameworks and proposes an improved lightweight authentication protocol. In order to make the protocol can really be used in low-cost tags, the protocol chooses the hash function as the encryption algorithm which not directly choosing the traditional hash function to encrypt directly, but deforms in a certain way. The protocol skillfully utilizes the Hamming weight of the encryption parameters without increasing storage space and computation. According to the comparison of Hamming weight, deformation in different ways, or take the reverse operation, or exchange operation, which can increase the difficulty for the attacker to crack.

3 Deformation Hash Function Design

The deformation Hash function convention is represented by the symbol $H(X,Y)$ and is implemented as defined below:

1) (X,Y,Z) is a binary string of length L bits, $w(X), w(Y)$ is the Hamming weight of X, Y respectively, and d is the absolute value of the difference of $w(X), w(Y)$.

2) When $w(X) \leq w(Y)$, where $Z = X \oplus Y$, according to the above can be learned from $d = w(Y) - w(X)$. The first d bits and the last d bits of Z will be reversed, and when the reverse operation is completed, the hash operation will be performed to output the final result.

3) When $w(X) > w(Y)$, according to the above, we can know $d = w(X) - w(Y)$. The first d bits and the last d bits of X will be reversed, and the first d bits and the last d bits of Y will be exchanged, and when the reverse and exchange operations are completed, then the XOR operation will be performed to obtain Z,

and finally the hash operation will be performed on Z to output the final result.

In order to clarify the implementation process of this encryption algorithm, the following two examples will be combined to illustrate it.

Example 1, take $L = 12, X = 010010010010, Y = 101110010110$, can be known $w(X) = 4, w(Y) = 7$, to meet the case of $w(X) \leq w(Y)$, according to definition, we further get $Z = X \oplus Y = 111100000100, d = w(Y) - w(X), Z' = 000100000011$, and finally calculate $H(Z')$, and output the results.

Example 2, take $L = 12, X = 110101110011, Y = 000110011000$, can be known $w(X) = 8, w(Y) = 4$, to meet the case of $w(X) > w(Y)$, according to the above definition to get $X' = 001001111100, Y' = 100010010001, d = w(X) - w(Y) = 4$, according to definition, we further get $Z' = X' \oplus Y' = 101011101101$, and finally calculated $H(Z')$, and output the results.

4 Lightweight RFID Protocol Definition

Given the many problems with the protocol designed by He *et al.*, this paper proposes an improved security protocol based on their protocol framework. The server and the reader inter-link security, can be viewed as a whole; the link between reader and the tag is open, there are hidden problems, need to use the security protocol for information interaction.

1) Lightweight RFID protocol symbols and meanings. There are many symbols for security protocols in the text, and the meaning of each symbol will be given below.

   DB/R denotes that the server and the read-writer consist of the whole;

   T denotes tag;

   K denotes that the key is shared between DB/R and T;

   K^{new} denotes the current shared key between DB/R and T;

   K^{old} denotes the last round of shared keys between DB/R and T;

   IDS denotes pseudonym for T;

   IDS^{new} denotes the current pseudonym of T;

   IDS^{old} denotes the upper round pseudonym of T;

   TID denotes the identifier of T;

   a denotes the random number generated by DB/R;

   b denotes the random number generated by T;

   $H(X), H(Y)$ denotes deformation Hash function operations;

   $\oplus$ enotes a XOR operation;

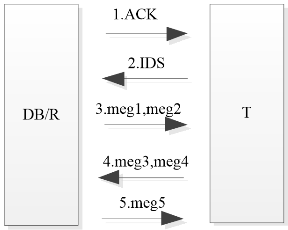


Figure 1: Schematic Diagram of Lightweight RFID Security Protocol

   $\&$ denotes and operation;

   $count$ denotes the statistician at the T end;

   $megi$ denotes message;

   ACK denotes request acknowledgment message.

2) Lightweight RFID protocol implementation. A schematic of the lightweight RFID security protocol is shown in Figure 1.

   The specific steps of the security protocol in the text are described according to Figure 1 as follows:

   Step 1. When T is close to DB/R and within the readable range of DB/R, DB/R sends a ACK message to T to enable the security authentication protocol.

   Step 2. T receives the ACK message and sends IDS to DB/R in response.

   Step 3. DB/R receives a message from IDS, DB/R looks for the existence of data in the database equivalent to IDS.

   If the data does not exist, T is counterfeit, the security protocol is stopped.

   If the data exists, T is verified and DB/R fetches the other stored data information corresponding to IDS. DB/R generates the random number a, and the messages $meg1$ and $meg2$ are calculated according to the agreed operation rules, and the messages $meg1$ and $meg2$ are sent to T.

   $$\text{meg1} = a \oplus TID, meg2 = H(a\&IDS, TID).$$

   Step 4. When T receives $meg1$ and $meg2$ messages, firstly deform $meg1$ to get $a' = meg1 \oplus TID$, then combine d, IDS and TID to get $meg2'$ according to the same algorithm, and compare $meg2$ and $meg2'$ in terms of size.

   The two are different, DB/R is counterfeit, and the security protocol is stopped.

   Both are the same, DB/R passes the verification. T generate the random number b, and the messages $meg3$ and $meg4$ are calculated in turn, and finally the messages $meg3$ and $meg4$ are sent to DB/R.

$$\begin{aligned} a' &= meg1 \oplus TID, \\ meg2' &= H(a'\&IDS, TID) \\ &= H((meg1 \oplus TID)\&IDS, TID), \\ meg3 &= b \oplus (a\&TID), \\ meg4 &= H(a\&K, b). \end{aligned}$$

Step 5. When DB/R receives $meg3$ and $meg4$ messages, firstly deform $meg3$ to get $b' = meg3 \oplus (a\&TID)$, then combine b', a and K to get $meg4$ and $meg4'$ in terms of size.

The two are different, T is counterfeit, and the security protocol is stopped.

Both are the same, T passes the verification. DB/R calculate to get the message $meg5$ and then the update of the message is started ,finally sending the message $meg5$ to T.

When DB/R verifies the success of T with $K = K^{new}$ and $IDS = IDS^{new}$, DB/R updates the information as follows:

$$\begin{aligned} K^{old} &= K^{new}, \\ K^{new} &= H(a, a\&b\&K^{new}), \\ IDS^{old} &= IDS^{new}, \\ IDS^{new} &= H(b\&IDS^{new}, a). \end{aligned}$$

When DB/R verifies the success of T with $K = K^{old}$ and $IDS = IDS^{old}$, DB/R, DB/R updates the information as follows:

$$\begin{aligned} K^{old} &= K^{old}, \\ K^{new} &= H(a, a\&b\&K^{old}), \\ IDS^{old} &= IDS^{old}, \\ IDS^{new} &= H(b\&IDS^{old}, a). \\ b' &= meg3 \oplus (a\&TID), \\ meg4' &= H(a\&K, b') \\ &= H(a\&K, meg3 \oplus (a\&TID)), \\ meg5 &= H(a\&b \oplus K, b\&TID). \end{aligned}$$

Step 6. When T receive the message of $meg5$, first check if the statistician *count* value corresponding to the $meg5$ value for the current round is 0?

If it is not 0, the tag has been received before, so no operation is performed after this time to resist replay attacks.

If it is 0,the tag has not received the message before, and set the count value of the statistician corresponding to the meg5 value to 1. The combination of a', K, b, and TID is calculated according to the same algorithm to obtain $meg5'$ and compare the size relationship between $meg5$ and $meg5'$.

The two are different, DB/R is counterfeit, and the security protocol is stopped.

Both are the same, DB/R passed the verification. T Start updating the information $K = H(a, a\&b\&K)$ and $IDS = H(b\&IDS, a)$, when the information update is completed, the security protocol ends normally.

$$\begin{aligned} a' &= meg1 \oplus TID, \\ meg5' &= H(a'\&b \oplus K, b\&TID) \\ &= H((meg1 \oplus TID)\&b \oplus K, b\&TID). \end{aligned}$$

5 Lightweight RFID Protocol Security Analysis

This section analyzes the security of the protocols in the text from multiple attack perspectives, including replay, impersonation, and asynchrony.

Replay attack. The He *et al.* protocol simply sends the *Accept* message in the last step without completing the inter-authentication, which allows the attacker to send the intercepted message repeatedly, thus making the protocol not resistant to replay attacks.

The protocol in the text introduces a statistic *count* at the T end, which will be used to determine whether the currently received message has been received before based on whether the value of the statistic *count* is 0. The tag will follow up when and only when the value of *count* is 0. If the value of *count* is not 0, it means that the message has been received at least once before and the tag will discard the message and stop the protocol to prevent possible replay attacks. Following this approach, the protocol in the text is resistant to replay attacks.

Rearward safety. The tag sends out the message all adds the random number b which generates by itself , the server and the reader composition whole adds the random number a which generates by itself, can guarantee each message freshness by the above way. When the attacker obtains the complete message of a certain round by eavesdropping, the attacker tries to reverse the hidden private data in the message of the previous round or a certain round from the current message, the protocol in the text cannot succeed in terms of the attacker. The addition of the random number a or the random number b makes the value of the message different each time, although the rules for calculating each message are the same, but the values are different, and there is no correlation between the values calculated in the two previous rounds, making it impossible for the attacker to backtrack.

Counterfeit attacks. The success of an attacker to perform a fake attack cannot be in the following ways:

one, replaying a previously intercepted message that passes authentication; two, one party fails to authenticate the other; and three, some private data can be obtained and the attacker calculates the correct message value by himself.

For one of the attacks, which was analyzed in detail in the previous replay attack, the attacker could not succeed.

For the second attack method, the protocol proposed by He *et al.* is this case, which makes the protocol unable to resist the impersonation attack. However, before each step of the protocol in the paper is performed, the receiver first verifies the source of the message, and only if the verification is passed, the receiver will proceed to the subsequent steps, so the impersonation attack fails in this case.

For its three attack methods, the attacker cannot obtain any useful privacy data in the text protocol at all, and the messages in the text are sent after encryption, even if the attacker can obtain the tag pseudonym, but the pseudonym is not important privacy data, and the attacker obtains it without causing danger to the user. The attacker can only randomly select parameters for calculation without obtaining useful data, but the counterfeit attack will only fail in this case.

Exhaustive attack. When an attacker is unable to obtain useful data by other means, he may think of the simplest and most direct means to crack, that is, by exhausting all possible values of the parameter by way of exhaustive enumeration. For example: the attacker by eavesdropping on a whole communication, can know the message $meg1, meg2, meg3, meg4, meg5$, and then according to the protocol implementation steps, the attacker will try to put the message $meg1, meg2$ together for exhaustive attack, the specific process can be analyzed as follows.

The attacker first deformed the message $meg1$, processing to get $a' = meg1 \oplus TID$ (here the attacker temporarily do not know the identifier of the tag, first choose a random value instead), and then the processing results combined with other parameters information in accordance with the rules of the protocol implementation steps to participate in the operation to get $meg2' = H(a'\&IDS, TID) = H((meg1 \oplus TID)\&IDS, TID)$, The attacker has now mastered $meg1, IDS$ data values, the attacker mistakenly believes that in $meg2'$ only a TID parameter is not known, so the attacker uses exhaustive means to try to exhaust all possible values of TID, and finally get the correct value of TID.

The above idea is only the attacker's own idea, and the attacker cannot break this protocol successfully in the protocol proposed by this paper. According to the definition of deformation Hash function in the text, the encryption operation will also involve the Hamming weight of the two parameters $(meg1 \oplus TID)\&IDS$ and TID. Without knowing the corresponding Hamming weight of these two parameters, the attacker cannot perform the encryption operation, and thus cannot perform the exhaustive enumeration.

Based on the above elaboration, the attacker exhaustive enumeration fails and hence the protocol is resistant to exhaustive attacks.

Positioning attack. When an attacker is unable to obtain important private information, the attacker may intentionally corrupt user resources, for example: the attacker intentionally damages the tag. If the attacker wants to intentionally damage the tag, or decrypt the location of the tag, the attacker must always have the exact location of the tag.

In the protocol, there are three messages sent out from the tag end of each round of communication, namely $IDS, meg3$ and $meg4$, Although IDS is sent in plaintext, it is only a pseudonym of the tag, and it is updated after each round of communication, so that the values sent out in each round are different; $meg3$ and $meg4$ are sent after encryption, and the attacker intercepts the ciphertext, and random numbers are added in the process of encryption for both. According to the above description, each round of communication, the value of the message sent by the tag is different and changing, and the attacker intercepts that the value is different each time, so it is impossible to locate the tag specifically, giving the attacker the impression that the tag is in a constant state of flux. Therefore, the location attack fails.

De-synchronization. It is necessary to ensure the consistency of shared information between different communication entities. When the shared information loses its consistency, authentication between the two is impossible and protocol authentication will only fail.

The protocol is designed to store the shared key K and pseudonym IDS for each round of communication at DB/R end to ensure the consistency of information between them. Specifically, step 5 shows that when DB/R cannot verify the tag with the current shared key K and pseudonym $IDS, DB/R$ will retrieve the shared key K and pseudonym IDS used in the previous authentication and will verify the tag again to restore the consistency between them.

The results of comparing the security aspects of other classical protocols with the protocol in the paper are shown in Table 1.

Table 1: Security Comparison Between Multiple Protocols

Attack Type	**Ref [11]**	**Ref [12]**	**Ref [13]**	**Ref [8]**	**Our Prot-ocol**
Replay Attack	√	√	√	×	√
Rearward safety	√	√	√	√	√
Counterfeit attack	×	√	×	×	√
Exhaustive attack	√	√	√	√	√
Positioning attack	√	√	√	√	√
Desynchron-ization	√	×	√	×	√

Notes: √ means can resist, × means cannot resist.

6 Logic Lightweight RFID Protocol Performance Analysis

This section will analyze the performance of the protocol in the paper and other comparison protocols from multiple perspectives. The tag is chosen as the object here because the tag cost is limited, so the tag computational power and storage space are not as powerful as the server end; the amount of computation at the tag end (abbreviated as metric A), the amount of storage at the tag end (abbreviated as metric B), the computation amount for a complete communication process (abbreviated as metric C), and the number of messages for a complete communication process (abbreviated as metric D) are chosen as the metrics for performance analyzed. The results of the comprehensive analysis and comparison are shown in Table 2.

The meanings of the different symbols appearing in Table 2 are as follows:$H()$ symbol represents the computation of the hash function operation, xor symbols represents the computation of the xor operation, $ECC()$ symbols indicates the computation of the elliptic curve ECC operation, AND symbol indicates the computation of the add operation, $PUF()$ symbol indicates the computation of the physical unclonable function operation, *mod* symbol indicates the computation of the mode operation, *or* symbol indicates the computation of the or operation, $PRNG()$ symbol indicates the computation of the pseudo-random function operation. The *and* symbol indicates the computation of with operations, and the L symbol indicates the message length.

The amount of computation at one end of the tag: the encryption algorithms used between the comparison protocols are not exactly the same, such as hash function, elliptic curve ECC, mode operation, physical unclonable function, pseudo-random function and other encryption algorithms, the computation of elliptic curve ECC and mode operation in the above encryption algorithms will be greater than other encryption operations. Although the hash function used in the protocol is not the traditional hash function, but an improved deformation hash function, it can be found that the deformed hash function is improved by adding partial reverse or exchange or XOR operations, and the computation required for these operations is negligible compared with the computation required for the hash function itself. Through Table 2 metric A can be found that the protocol in the text of the tag end of the computation is less than other comparative protocols, there is a certain advantage.

The amount of storage at one end of the tag: the protocol in the text needs to introduce a statistician at one end of the tag *count*, the statistician does not need to occupy a large space, only 1 *bit* can be used. It can be concluded from Table 2, metric B, that the storage space size at one end of the tag is comparable for each protocol.

The amount of computation for a complete communication: the amount of computation at the tag end is roughly equal to that at the server end, so the amount of computation for a complete communication is roughly equivalent to twice the amount of computation at the tag end. Still, it can be found from Table 2 that the protocol in the text has the least amount of computation in a complete session, which is still an advantage.

Number of messages for a complete communication: different protocols send messages such as $ACK, ACCEPT$ during the exchange process. These messages do not need to be stored in a space of length L, but only 1 *bit*, Table 2 shows that the number of messages for a complete communication is also less than that of the other protocols in the comparison.

7 Logical Formal Analysis of Lightweight RFID Protocol

This chapter will use GNY [1] logic formalization to analyze the protocol in the text from a formal perspective.

Formal Model.

The protocol is formally modeled using GNY logic as follows:

$Msg1 : DB/R \rightarrow T : ACK$

$Msg2 : T \rightarrow DB/R : IDS$

$Msg3 : DB/R \rightarrow T : meg1, meg2$

$Msg4 : T \rightarrow DB/R : meg3, meg4$

$Msg5 : DB/R \rightarrow T : meg5$

The protocol is formally modeled using GNY logic as follows:

$Msg1 : T \triangleleft *ACK \sim | \rightarrow DB/R| \equiv \#ACK$

$Msg2 : DB/R \triangleleft *IDS \sim | \rightarrow T| \equiv \#IDS$

Table 2: Performance Analysis Between Multiple Protocols

Impact indicators	Metric A	Metric B	Metric C	Metric D
Ref [11]	$5PRNG() + 4xor + 2and$	$4L$	$11PRNG() + 9xor + 5and$	$8L + 1bit$
Ref [12]	$6PUF() + 3xor + 1or$	$4L$	$11PUF() + 6xor + 3or$	$7L + 1bit$
Ref [13]	$3PUF() + 2mod + 4xor$	$3L$	$6PUF() + 5mod + 9xor$	$7L$
Ref [8]	$4ECC() + 4H() + 1AND$	$5L$	$8ECC() + 8H() + 2AND$	$7L + 2bit$
Our protocol	$5H() + 2xor$	$3L + 1bit$	$10H() + 4xor$	$6L + 1bit$

$Msg3 : T \lhd *(meg1, meg2) \sim | \to DB/R| \equiv \#(meg1, meg2)$

$Msg4 : DB/R \lhd *(meg3, meg4) \sim | \to T| \equiv \#(meg3, meg4)$

$Msg5 : T \lhd *meg5 \sim | \to DB/R| \equiv \#meg5$

Initialization Assumptions.

$A1 : DB/R \ni K$

$A2 : DB/R \ni a$

$A3 : DB/R \ni TID$

$A4 : DB/R \ni IDS$

$A5 : T \ni b$

$A6 : T \ni K$

$A7 : T \ni TID$

$A8 : T \ni IDS$

$A9 : DB/R| \equiv \#(a)$

$A10 : T| \equiv \#(b)$

$A11 : DB/R| \equiv DB/R \overset{K}{\longleftrightarrow} T$

$A12 : DB/R| \equiv DB/R \overset{TID}{\longleftrightarrow} T$

$A13 : DB/R| \equiv DB/R \overset{IDS}{\longleftrightarrow} T$

$A14 : T| \equiv T \overset{K}{\longleftrightarrow} DB/R$

$A15 : T| \equiv T \overset{TID}{\longleftrightarrow} DB/R$

$A16 : T| \equiv T \overset{IDS}{\longleftrightarrow} DB/R$

Demonstrate Objectives.

The following five objectives need to be proved:

$G1 : DB/R| \equiv T| \sim \#(meg4)$

$G2 : DB/R| \equiv T| \sim \#(meg3)$

$G3 : T| \equiv DB/R| \sim \#(meg1)$

$G4 : T| \equiv DB/R| \sim \#(meg2)$

$G5 : T| \equiv DB/R| \sim \#(meg5)$

Reasoning Proof.

Although there are five proving objectives to be proved, the proving process of the five proving objectives is roughly the same. In addition to the limited space and other factors in the text, only the proving objectives are selected as examples for the derivation and proof. The proving process is as follows:

$\because A9 : DB/R| \equiv \#(a), F1 : \frac{P|\equiv\#(X)}{P|\equiv\#(X,Y), P|\equiv\#(F(X))}$

$\therefore DB/R| \equiv \#(a, K)$

$\because Msg4 : DB/R \lhd *(meg3, meg4) \sim | \to T| \equiv \#(meg3, meg4), DB/R \lhd *a$, approach $DB/R \ni a$

$\because A1 : DB/R \ni K, A2DB/R \ni a, A3 : DB/R \ni TID, A4 : DB/R \ni IDS, P2$

$\therefore DB/R \ni (a, K)$

$\because DB/R| \equiv \#(a, K), DB/R \ni (a, K)$

$\because F10 : \frac{P|\equiv\#(X), P \ni X}{P|\equiv\#(H(X,Y))}$

$\therefore DB/R| \equiv \#(meg4)$, approach $DB/R| \equiv \#(H(a\&K, b))$

$\because Msg4 : DB/R \lhd *(meg3, meg4) \sim | \to T| \equiv \#(meg3, meg4), A11 : DB/R| \equiv DB/R \overset{K}{\longleftrightarrow} K, I3$

$\because DB/R \ni (a, K), DB/R| \equiv \#(H(a\&K, b))$

$\therefore DB/R| \equiv T| \sim \#(meg4)$, approach $DB/R| \equiv T| \sim \#(H(a\&K, b))$

$\therefore G1 : DB/R| \equiv T| \sim \#(meg4)$, approach $G1 : DB/R| \equiv T| \sim \#(H(a\&K, b))$

Complete proof.

8 Conclusion

The security problems in the application of RFID technology are introduced. This paper focuses on the analysis of many problems or areas to be discussed in the design protocol of He *et al.* in recent years, and proposed an improved security protocol. The original protocol uses elliptic curve ECC and hash function to encrypt data transmission, but the low-cost tags cannot meet the above requirements, so the security protocol in this paper abandons the original protocol encryption algorithm and adopts a self-designed deformation Hash function algorithm to achieve important data encryption. In this paper, we give the steps of how to implement the deformation Hash function algorithm, and cleverly use the encryption parameters themselves Hamming weight. This protocol can reduce the amount of space storage and increase the attack difficulty without adding new parameters. The analysis of the security and performance metrics of the proposed protocols is carried out from different aspects. According to the results of the comparative analysis, it can be shown that the proposed protocols are resistant to many common types of attacks, meanwhile, the results shows that the performance in terms of the computation amount at the tag and the total computation amount for complete communication are better than other comparative protocols.

References

[1] M. Bellare, P. Rogaway, "Random oracles are practical: A para-digm for designing efficient protocols," in *Proceedings of the first Annual Conference on Computer and Communications Security, Virginia, United States, November 03-05, 1993*, pp. 62–73, New York, NY: ACM, 1993.

[2] Y. C. Chen, W. L. Wang, M. S. Hwang, "RFID authentication protocol for anti-counterfeiting and privacy protection", in *The 9th International Conference on Advanced Communication Technology*, pp. 255-259, 2007.

[3] J. Y. Chun, G. Noh, "Privacy-preserving RFID-based search system," *Electronics*, vol. 10, no. 1, pp. 1–13, 2021.

[4] N. Dinarvand, H. Barati, "An efficient and secure RFID authentication protocol using elliptic curve cryptography," *Wireless Networks*, vol. 25, no. 1, pp. 415–428, 2019.

[5] K. Fan, S. Zhu, K. Zhang, et al, "A lightweight authentication scheme for cloud-based RFID healthcare systems," *IEEE Network*, vol. 33, no. 2, pp. 44–49, 2019.

[6] M. Y. Fan, Q. K. Dong, L. Wang, et al, "Ttp-free weighted muti-owner RFID tag authentication protocol," *Journal of Xidian University*, vol. 48, no. 1, pp. 133–140, 2021.

[7] P. Gope, J. Lee, T. Q. S. Quek, "Lightweight and practical anonymous authentication protocol for RFID systems using physically unclonable functions," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 11, pp. 2831–2843, 2018.

[8] J. Q. He, C. G. Peng, Z. J. Fu, et al, "Lightweight bidirectional authentication protocol for rfid," *Computer Engineering and Applications*, vol. 32, no. 4, pp. 194–199, 2023.

[9] M. S. Hwang, C. H. Wei, C. Y. Lee, "Privacy and security requirements for RFID applications", *Journal of Computers*, vol. 20, no. 3, pp. 55-60, Oct. 2009.

[10] T. Jager, J. Schwenk, J. Somorovsky, "Practical invalid curve attacks on tls-ecdh," pp. 407–425, 2015.

[11] V. Kumar, M. Ahmad, D. Mishra, et al, "RSEAP: RFID based secure and efficient authentication protocol for vehicular cloud computing," *Vehicular Communications*, vol. 22, no. 2, pp. 210–213, 2020.

[12] T. Li, Y. L. Liu, "A double puf-based RFID authentication protocol," *Journal of Computer Research and Development*, vol. 58, no. 8, pp. 1801–1810, 2021.

[13] W. Liang, S. Xie, J. Long, et al, "A double puf-based RFID identity authentication protocol in service-centric internet of things environments," *Information Sciences*, vol. 50, no. 3, pp. 129–147, 2019.

[14] D. W. Liu, J. Ling, "An improved RFID authentication protocol with backward privacy," *Computer Science*, vol. 43, no. 8, pp. 128–130, 2016.

[15] Y. L. Liu, X. C. Yin, Y. Q. Dong, et al, "Lightweight authentication scheme with inverse operation on passive RFID tags," *Journal of the Chinese Institute of Engineers*, vol. 42, no. 1, pp. 74–79, 2019.

[16] S. Q. Mei, X. R. Deng, "Mobile RFID bidirectional authentication protocol based on shared private key and bitwise operation," *Computer Applications and Software*, vol. 37, no. 7, pp. 302–308, 2020.

[17] G. F. Shen, S. M. Gu, D. W. Liu, "An anti-counterfeit complete RFID tag grouping proof generation protocol," *International Journal of Network Security*, vol. 21, no. 6, pp. 889–896, 2019.

[18] F. Tan, "An improved RFID mutual authentication security hardening protocol," *Control Engineering of China*, vol. 26, no. 4, pp. 783–789, 2019.

[19] J. Q. Wang, Y. F. Zhang, D. W. Liu, "Provable secure for the ultra-lightweight RFID tag ownership transfer protocol in the context of IoT commerce," *International Journal of Network Security*, vol. 22, no. 1, pp. 12–23, 2020.

[20] C. H. Wei, M. S. Hwang, Augustin Y. H. Chin, "A secure privacy and authentication protocol for passive RFID tags," *International Journal of Mobile Communications*, vol. 15, no. 3, pp. 266–277, 2017.

[21] C. H. Wei, M. S. Hwang, Augustin Y. H. Chin, "An improved authentication protocol for mobile agent device in RFID," *International Journal of Mobile Communications*, vol. 10, no. 5, pp. 508–520, 2012.

[22] R. Xie, B. Y. Jian, D. W. Liu, "An improved ownership transfer for RFID protocol," *International Journal of Network Security*, vol. 20, no. 1, pp. 149–156, 2018.

[23] R. Xie, J. Ling, D. W. Liu, "A wireless key generation algorithm for RFID system based on bit operation," *International Journal of Network Security*, vol. 20, no. 5, pp. 938–950, 2018.

Biography

Fang-ming Cao graduated from South China University of technology with a master's degree in 2016. Now working in Guangdong Peizheng University, he is a full-time teacher of computer major, and also a lecturer. At present, the research direction is mainly in the field of information security, etc.

Xiao-ping He received a master's degree from software engineering of Huazhong University of Science and Technology (China) in June 2013. He is now an associate professor and works in Guangdong Peizheng College. His curnet research interest fields include information security and network security.

An Improved Secure RFID Authentication Protocol Using Elliptic Curve Cryptography

Wan-Rong Liu, Zhi-Yong Ji, and Cheng-Chen Chu

(Corresponding author: Cheng-Chen Chu)

Shanghai Sixth People's Hospital Affiliated to Shanghai Jiao Tong University School of Medicine
Shanghai 201306, China
Email: 1052529220@qq.com

(Received Mar. 3, 2023; Revised and Accepted Oct. 26, 2023; First Online Dec. 24, 2023)

Abstract

Radio-frequency identification technology is one of the critical technologies of the Internet of Things. It is a non-contact automatic identification technology. People tend only to pay attention to the accelerated development of its technology and ignore the security problems of the technology; radio-frequency identification technology development is no exception. Its security problems are becoming increasingly prominent. Because the radio-frequency identification system in the Internet of Things carries a large amount of user privacy information, its security problem will directly restrict the promotion and application of the Internet of Things. Therefore, the research of radio-frequency identification security authentication technology is significant to the healthy development of the Internet of Things. We propose an improved secure RFID authentication protocol using elliptic curve cryptography based on Izza *et al.*'s protocol. BAN logic is used to prove the improved protocol. Meanwhile, we carry out a comparative analysis of performance and efficiency. The results show that the improved protocol has higher security and lower calculation costs.

Keywords: Anonymity; Authentication; Elliptic Curve Cryptography; Encryption; Radio-Frequency Identification

1 Introduction

The Internet of Things (IoT) is a network that uses Radio-Frequency Identification (RFID) technology, wireless communication technology and computer technology to realize the interconnection between things. It makes the real-time information exchange between machines without human participation, which greatly facilitates people's work and life. RFID technology is one of the key technologies of the IoT. It is a contactless automatic identification technology. It uses radio frequency to carry out non-contact two-way data communication, identify the target entity and obtain relevant data, so as to realize the tracking and information sharing of objects. Its biggest advantage for the IoT is non-contact, with fast identification speed, small size, strong penetration and other characteristics [16, 17].

RFID tags are mainly divided into two types, namely active tags and passive tags [5, 11, 26–29]. The power supply of the active tag depends on the internal battery power supply. Passive tags rely on the power of the radio waves generated by the card reader, by converting some of the energy of the waves into direct current for working use. With the continuous development of technology, semi-passive electronic tags have been developed. The ultimate purpose of electronic tags is to identify objects regardless of the power supply, but the two types of tags are used in different scenarios.

The RFID system is generally composed of RFID tag, reader and back-end database system. See the Figure 1 for its basic structure. RFID tags and readers exchange data through radio frequency signals. When the RFID tag enters the magnetic field range of the RFID reader, it receives the radio frequency signal from the reader. At this time, the tag will generate enough induced current to obtain energy to activate the tag. The label then sends out the product information stored in the chip in the form of a carrier signal (passive label), or an active frequency signal (active label). After receiving the signal, the reader reads the information it contains and decodes it. It is then sent to the back-end database for related data processing.

Compared with traditional barcode tags, RFID tags are more reasonable and resourceful. It has also been increasingly widely used in the medical system, such as the application of supply, processing, and distribution (SPD). It introduces the concept of "Internet +" supply chain management and modern pharmaceutical logistics management into the purchase and management of hospital drugs. Optimize and reorganize the whole supply chain management process of hospitals from medical consumables procurement, warehousing, distribution to clinical consumption, and establish a new supply chain manage-

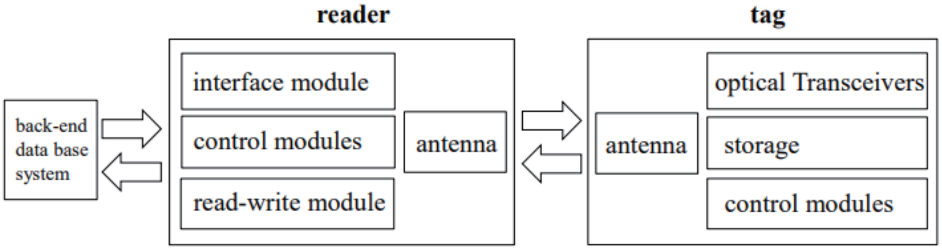


Figure 1: Basic structure of RFID system

ment mode. However, people tend to focus only on the accelerated development of its technology and ignore the security problems of the technology. The development of RFID technology is no exception, and its security is becoming increasingly prominent. Because the RFID system in the IoT carries a large amount of user privacy information, its security problem will directly restrict the promotion and application of the IoT. Therefore, the research of RFID security authentication technology is of great significance to the healthy development of the IoT.

Alamr *et al.* [2] proposes a protocol using elliptic-curve Diffie-Hellman (ECDH) for RFID system. Naeem *et al.* [21] indicates that Alamr *et al.*'s protocol is not scalable and the scheme can accommodate only one tag. There is only one tag at all is not practical. In real-world scenarios, a reader is expected to work with hundreds of even thousands of tags. Dinarvand *et al.* [7] believes that Alamr *et al.*'s protocol cannot resist de-synchronization and Denial of Service attack. Further, based on the schema architecture of Alamr *et al.*, Naeem *et al.* makes minor modification to the initialization and authentication phases. Abughazalah *et al.* [1] proposes an improved cloud-based RFID authentication protocol. Fan *et al.* [8] points out that $H(ID_i)$ is transmitted in plain text, and the next authentication is equal to $H(ID_i)$ in Abughazalah *et al.*'s protocol, so that the protocol cannot provide the backward security and tags may be tracked. In 2020, Fan *et al.* proposes a cloud-based RFID authentication protocol to ensure that the tag overhead is lightweight. Nagarajan *et al.* [22] points out that Fan *et al.*'s protocol cannot resist to known session-specific temporary information attack, impersonation attack, and man-in-the-middle attack in 2022. Then, Nagarajan *et al.* patches the security threats in the Extended Tiny Encryption Algorithm (XTEA) by applying domain-specific customization, random number utilization, and undisclosed key renewal techniques. Mishra *et al.* [20] presents a chaotic map-based mutual authentication framework for RFID-authentication. Safkhani *et al.* [25] indicates Mishra *et al.*'s protocol cannot resist tag/user impersonation attack, reader impersonation attack and man-in-the-middle attack. Meanwhile, Safkhani *et al.* points out that Jang *et al.*'s protocol [13] cannot resist relay attack. In 2021, Safkhani *et al.* proposes an improved protocol, called an enhanced version of RFID based secure and efficient authentication protocol (RSEAP2). Their research shows that RSEAP2 is more efficient (computation and communication costs) than the original RSEAP which proposed by Kumar *et al.* [15]. In 2021, Gabsi *et al.* [9] performs a comparative study between partial ECC-based RFID authentication protocols in terms of security and performance.

Izza *et al.* [12] proposes a RFID authentication protocol for WBAN within an IoT environment in 2021. We find that there are some problems with Izza *et al.*'s protocol: (1) Failure to offer tag anonymity. (2) Failure to provide forward secrecy. (3) Failure to offer clock synchronization mechanism. We believe that the agreement of Izza *et al.*'s protocol has a good framework. Based on Izza *et al.*'s protocol, we propose an improved secure RFID authentication protocol using elliptic curve cryptography (ECC) [18].

ECC is an algorithm for establishing public key encryption, based on elliptic curve mathematics. The essence of ECC encryption algorithm is the equation calculation of the curve on the number axis, through the number calculation to get encryption/decryption. Public key algorithms are always based on a mathematical puzzle. The following are a few recognized "hard problems" within elliptic curves.

1) Discrete Logarithm problem: given $P, aP \in E/Fq$, for unknown $a \in Z_n*$, the probability of success of finding the value of is negligible.

2) Bilinear Diffie-Hellman problem: given (P, aP, bP, cP), $a, b, c \in Z_n*$, the probability of success of judging whether the equation $e(P,P)^d = e(P,P)^{abc}$is correct is negligible.

3) Computational Diffie-Hellman problem: given P, aP, $bP, P \in E/Fq$, for unknown $a, b \in Z_n*$, the probability of success of finding the value of is negligible.

The main contributions of this paper: (1) Reviewing and analyzing the existing problems of Izza *et al.*'s protocol. (2) Based on ECC, an improved authentication protocol is proposed. (3) BAN logic is used to prove the improved protocol. (4) Comparative analysis of performance and efficiency. The results show that the improved protocol has higher security and lower calculation cost.

2 Izza *et al.*'s Protocol Description

The notations used in the paper are summarized in Table 1.

Table 1: Notations

Symbol	Definition
U_j	Network user
x_j	User secret key
y_j	User public key
β	NM public key
α	NM secret key
P_{rR}/P_{rS}	Reader/sever private key
P_{uR}/P_{uS}	Reader/sever public key
n	Number of tags
ID_{T_i}	Tad identity of i
PID_i	Tag's pseudo password of i
P	A point on the elliptic curve
$h(\cdot)$	One-way hash function
$\oplus$	Bitwise XOR operation
NM	Network manager
$\|\|$	Concatenation operation
SK_{TR}/SK_{RT}	Session-key between tag and reader
T	The current time of system

2.1 Initialization Phase

A trust worthy authority Network Manager (NM) selects an elliptic curve E_q over a prime field F_q where q is a prime number and is the elliptic curve base point of order n that is shared with all the network parties.

Step 1: NM selects a random integer $\alpha \in [1, n-1]$, then calculates $\beta = \alpha.P$ as its public key.

Step 2: Each entity generates its own random integer $c_j \in [1, n-1]$ and calculates $d_j = c_j.P$, then sends (d_j, ID_j) to NM.

Step 3: NM selects a random integer $k_j \in [1, n-1]$ and computes $y_j = k_j.P + d_j$ and $z_j = k_j + ((y_j)_x + ID_j)\alpha \bmod n$.

Step 4: NM returns back (y_j, z_j) with which the user can compute its secret key $x_j = z_j + c_j \bmod n$, then checks $x_j.P = y_j + ((y_j)_x + ID_j)\beta$.

2.2 Registration Phase

Step 1: The sever selects the tag identities ID_{T_i}, where $i = \{1, 2, ..., n\}$, and calculate $PID_{Told} = h(ID_{T_i})$ then stores ID_{T_i} and PID_{Told} in the tag's and the reader's memory.

Step 2: S selects a random value *init* and inserts it in both the tag's and the reader's memory. Then S initializes $PID_{Tinew} = h(PID_{Tiold}||init)$ and stores it in the reader's memory.

Step 3: S selects its identity ID_S, the reader's identity ID_R and computes $PID_{Rold} = h(ID_R)$, then stores them in the reader's memory, as well as in its memory.

Step 4: Then S initializes $PID_{Rnew} = h(PID_{Tiold}||init)$ and stores it in its memory.

Step 5: S picks a random number P_{rR} that represents the reader's secret key and assigns $P_{uR} = P_{rR}.P$ as being the public key. Then, it stores the key pair in the reader's memory. Sselects a random number P_{rS} as being its secret key and computes $P_{uS} = P_{rS}.P$ that represents its public key, then stores the key pair in its database.

Step 6: S inserts P_{uR} and P_{uS} in the tag's memory. It also inserts P_{uS} in the reader's database and P_{uR} in its database.

2.3 Authentication and Data Transmission Phase

Step 1: The reader R generates r_1 and computes $R_{r_1} = r_1.P$ and sends $\{R_{r_1}\}$ to the tag T.

Step 2: T generates t_1 and computes $C_1 = t_1.P$, $R_{t_1} = t_1.P_{uR}$. Then T initializes $PIDT_{new} = h(PID_{Tiold}||init)$, $C_2 = PID_{Tinew} + h((R_{t1})_x|| (R_{r1})_x||(C_1)_x||T_1)$ and sends $\{C_1, C_2, T_1\}$ to R.

Step 3: R checks whether the condition $|T_2 - T_1| < \Delta T$ holds or not, if it holds, R calculates $R^*_{t1} = C_1.P_{rR}$, $PID^*_{T_i} = C_2 - h((R^*_{t1})x||(R_{r1})_x||(C_1)_x||T_1)$ and checks if the pseudo identity corresponds to the one existing in its database to authenticate the tag.

Step 4: R computes $N_1 = r_1.P_{uS}$, initializes$PID_{Rnew} = h(PID_{Rold}||init)$, $N_2 = PID_{Rnew} + h((R_{r_1})_x|| ID_R||(N_1)_x||T_2)$, then sends $\{N_2, R_{r1}, T_2\}$ to S.

Step 5: S checks whether the condition $|T_3 - T_2| < \Delta T$ holds or not, if it holds, S computes $N^*_1 = R_{r1}.P_{rS}$,$PID^*_R = N_2 - h((R_{r1})_x||ID_R||(N^*_1)_x||T_2)$, checks PID^*_R in database. S generates S_1 and computes $S_1 = s1.P$, $R_{s1} = s1.P_{uR}$, $N_3 = h((R_{s1})_x||PID^*_R||T_2||T_3) + ID_S$. If the condition $PID^*_R = PID_{Rold}$ holds, it updates $PID_{Rold} \leftarrow PID_{Rold}$, $PID_{Rnew} \leftarrow h(PID_{Rold}||(N_1)_x)$, else if the condition $PID^*_R = PID_{Rnew}$ holds, updates $PID_{Rold} \leftarrow PID_{Rnew}$, $PID_{Rnew} \leftarrow h(PID_{Rnew}||(N_1)_x)$. Then, S sends $\{N_3, S_1, T_3\}$ to R.

Step 6: R. checks whether the condition $|T_4 - T_3| < \Delta T$ holds or not, if it holds, computes $R^*_{S1} = S_1.P_{rR}$, $ID^*_S = N_3 - h((R^*_{S1})_x||PID_{Rnew}||T_2||T_3)$. Then, it authenticates the server if ID^*_S corresponds to the

ID_S stored in its database, otherwise, the session is canceled. After a successful server authentication, R computes $C_3 = h(ID_{T_i}||T_3||T_4) + PID_{Rnew}$, $C_4 = h((R^*_{t_1})_x||PID_{Rnew}||(R_{r1})_x||T_4)$. If the condition $PID^*_{T_i} = PID_{Told}$ holds, it updates $PID_{Told} \leftarrow PID_{Told}$, $PID_{Tnew} \leftarrow h(PID_{Told}||(R_{t1})_x)$, else if the condition $PID^*_{T_i} = PID_{Tnew}$ holds, it updates $PID_{Told} \leftarrow PID_{Tnew}$, $PID_{Tnew} \leftarrow h(PID_{Tnew}||(R_{t_1})_x)$.

Step 7: R updates $PID_{Rnew} \leftarrow h(PID_{Rnew}||(N_1)_x)$, generates the shared session key $SK_{RT} = h(ID_{Ti}||PID_{Tinew}||(r_1.C_1)_x)$. R sends $\{C_4, C_3, T_3, T_4\}$ to T.

Step 8: T checks whether the condition $|T_5 - T_4| < \Delta T$ holds or not, if it holds, T computes $PID^*_{Rnew} = C_3 - h(ID_{T_i}||T_3||T_4)$, $C^*_4 = h((R_{t_1})_x||PID^*_{Rnew}||(R_{r_1})_x||T_4)$. T checks whether the condition $C^*_4 \overset{?}{=} C_4$ holds or not, if it holds, updates $PID_{Tinew} \leftarrow h(PID_{Tinew}||(R_{t_1})_x)$, generates the shared session key $SK_{TR} = h(ID_{T_i}||PID_{Tinew}||(t_1.R_{r_1})_x)$.

Once the shares session key is derived, $T$and can exchanged sensitive data security thanks to the symmetric encryption. The procedures are described as follow:

Step 1: T generates a menssage m_i and computes $M_i = E_{SK}(m_i)$. Then T sends $\{M_i, T_5\}$ to R.

Step 2: R checks whether the condition $|T_6 - T_5| < \Delta T$ holds or not, if it holds, R computes $m_i = D_{SK}(M_i)$, makes $r_0 = 0$. Then R selects a random number $k \in [1, n-1]$ and calculates $r_i = m_i + h(r_i - 1 \oplus (k(y_S + ((y_S)_x + IDS)\beta))_x) \bmod n$. Then, it deduces $r = h(r_1||r_2||r_3||...||r_n)$ and it computes $z = k - rxR \bmod n$. R sends $\{r, z, r_1, r_2, r_3, ..., r_n, T_6\}$ to S.

Step 3: S checks whether the condition $|T_7 - T_6| < \Delta T$ holds or not, if it holds, S computes $r^* = h(r_1||r_2||r_3||...||r_n)$ and checks whether $r \overset{?}{=} r^*$, if it holds, S extracts the set of message $m_1, m_2, ..., m_n$ as follow: $m_i = r_i - h(r_{(i-1)} \oplus (zP + r(y_R + ((y_S)_x + ID_R)\beta)xS)x) \bmod n$, where $i = 1, 2, ..., n$ and $r_0 = 0$.

3 Vulnerabilities of Izza *et al.*'s Protocol

3.1 Weakness 1: No Anonymity

Arslan has shown in his paper [3] that the protocol suffers particularly from the existing relationship between the C_2 message and the long-term identity of the tag ID_T. During the Step 2 of Izza *et al.*'s protocol, the inputs of the first hash function, computed in $PID_{Tinew} = h(PID_{Tiold}||init)$, combine the pseudo tag identity and the random variable . The second value $C_2 = PID_{Tinew} + h((R_{t1})_x||(R_{r1})_x||(C_1)_x||T_1)$ performs the hash operation between the values $(R_{t1}\ and\ C_1)$ generated by the tag and the public value R_{r1}. Consequently, since the *init* and R_{r1} value can be manipulated by an attacker, these two hash operations can be the target of a SCA attack in order to extract the pseudo identity of the tag. Therefore, Izza *et al.*'s scheme cannot offer tag anonymity.

3.2 Weakness 2: Provide Forward Secrecy

Let we assume that this is a legitimate tag called T_a and a legitimate R in this system. R executes several sessions with randomly selected a tag in a time interval.

Step 1: The adversary A records the parameters of the consecutive protocol session S_a executing between T_a and R. S_a includes the following set of protocol transaction parameters:

$$SPa : [{}^aRr1, {}^aC1, {}^aC2, {}^aC3, {}^aC4, {}^aT1, {}^aT2, {}^aT3, {}^aT4, {}^aT5, {}^aN2, {}^aN3, {}^aS1, {}^aMi]$$

Step 2: A arbitrarily selects a tag, called T_a and obtains the internal knowledge of T_b, called ϕ^{Tb}. ϕ^{Tb} : $[ID_{Tb}, PID_{Tbold}, n, P, P_{uR}, P_{uS}, init]$.

Step 3: A computes $PID_{Tbnew} = {}^aC_3 - h(ID_{Tb}||{}^aT_3||{}^aT_4)$, $PID_{Tbnew'} = h(PID_{Tbold}||init)$. Then A checks $PID_{Tbnew} \overset{?}{=} PID_{Tbnew'}$, if the verification is succeeded, A claims that $T_a = T_b$.

Therefore, this scheme does not provide forward secrecy.

4 Proposed Protocol

4.1 Initialization and Registration Phase

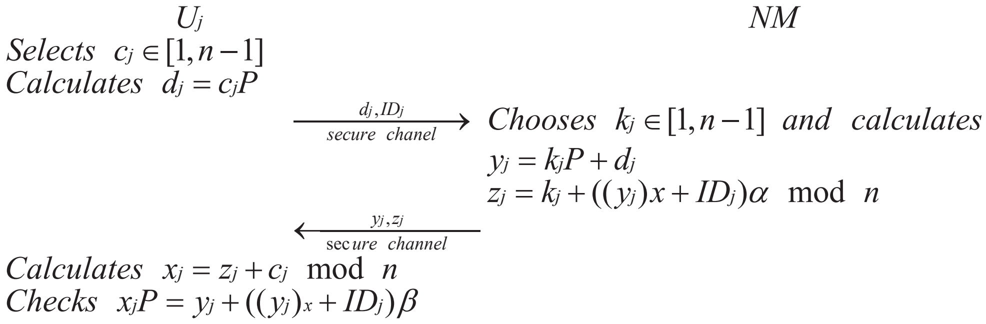
The initialization and registration phase of the proposed scheme is shown in Figure 2.

4.2 Authentication Phase

The Authentication phase of the proposed scheme is shown in Figure 3.

4.3 Digital Signature and Data Transmission Phase

The digital signature and data transmission phase of the proposed scheme is shown in Figure 4.



U_j $\quad$ NM

$Selects \;\; c_j \in [1, n-1]$

$Calculates \;\; d_j = c_j P$

$\xrightarrow[secure\;\; chanel]{d_j, ID_j}$ $Chooses \;\; k_j \in [1, n-1] \;\; and \;\; calculates$

$y_j = k_j P + d_j$

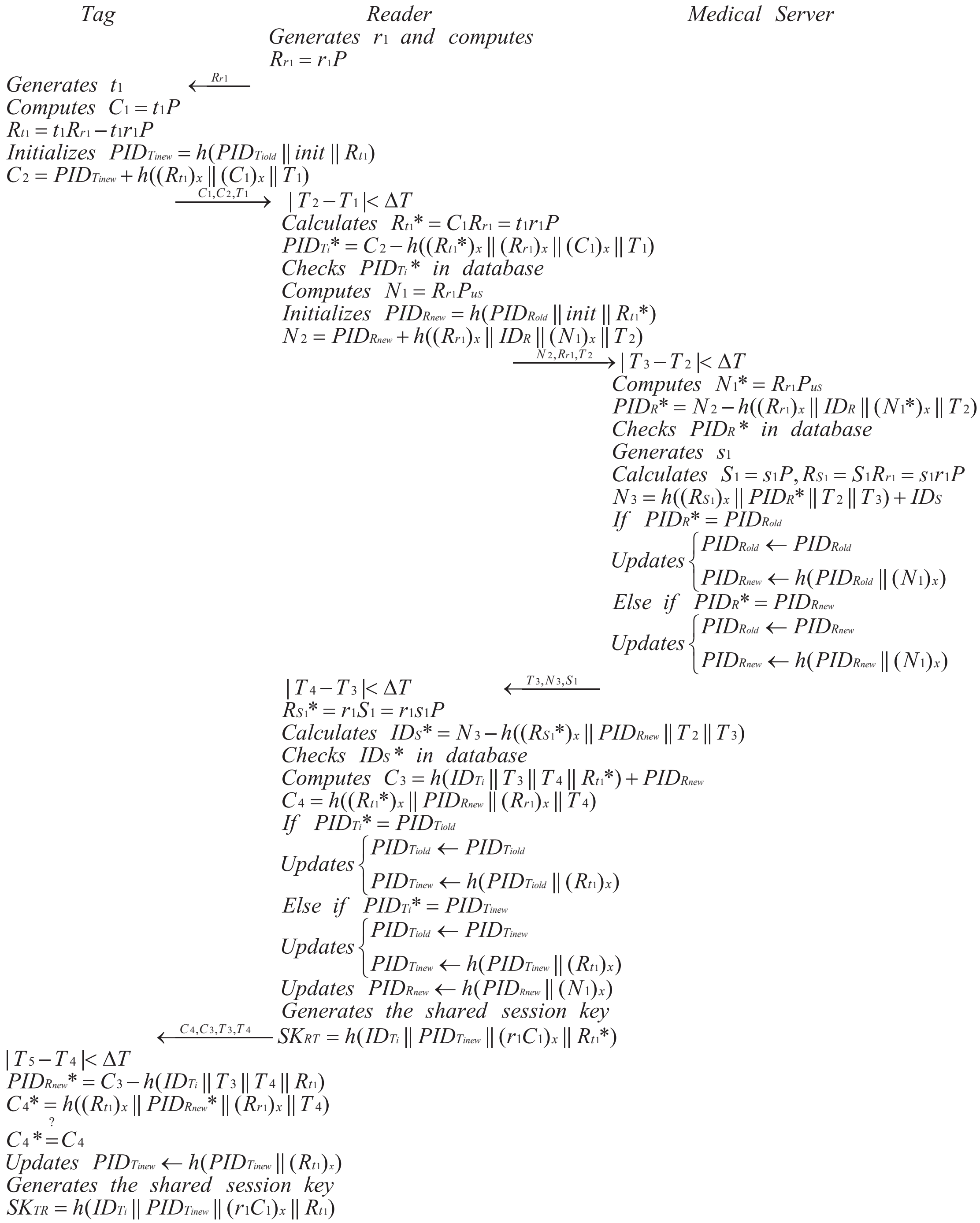
$z_j = k_j + ((y_j)x + ID_j)\alpha \mod n$

$\xleftarrow[\sec ure\;\; channel]{y_j, z_j}$

$Calculates \;\; x_j = z_j + c_j \mod n$

$Checks \;\; x_j P = y_j + ((y_j)_x + ID_j)\beta$

Figure 2: Initialization and registration phase



Tag $\quad$ $Reader$ $\quad$ $Medical \;\; Server$

$Generates \;\; r_1 \;\; and \;\; computes$

$R_{r1} = r_1 P$

$Generates \;\; t_1$ $\xleftarrow{R_{r1}}$

$Computes \;\; C_1 = t_1 P$

$R_{t1} = t_1 R_{r1} - t_1 r_1 P$

$Initializes \;\; PID_{Tinew} = h(PID_{Tiold} \| init \| R_{t1})$

$C_2 = PID_{Tinew} + h((R_{t1})_x \| (C_1)_x \| T_1)$

$\xrightarrow{C_1, C_2, T_1}$ $|T_2 - T_1| < \Delta T$

$Calculates \;\; R_{t1}^* = C_1 R_{r1} = t_1 r_1 P$

$PID_{Ti}^* = C_2 - h((R_{t1}^*)_x \| (R_{r1})_x \| (C_1)_x \| T_1)$

$Checks \;\; PID_{Ti}^* \;\; in \;\; database$

$Computes \;\; N_1 = R_{r1} P_{uS}$

$Initializes \;\; PID_{Rnew} = h(PID_{Rold} \| init \| R_{t1}^*)$

$N_2 = PID_{Rnew} + h((R_{r1})_x \| ID_R \| (N_1)_x \| T_2)$

$\xrightarrow{N_2, R_{r1}, T_2}$ $|T_3 - T_2| < \Delta T$

$Computes \;\; N_1^* = R_{r1} P_{uS}$

$PID_R^* = N_2 - h((R_{r1})_x \| ID_R \| (N_1^*)_x \| T_2)$

$Checks \;\; PID_R^* \;\; in \;\; database$

$Generates \;\; s_1$

$Calculates \;\; S_1 = s_1 P, R_{S1} = S_1 R_{r1} = s_1 r_1 P$

$N_3 = h((R_{S1})_x \| PID_R^* \| T_2 \| T_3) + ID_S$

$If \;\; PID_R^* = PID_{Rold}$

$Updates \begin{cases} PID_{Rold} \leftarrow PID_{Rold} \\ PID_{Rnew} \leftarrow h(PID_{Rold} \| (N_1)_x) \end{cases}$

$Else \;\; if \;\; PID_R^* = PID_{Rnew}$

$Updates \begin{cases} PID_{Rold} \leftarrow PID_{Rnew} \\ PID_{Rnew} \leftarrow h(PID_{Rnew} \| (N_1)_x) \end{cases}$

$|T_4 - T_3| < \Delta T$ $\xleftarrow{T_3, N_3, S_1}$

$R_{S1}^* = r_1 S_1 = r_1 s_1 P$

$Calculates \;\; ID_S^* = N_3 - h((R_{S1}^*)_x \| PID_{Rnew} \| T_2 \| T_3)$

$Checks \;\; ID_S^* \;\; in \;\; database$

$Computes \;\; C_3 = h(ID_{Ti} \| T_3 \| T_4 \| R_{t1}^*) + PID_{Rnew}$

$C_4 = h((R_{t1}^*)_x \| PID_{Rnew} \| (R_{r1})_x \| T_4)$

$If \;\; PID_{Ti}^* = PID_{Tiold}$

$Updates \begin{cases} PID_{Tiold} \leftarrow PID_{Tiold} \\ PID_{Tinew} \leftarrow h(PID_{Tiold} \| (R_{t1})_x) \end{cases}$

$Else \;\; if \;\; PID_{Ti}^* = PID_{Tinew}$

$Updates \begin{cases} PID_{Tiold} \leftarrow PID_{Tinew} \\ PID_{Tinew} \leftarrow h(PID_{Tinew} \| (R_{t1})_x) \end{cases}$

$Updates \;\; PID_{Rnew} \leftarrow h(PID_{Rnew} \| (N_1)_x)$

$Generates \;\; the \;\; shared \;\; session \;\; key$

$\xleftarrow{C_4, C_3, T_3, T_4}$ $SK_{RT} = h(ID_{Ti} \| PID_{Tinew} \| (r_1 C_1)_x \| R_{t1}^*)$

$|T_5 - T_4| < \Delta T$

$PID_{Rnew}^* = C_3 - h(ID_{Ti} \| T_3 \| T_4 \| R_{t1})$

$C_4^* = h((R_{t1})_x \| PID_{Rnew}^* \| (R_{r1})_x \| T_4)$

$C_4^* \overset{?}{=} C_4$

$Updates \;\; PID_{Tinew} \leftarrow h(PID_{Tinew} \| (R_{t1})_x)$

$Generates \;\; the \;\; shared \;\; session \;\; key$

$SK_{TR} = h(ID_{Ti} \| PID_{Tinew} \| (r_1 C_1)_x \| R_{t1})$

Figure 3: Authentication phase

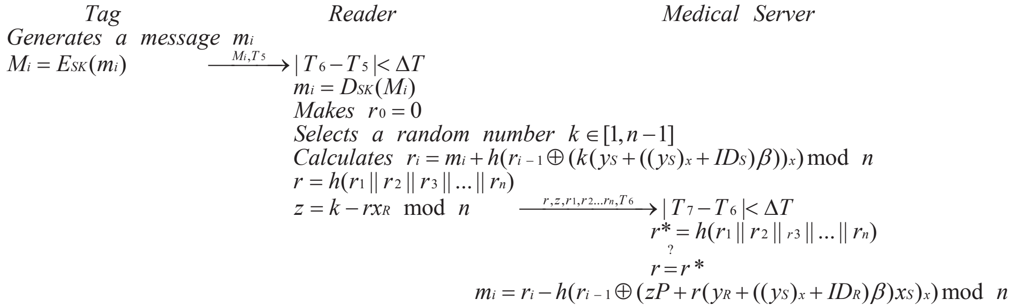

Tag
Reader
Medical Server
Generates a message m_i
$M_i = E_{SK}(m_i)$
M_i, T_5
$|T_6 - T_5| < \Delta T$
$m_i = D_{SK}(M_i)$
Makes $r_0 = 0$
Selects a random number $k \in [1, n-1]$
Calculates $r_i = m_i + h(r_{i-1} \oplus (k(y_S + ((y_S)_x + ID_S)\beta))_x) \bmod n$
$r = h(r_1 || r_2 || r_3 || \ldots || r_n)$
$z = k - rx_R \bmod n$
$r, z, r_1, r_2 \ldots r_n, T_6$
$|T_7 - T_6| < \Delta T$
$r^* = h(r_1 || r_2 || r_3 || \ldots || r_n)$
$r \overset{?}{=} r^*$
$m_i = r_i - h(r_{i-1} \oplus (zP + r(y_R + ((y_S)_x + ID_R)\beta)x_S)_x) \bmod n$

Figure 4: Digital signature and data transmission phase

5 Security Analysis

5.1 Informal Security Analysis

Anonymity

In the authentication phase, $PID_{Tinew}{=}h(PID_{Told}|| init||R_{t1})$, $C_2 = PID_{Tinew}{+}h((R_{t1})_x||(R_{r1})_x||(C_1)_x||T_1)$. Even if an attacker manipulates the value of $init$ and R_{r1}, the improved proposed protocol can offer tag anonymity. Because in the improved protocol we using computational Diffie-Hellman problem. According to computational Diffie-Hellman problem, we know that given $P, t_1P, r_1P, P \in E/F_q$, for unknown $t_1, r_1 \in Z_n^*$, the probability of success of finding the value of $R_{t_1} = t_1r_1P$ is negligible.

Provide Forward Secrecy

Let we assume that a legitimate tag called T_a and a legitimate R in this system. R executes several sessions with randomly selected a tag T_b in a time interval. A arbitrarily selects a tag and computes $PID_{Tbnew} = {}^aC_3 - h(ID_b||{}^aT_3||{}^aT_4||R_{t1})$, $PID_{Tbnew'} = h(PID_{Tiold}||init||R_{t1})$. Even if A know ${}^aC_3, {}^aT_3, {}^aT_4$, A cannot derive the correct PID_{Tbnew} without R_{t1}.

Resistance to Know Session-specific Temporary Information Attack

In the proposed protocol, the secret session key is calculated using the temporary secret parameters t_1, PID_{Tinew} and R_{t1}. Even if A obtains the short term identity PID_{Tinew} from insecure channel and come to guess the secret value t_1, he/she still cannot calculate R_{t1}. The proposed protocol is resilient to the know session-specific temporary information attack.

Resistance to Replay Attack

The timestamp and new random secret nouns are used in proposed protocol. Furthermore, the pseudo identities are updated for each session. Even if A logs in at the same time as the tag, he/she cannot compute $R_{t1} = t_1r_1P$. The proposed protocol can resist to replay attack.

Security Against Privileged Insider Attack

Assuming A is at the reader level. Even if A gets the tag's real identity, he/she cannot impersonate the tag. Because the shared session key $SK_{TR} = h(ID_{T_i}||PID_{Tinew}||(t_1R_{r_1})x||R_{t_1})$ is calculated using t_1,ID_i, PID_{Tinew} and R_{t1}. It is impossible for A to guess all parameter of the session. The proposed protocol can resist to privileged insider attack.

5.2 Formal Verification Using BAN-logic

Basic BAN-logic rules
Message meaning rule, seeing rule, jurisdiction rule, freshness rule, belief rule, and nonce verification rule of BAN-logic is used.
Protocol goals
Using the Ban-logic rules and the protocol assumptions, the following goals should be achieved.

$$G_1 : T| \equiv R \overset{SK}{\longleftrightarrow} T$$

$$G_2 : T| \equiv R| \equiv R \overset{SK}{\longleftrightarrow} T$$

$$G_3 : R \mid\equiv R \overset{SK}{\longleftrightarrow} T$$

$$G_4 : R| \equiv T| \equiv R \overset{SK}{\longleftrightarrow} T$$

$$G_5 : S \mid\equiv N_2$$

$$G_6 : R \mid\equiv m_i$$

Initial assumptions on the protocol
There are some logical assumptions on the entities participating in our protocol.

$$A1 : T| \equiv \#(T_{t),} T| \equiv \#(T_{r)}$$

$$A2 : T| \equiv \#(t1)$$

$$A_3 : R| \equiv T \Rightarrow C_1, R| \equiv T| \Rightarrow R_{t1}, R| \equiv T| \Rightarrow m_i$$

$$A_4 : T \equiv\xrightarrow{P_{uuR}} R, T \mid\xrightarrow{P_{us}} S$$

$$A_5 : T \mid\equiv T \overset{SK}{\longleftrightarrow} R$$

$$A_6 : R \equiv \#(T_t), R \mid\equiv \#(T_r), R| \equiv \#(T_s)$$

$$A_7 : R \mid\equiv \#(r_1)$$

$$A_8 : T| \equiv R| \Rightarrow R_{r1}, S \mid\equiv R \Rightarrow N_1$$

$$A_9 : R| \equiv\xrightarrow{P_uS} S, R \equiv T \overset{SK}{\longleftrightarrow} R$$

$$A_{10}: S \mid\equiv \#\left(T_r\right), R\mid\equiv \#\left(T_s\right)$$

$$A_{11}: S\mid\equiv \xrightarrow{P_{uR}} R$$

$$A_{12}: R\mid\equiv S\mid\Rightarrow R_{S1}, R\mid\equiv S\mid\Rightarrow S_1$$

Nothing that T_t, T_r, T_s are the current timestamp of the tag, the reader, and the server respectively.
Idealized model of the protocol
The idealized from of the communication messages between the tag, reader and server is described as followed.

$$\begin{aligned}
M_1: &\quad R \to T: \langle R_{r1}\rangle \\
M_2: &\quad T \to R: \langle C_1, C_2, T_1\rangle = \{C_1, \langle PIDTinew \\
&\qquad +\{R_{t1}\}P_{uR}, R_{r1}, C_1, T_1\rangle PID_{Tinew}, T_1\} \\
M_3: &\quad R \to S: \langle N_2, R_{r1}, T_2\rangle = \{\langle PID_{Rnew} + R_{r1}, \\
&\qquad ID_R, \{N_1\}P_{uS}, T_2\rangle PID_{Rnew}, ID_R, \\
&\qquad R_{r1}, T_2\} \\
M_4: &\quad S \to R: \langle S_1, N_3, T_3\rangle = \{S_1, \langle\{R_{s1}\}P_{uR}, PID_{Rnew}, \\
&\qquad T_2, T_3 + ID_S\rangle ID_S, PID_{Rnew}, T_3\} \\
M_5: &\quad R \to T: \langle C_3, C_4, T_3, T_4\rangle = \{\langle ID_{Ti}, T_3, T_4, \\
&\qquad R^*_{t1} + PID_{Rnew}\rangle ID_{Ti}, PID_{Rnew}, \\
&\qquad \langle\{R_{t1}\}P_{uR}, PID_{Rnew}, R_{r1}, T_4\rangle PID_{Rnew}, \\
&\qquad T_3, T_4\} \\
M_6: &\quad T \to R: \langle\{m_i\}\text{SK}, T5\rangle \\
M_7: &\quad R \to S: \langle r, z, r_i, T_6\rangle
\end{aligned}$$

Verification process
From M_2, we have

$$S_1: R \triangleleft \{C_1, \langle PID_{Tinew} + \{R_{t1}\}P_{uR}, R_{r1}, C_1, T_1\rangle \\ PID_{Tinew}, T_1\}$$

According to seeing rule, message meaning rule and A_4, we have

$$S_2: R\mid\equiv T \sim \langle PID_{Tinew} + \{R_{t1}\}P_{uR}, R_{r1}, C_1, \\ T_1\rangle PID_{Tinew}$$

According to A_6, S_2, freshness rule and nonce verification rule, we have

$$S_3: R\mid\equiv T\mid\equiv \langle PID_{Tinew} + \{R_{t1}\}P_{uR}, R_{r1}, C_1, \\ T_1\rangle PID_{Tinew}$$

According to S_3, A_3 and jurisdiction, we have

$$S_4: R\mid\equiv \langle PID_{Tinew} + \{R_{t1}\}P_{uR}, R_{r1}, C_1, T_1\rangle PID_{Tinew}$$

According to M_3, we have

$$S_5: S \triangleleft \{\langle PID_{Rnew} + R_{r1}, ID_R, \{N_1\}P_{uS}, \\ T_2\rangle PID_{Rnew}, ID_R, R_{r1}, T_2\}$$

According to message meaning rule, seeing rule and A_9, we have

$$S_6: S\mid\equiv R \sim \langle PID_{Rnew} + R_{r1}, ID_R, \{N_1\}P_{uS}, \\ T_2\rangle PID_{Rnew}, ID_R$$

According to S_6, freshness rule, nonce verification rule and A_{10}, we have

$$S_7: S\mid\equiv R\mid\equiv \langle PID_{Rnew} + R_{r1}, ID_R, \{N_1\}P_{uS}, \\ T_2\rangle PID_{Rnew}, ID_R$$

According to S_7, jurisdiction rule and A_8, we have

$$\begin{aligned}
S_8: S &\mid\equiv \langle PID_{Rnew} + R_{r1}, ID_R, \{N_1\}P_{uS}, \\
&\qquad T_2\rangle PID_{Rnew}, ID_R \\
S_9: S &\mid\equiv N_2 (Goal5)
\end{aligned}$$

From M_4, we have

$$S_{10}: R \triangleleft \{S_1, \langle\{R_{s1}\}P_{uR}, PID_{Rnew}, T_2, T_3 \\ +ID_S\rangle PID_{Rnew}, ID_S, T_3\}$$

According to message meaning rule, A_9 and A_{11}, we have

$$S_{11}: R\mid\equiv S\mid\sim \langle\{R_{s1}\}P_{uR}, PID_{Rnew}, T_2, T_3 \\ +ID_S\rangle PID_{Rnew}, ID_S$$

According to S_{11}, A_6, nonce verification rule and freshness rule, we have

$$S_{12}: R\mid\equiv S\mid\equiv \langle\{R_{s1}\}P_{uR}, PID_{Rnew}, T_2, T_3 \\ +ID_S\rangle PID_{Rnew}, ID_S$$

According to S_{12}, A_{12}, belief rule and jurisdiction rule, we have

$$S_{13}: R\mid\equiv \langle\{R_{s1}\}P_{uR}, PID_{Rnew}, T_2, T_3 \\ +ID_S\rangle PID_{Rnew}, ID_S$$

The reader generates a session key SK_{RT} using ID_{Ti}, PID_{Tinew}, its random nonce r_1, and the tag public parameter C_1, we have

$$S_{14}: R \mid\equiv R \overset{SK}{\longleftrightarrow} T \quad \text{(Goal 3)}$$

The reader sends M_5 to the tag, hence, we have

$$S_{15}: R \equiv T \equiv R \overset{SK}{\longleftrightarrow} T \quad \text{(Goal 4)}$$

According to M_5, we have

$$S_{16}: T \triangleleft \{\langle ID_{Ti}, T_3, T_4, R^*_{t1} + PID_{Rnew}\rangle ID_{Ti}, PID_{Rnew}, \\ \langle\{R_{t1}\}P_{uR}, PID_{Rnew}, R_{r1}, T_4\rangle PID_{Rnew}, T_3, T_4\}$$

According to message rule, S_{16} and A_4, we have

$$S_{17}: T\mid\equiv R \sim \{\langle ID_{Ti}, T_3, T_4, R^*_{t1} + PID_{Rnew}\rangle ID_{Ti}, \\ PID_{Rnew}, \langle\{R_{t1}\}P_{uR}, PID_{Rnew}, R_{r1}, \\ T_4\rangle PID_{Rnew}\}$$

According to S_{17}, freshness rule, nonce verification rule and A_1, we have

$$S_{18}: T\mid\equiv R\mid\equiv \{\langle ID_{Ti}, T_3, T_4, R^*_{t1} + PID_{Rnew}\rangle ID_{Ti}, \\ PID_{Rnew}, \langle\{Rt1\}P_{uR}, PID_{Rnew}, R_{r1}, \\ T_4\rangle PID_{Rnew}\}$$

According to S_{18}, jurisdiction rule, A_8 and $T| \equiv R_{t1}$, we have

$$\begin{aligned} S_{19}: T \quad &| \equiv \quad \{\langle ID_{Ti}, T_3, T_4, R^*_{t1} + PID_{Rnew}\rangle ID_{Ti}, \\ &\qquad PID_{Rnew}, \langle\{R_{t1}\}P_{uR}, PID_{Rnew}, R_{r1}, \\ &\qquad T_4\rangle PID_{Rnew}\} \\ T \quad &\equiv \quad \{C_3, C_4\} \\ T \quad &\equiv \quad R \equiv\xleftrightarrow{SK} T \qquad \text{(Goal 2)} \\ T \quad &\equiv \quad R \xleftrightarrow{SK} \qquad \text{(Goal 1)} \end{aligned}$$

According to M_6, we have

$$S_{20}: R \triangleleft \{\{m_i\}\mathrm{SK}, \mathrm{T}_5\}$$

According to message meaning rule, A_5 and A_9, we have

$$S_{21}: R| \equiv T| \sim m_i$$

According to S_{21}, freshness rule and A_6, we have

$$S_{22}: R| \equiv T| \equiv m_i$$

According to S_{22}, jurisdiction rule and A_3, we have

$$S_{23}: R| \equiv m_i \qquad \text{(Goal 7)}$$

6 Comparison of Security Features and Efficiency Characteristics

To evaluate the computational time analysis, we account the time complexity of a one-way hash functionT_h ≈0.0023ms, the time complexity of a point multiplication operation on elliptic curveT_{mul} ≈2.226ms, and the time complexity of encryption or decryption funtionT_{fun} ≈0.0046ms as reported in [14] and [23]. According to the Table 2 and Table 3, our protocol provides more security features with the addition of a small amount of computation.

7 Conclusions

Based on the architecture of Izza *et al.*'s agreement, we propose an improved secure RFID authentication protocol using elliptic curve cryptography. Our improved protocol uses the knowledge of elliptic curve cryptography, which is an algorithm for establishing public key encryption. Public key algorithms are always based on a mathematical puzzle. We construct our protocol by using some recognized mathematical puzzle within elliptic curves. Its main advantage is that in some cases it provides equivalent or higher security than other methods using smaller keys. BAN logic is used to prove the improved protocol. Meanwhile, we carry out comparative analysis of performance and efficiency. The results show that the improved protocol has higher security and lower calculation cost.

Acknowledgments

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

References

[1] S. Abughazalah, K. Markantonakis, K. Mayes, "Secure improved cloud-based RFID authentication protocol," *in: Data Privacy Management, Autonomous Spontaneous Security, and Security Assurance*, pp. 47–164, 2015.

[2] A. A. Alamr, F. Kausar, J. Kim, *et al.*, "A secure ECC-based RFID mutual authentication protocol for Internet of Things," *Journal of supercomputing*, vol. 74, no. 9, pp. 4281-4294, 2018.

[3] A. Arslan, M. A. Bingöl, "Cryptanalysis of Izza *et al.*.'s Protocol: An enhanced scalable and secure RFID authentication protocol for WBAN within an IoT environment," *IACR Cryptology ePrint Archive*, 2021.

[4] B. Chander, K. Gopalakrishnan, "A secured and lightweight RFID-tag based authentication protocol with privacy-preserving in Telecare medicine information system," *Computer Communication*, vol. 191, no. 2022, pp. 425-437, 2022.

[5] Y. C. Chen, W. L. Wang, M. S. Hwang, "RFID authentication protocol for anti-counterfeiting and privacy protection", in *The 9th International Conference on Advanced Communication Technology*, pp. 255-259, 2007.

[6] D. Dharminder, D. Mishra, X. Li, "Construction of RSA-based authentication scheme in authorized access to healthcare services, "*Journal of Medical System*, vol. 44, no. 1, pp.1-9, 2020.

[7] N. Dinarvand, H. Barati, "An efficient and secure RFID authentication protocol using elliptic curve cryptography," *Wireless Networks*, vol. 25, pp. 415–428, 2019.

[8] K. Fan, Q. Luo, K. Zhang, *et al.*, "Cloud-based lightweight secure RFID mutual authentication protocol in IoT," *Information Sciences*, vol. 527, pp. 329–340, 2020.

[9] S. Gabsi, V. Beroulle, Y. Kieffer, *et al.*, "Survey: Vulnerability Analysis of Low-Cost ECC-Based RFID Protocols against Wireless and Side-Channel Attacks," *Sensors*, vol. 21, no. 17, 2021.

[10] S. J. Hussain, M. Irfan, N. Z. Jhanjhi, *et al.*, "Performance enhancement in wireless body area networks with secure communication, "*Wireless Personal Comminocations*, vol. 116, no.1, pp. 1-22, 2021.

[11] M. S. Hwang, C. H. Wei, C. Y. Lee, "Privacy and security requirements for RFID applications", *Journal of Computers*, vol. 20, no. 3, pp. 55-60, Oct. 2009.

[12] S. Izza, M. Benssalah, K. Drouiche, "An enhanced scalable and secure RFID authentication protocol for WBAN within an IoT environment," *Journal of*

Table 2: Comparison of security features

Performance	Renuka	Madhusudhan *et al.*	Dharminder *et al.*	Izza *et al.*	Hussain *et al.*	Chander *et al.*	Our protocol
	et al. [24]	*et al.* [19]	*et al.* [6]	*et al.* [12]	*et al.* [10]	*et al.* [4]	
F1	No	No	No	Yes	Yes	Yes	Yes
F2	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F3	No	Yes	Yes	No	Yes	No	Yes
F4	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F5	Yes	No	Yes	Yes	Yes	No	Yes
F6	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F7	Yes	Yes	Yes	No	Yes	Yes	Yes
F8	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F9	Yes	Yes	Yes	No	Yes	Yes	Yes
F10	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F11	Yes	No	Yes	Yes	No	Yes	Yes

F1: Resist against impersonation attack; F2: Resist forgery attack; F3: Anonymity; F4: Defend known session-specific temporary information attack; F5: Resist man-in-the-middle attack; F6: Resist the tag's identity reveal; F7: Clock synchronization mechanism; F8: Resist offline guessing attack; F9: Resist forward secrecy; F10: Session key agreement; F11: Resist replay attack.

Table 3: Comparisons in terms of efficiency

Protocols	Computation time
Renuka *et al.* [24]	0.0437ms
Madhusudhan *et al.* [19]	0.046ms
Dharminder*et al.* [6]	0.0322ms
Izza *et al.* [12]	22.2646ms
Hussain *et al.* [10]	0.0874ms
Chander *et al.* [4]	0.0529ms
Our protocol	24.9106ms

Information Security and Applications, vol. 58, pp. 102705.1-102705.15, 2021.

[13] Q. Jiang, N. Zhang, J. Ni, *et al.*, "Unified biometric privacy preserving three-factor authentication and key agreement for cloud-assisted autonomous vehicles," *IEEE Transactions on Vehicular Technology*, vol. 69, no.9, 2020.

[14] H. H. Kilinc and T. Yanik, "A survey of SIP authentication and key agreement schemes," *IEEE Communication Surveys and Tutorials*, vol. 16, no. 2, pp. 1005–1023, 2014.

[15] V. Kumar, M. Ahmad, D. Mishra, *et al.*, "RSEAP: RFID based secure and efficient authentication protocol for vehicular cloud computing," *Vehicular Communications*, vol. 22, pp. 100213.1-100213.13, 2020.

[16] C. T. Li, M. S. Hwang, "A lightweight anonymous routing protocol without public key en/decryptions for wireless ad hoc networks",*Information Sciences*, vol. 181, no. 23, pp. 5333-5347, 2011.

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

[18] W. R. Liu, B. Li, Z. Y. Ji, "An improved three-factor remote user authentication protocol using elliptic curve cryptography, "*International Journal of Network Security*, vol. 24, no. 3, pp.521-532, 2021.

[19] R. Madhusudhan, C.S. Nayak, "A robust authentication scheme for telecare medical information systems," *Multimedia Tools and Applications*, vol. 78, no.11, pp15255-15273, 2019.

[20] D. Mishra, V. Kumar, D. Dharminder, *et al.*, "SFVCC: chaotic map-based security framework for vehicular cloud computing," *IET lntelligent Transport Systems*, vol. 14, no.4, pp. 241-249, 2020.

[21] M. Naeem, S. A. Chaudhry, K. Mahmood, *et al.*, "A scalable and secure RFID mutual authentication protocol using ECC for Internet of Things," *Inernational Jornal of Communication Systems*, vol. 33, no. 13, pp. e3906.1-e3906.13, 2020.

[22] M. Nagarajan, M. Rajappa, "Simple Yet Secure Encoder Architecture and Ultralightweight Mutual Authentication Protocol for RFID Tags in IoT," *Journal of Circuits Systems and Computers*, vol. 32, no. 07, 2350118, 2023.

[23] A. Ostad-Sharif, A. Babamohammadi, D. Abbasinezhad-Mood, *et al.*, "Efficient privacy-preserving authentication scheme for roaming consumer in global mobility networks," *International Journal of Communication Systems*, vol. 32, no. 5, pp. e3904, 2019.

[24] K. Renuka, S. Kumari, X. Li, "Design of a secure three-factor authentication scheme for smart healthcare, " *Journal of Medical Systems*, vol. 43, no.5, 2019.

[25] M. Safkhani, C. Camara, P. Peris-Lopez, *et al.*, "RSEAP2: An enhanced version of RSEAP, an RFID-based authentication protocol for vehicular cloud computing," *Vehicular Communications*, vol. 28, 2021.

[26] C. H. Wei, M. S. Hwang, Augustin Y. H. Chin, "A secure privacy and authentication protocol for passive RFID tags," *International Journal of Mobile Communications*, vol. 15, no. 3, pp. 266–277, 2017.

[27] C. H. Wei, M. S. Hwang, Augustin Y. H. Chin, "An authentication protocol for low-cost RFID tags", *In-*

ternational Journal of Mobile Communications*, vol. 9, no. 2, pp. 208–223, 2011.

[28] C. H. Wei, M. S. Hwang, Augustin Y. H. Chin, "An improved authentication protocol for mobile agent device in RFID," *International Journal of Mobile Communications*, vol. 10, no. 5, pp. 508–520, 2012.

[29] C. H. Wei, M. S. Hwang, Augustin Y. H. Chin, "Security analysis of an enhanced mobile agent device for RFID privacy protection," *IETE Technical Review*, vol. 32, no. 3, pp. 183–187, 2015.

Biography

Liu Wanrong received her master's degree from Shanghai Ocean University in 2021. At present, she has worked in Shanghai Sixth People's Hospital Affiliated to Shanghai Jiao Tong University School of Medicine. Her main research is communication security and Internet of things technology.

Ji Zhiyong received his bachelor's degree from Nanjing University of Aeronautics and Astronautics in 2012. He received his MS degree Jiangsu University in 2017. He is the master's supervisor of mechanical engineering of Shanghai Ocean University. He is also the medical equipment senior engineer and deputy director of Shanghai Sixth People's Hospital Affiliated to Shanghai Jiao Tong University School of Medicine. His research directions include the development and application of wearable medical devices based on the Internet of things and the information security of the medical Internet of things.

Chu Chengchen received the M.S. degree in biomedical engineering from Shanghai Jiao Tong University (SJTU), China. Since 2017, he has worked in the Medical Equipment Department of the Sixth People's Hospital affiliated to Shanghai Jiao Tong University School of Medicine. His research interests include medical equipment reliability research and the information security of the medical Internet of things.

Parameter Obfuscation and Restoration for Secure Federated Learning Aggregation

Xiangxiang Ma[1,2], Linming Gong[1,2], Jian Chen[1,2], and Daoshun Wang[3]
(Corresponding author: Xiangxiang Ma)

The Shaanxi Key Laboratory of Clothing Intelligence School of Computer Science, Xi'an Polytechnic University[1]
State and Local Joint Engineering Research Center for Advanced Networking and Intelligent Information Services,
School of Computer Science, Xi'an Polytechnic University[2]
Xi'an 710048, China
Department of Computer Science and Technology, Tsinghua University[3]
Beijing 100084, China
Email: maxiangxiang2023@163.com

(Received Mar. 8, 2023; Revised and Accepted Nov. 22, 2023; First Online Dec. 24, 2023)

Abstract

Federated Learning (FL) is a privacy-preserving solution that addresses the issue of data security. It achieves model aggregation by uploading only the locally trained parameters of each party to a central server, thus preserving the original privacy of each entity. However, recent studies indicate that the global parameters aggregated by the server and the local parameters submitted by users may inadvertently reveal private data. In this paper, we propose a secure FL aggregation protocol based on parameter confusion and restoration. Our protocol ensures the protection of local parameters uploaded by a single client by uploading only blurred parameters to the server. The global model aggregated by the server remains unaffected by the interference data, ensuring the high accuracy of the final model. Furthermore, the aggregation server remains unaware of any parameter information during training. We demonstrate the security of our protocol against a semi-honest adversary environment and prove that even if n-2 participants collude with the server, it is impossible to learn the private information of the honest participants. Our tests on the MNIST and cifar10 datasets show that our system achieves accuracy comparable to the plaintext baseline.

Keywords: Federated Learning; Parameter Confusion; Privacy; Restoration; Secure Aggregation

1 Introduction

With the introduction of cloud computing [5,8,18,19] and big data [11] technologies, the performance of machine learning algorithms has improved significantly. In many fields [12,32], the accuracy of a machine learning model trained on a significant amount of real data, especially a deep learning model [9], approaches or exceeds that of a human. Enterprises and organizations collect a significant amount of richer and more sensitive user data to build better models.

However, the collection of large amounts of data inevitably leads to privacy issues. Consider the Facebook data leak in April 2021. Members of a hacker community leaked hundreds of millions of Facebook user records online, including phone numbers and other private information. Centralizing data storage increases the risk of unwarranted security and privacy issues.

Federated learning [3,17,33,35], as a distributed machine learning paradigm, does not require that all data sets be brought together to train the model. Instead, each party trains the model on the local dataset. After each round of training, the server aggregates the local model parameters, and returns the updated global parameters to each party to continue iterative training until the model converges. Each client can keep its data set local, which reduces the risk of privacy leakage caused by traditional centralized data model training.

The work [21,22] provides federated averaging algorithm. In federated averaging, the training data resides on the clients, never uploaded to the server. The client uploads the parameters to the aggregation server after local training, and the server aggregates the received parameters and sends the updated model parameters to the client. This process repeats until the model converges. Local data is maintained locally, eliminating direct leakage. However, a growing number of studies have shown that these shared model parameters can also leak private data information of parties. For example, in gradient inversion [14] or member inference attacks [24,29], model parameter information can lead to indirect leakage of client's local data [25,30]. To target this problem, the work [31] provides mechanisms to split the model, where the user trains the sensitive data with part of the model and uploads the trained non-sensitive data to the server, which uses the other part of the model to continue the training and aggregate the models. This method does not need all the user data, so it has certain privacy. However, training with only part of the data will reduce the accuracy of the model. At the same time, due to the strong correlation between the gradient and the data [25], the adversary can reverse the gradient [13] to obtain the user's private information. To solve these problems, the work of [28] proposes a privacy-preserving deep learning system. The client can optionally use part of the local dataset for training and upload a portion of the trained gradient to the server for aggregation. But even the partial gradient information can be used by the adversary to extract some relevant information about the client [2]. In addition, the accuracy of the model can be reduced by sharing only part of the parameters. To avoid sharing raw parameters, the work of [1] proposes to apply differential privacy techniques to avoid the leakage of plaintext gradients by adding noise to the gradients. The work of [4, 23] also combines differential privacy with the federated averaging algorithm to achieve a trade-off between privacy and training accuracy. Although differential privacy excludes the effects of inversion and inference attacks, the adversary can still learn private data using GAN (Generative Adversarial Networks) [15]. And this method will affect the accuracy of the model. In addition, as an encryption algorithm that satisfies the homomorphic property of ciphertext, homomorphic encryption also provides a feasible solution for protecting the client's gradient information. the works of [2, 27, 34, 36] use this technique to ensure the security of the model parameters at the client. However, there can be a high performance overhead due to the need to encrypt a large number of model parameters in each round, and these schemes often come with additional trust assumptions (e.g., sharing a private key among data owners). The work of [7] proposes a privacy-preserving protocol for federated learning that combines multiple cryptographic techniques. The scheme has strong security, even if the malicious server colludes with the malicious client, the honest user's private information is not leaked. However, this scheme is vulnerable to membership inference attacks [26], and it is insecure. Table 1 summarizes the comparison with most related works discussed above.

Focusing on the above problems, we propose a secure FL aggregation protocol based on parameter confusion and restoration. Our contributions are as follows.

1) We introduce a joint deep learning training framework that supports the parallel training process to achieve strong privacy and high model accuracy;

2) Both the private local original data set and the local individual model parameters of each honest party can be protected;

3) The proposed security protocol can resist collusion attack (even if n-2 of n parties collude with the server);

4) No global parameters are leaked to the aggregation server.

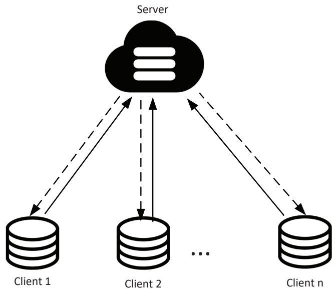


Figure 1: Example of federated learning

2 Preliminaries

In this section, we briefly introduce some of the technical backgrounds involved in our scheme.

2.1 Federated Learning

Federated learning [6, 16, 35] refers to the setting where multiple customers (such as mobile devices, institutions, organizations, etc.) collaborate [21, 22] to perform decentralized machine learning under the collaboration of one or more central servers [20]. In this approach, each user and server collaborate to train a unified neural network model. Each party trains its own individual model locally and then sends the model to some intermediate server party for aggregation of the training results. The server then returns the aggregated results to each user who uses the aggregated results for a new round of training. Ultimately, the server and each user will get the optimal network parameters, as shown in Figure 1. However, research shows [13, 29] that attackers can still indirectly obtain the sensitive information based on shared parameters and server aggregation results. Therefore, in this paper, we focus on protecting the privacy of users' local parameters and server aggregation results.

2.2 Paillier Encryption

Paillier encryption [10] is an additive homomorphic encryption based on the composite power residue class problem, which has been widely used in data processing. Generally, it consists of the following poly-time algorithms.

- **KeyGen**$(*) \rightarrow (pk, sk)$. Pick two independent large prime numbers p and q at random and satisfy $gcd(pq, (p-1)(q-1)) = 1$, calculate $n = pq$, $\lambda = lcm(p-1, q-1)$, choose a uniform number

Table 1: Comparison of federated learning systems with secure aggregation

	Features			
Proposed Approach	accuracy loss	Local parameter privacy	Global parameter privacy	Anti-collusion attack
[21,22]	low	×	×	×
[28]	high	×	×	×
[1]	high	√	√	√
[2]	low	√	√	×
[7]	low	√	×	√
[31]	high	×	×	×
Oursystem	low	√	√	√

$g \in \mathbb{Z}_{n^2}^*$. The public key is $pk = (n, g)$, and the private key is $sk = (\lambda)$.

- **Encryption**$(pk, m) \to c$. on input public key pk and a plaintext $m \in \mathbb{Z}_n$, choose a uniform number $r \in \mathbb{Z}_n^*$ and output the ciphertext:

$$c = g^m \cdot r^n \bmod n^2 \tag{1}$$

- **Decryption**$(sk, c) \to m$. on input a private key sk and a ciphertext c, compute:

$$m = L(c^\lambda \bmod n^2) \cdot \mu \bmod n \tag{2}$$

where $L(x) = \frac{x-1}{n}$ and $\mu = (L(g^\lambda \bmod n^2))^{-1} \bmod n$.

- **Additive homomorphic**. The ciphertext of plaintext m is defined as $E(m)$. And the ciphertexts $E(m_1 + m_2)$ satisfy:

$$E(m_1 + m_2) = E(m_1) \cdot E(m_2) \tag{3}$$

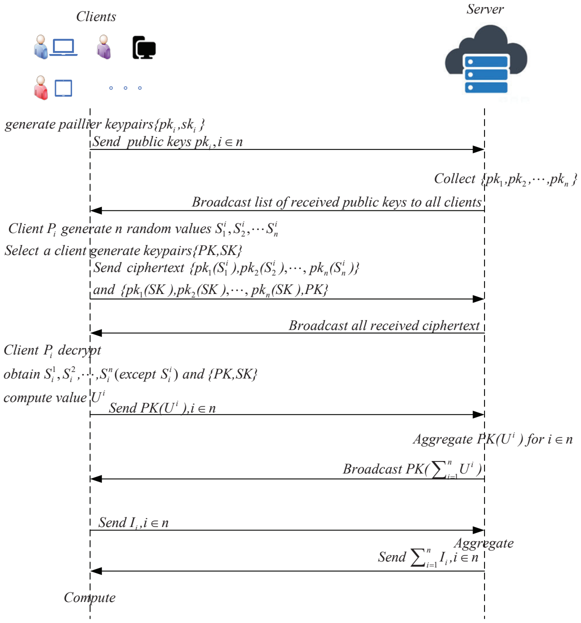

Clients
Server
generate paillier keypairs$\{pk_i, sk_i\}$
Send public keys $pk_i, i \in n$
Collect $\{pk_1, pk_2, \cdots, pk_n\}$
Broadcast list of received public keys to all clients
Client P_i generate n random values $S_1^i, S_2^i, \cdots S_n^i$
Select a client generate keypairs{PK,SK}
Send ciphertext $\{pk_1(S_1^i), pk_2(S_2^i), \cdots, pk_n(S_n^i)\}$
and $\{pk_1(SK), pk_2(SK), \cdots, pk_n(SK), PK\}$
Broadcast all received ciphertext
Client P_i decrypt
obtain $S_i^1, S_i^2, \cdots, S_i^n$(except S_i^i) and {PK,SK}
compute value U^i
Send $PK(U^i), i \in n$
Aggregate $PK(U^i)$ for $i \in n$
Broadcast $PK(\sum_{i=1}^n U^i)$
Send $I_i, i \in n$
Aggregate
Send $\sum_{i=1}^n I_i, i \in n$
Compute

Figure 2: The Overall Process of system

3 System Overview

3.1 Architecture

There are N decentralized participants $P = \{P_1, P_2, \cdots, P_n\}$, each owning a local private dataset $D_1, D_2, \cdots, D_n$, and a server S responsible for aggregating the data. Each participant establishes a TLS/SSL secure channel, different from the others, to communicate and protect the integrity of the homomorphic ciphertexts. The goal of the protocol is to jointly train a global model without exposing the client's private data set. The protocol uses local parameters W_i to represent the local model, G_i to represent the model gradient, and builds the global model W by sharing parameters with other parties. In our system, the entire process is shown in Figure 2.

Traditional federated learning algorithms use intermediate information, such as model parameters, to replace the original data to be transmitted between nodes. However, these intermediate information often contain the information of the original data, and the adversary can infer the original data through them. The privacy leakage targeted in this paper is mainly divided into local privacy leakage and global privacy leakage, which details are as follows.

- **Local privacy leakage.** In each training round, the client needs to upload the trained local model parameters which contain the original data characteristics. The adversary can obtain the original data information from the local parameters through some attack methods (such as reconstruction attack). To solve these problems, we obscure the local parameters uploaded by each client. The server does not need to know the specific local parameter information of each client. At the same time, the protocol ensures that the server can aggregate the uploaded parameters to obtain the correct aggregation results, so as to avoid local information leakage.

- **Global privacy leakage.** In addition to the local information uploaded by the client that can leak private data, research shows [29] that the adversary can also infer some data information from the global information, such as the aggregation result or the final output model of the server. To solve these problems, we use the homomorphic encryption technique to encrypt some parameters uploaded by the client to protect the global information and prevent the aggregated parameters from being exploited by the adversary.

3.2 Threat Model

Our secure federated learning system assumes that the server and all clients are honest but curious (also known as semi-honest or passive). They honestly follow the secure federated learning aggregation protocol, but may try to infer other clients' private data from the known information. In the process of model training, collusion between the server and the client may occur. For this case, our protocol is also able to guarantee the privacy of honest clients. We do not consider the extreme case where there is only one honest client (n-1 clients collude with the server to try to infer the private data of the only honest party). It is clear that for any practical aggregation scheme that securely computes the sum, if the aggregator merges with all but one client, the data of the merged client can be certainly removed from the sum, and the data of the honest client will be revealed. Therefore, our system does not provide data protection against extreme cases.

4 The Proposed FL Aggregation Solution

In this section, we describe our scheme and security analysis in detail.

4.1 Detailed Execution of Our Scheme

The detailed execution consists of three phases.as shown in Algorithm 1, proceeds as follows.

- **System initialization phase.** Each client used pailliar homomorphic encryption technology to generate a set of its own public and private keys $\{pk_n, sk_n\}$, and uploaded the public keys to the aggregation server S. The server collected all the public keys it received and sent the public key set $\{pk_1, pk_2, \cdots, pk_n\}$ to all clients in the form of broadcast. Each client P_i generates n random values $S_1^i, S_2^i, \cdots, S_n^i$ whose sum is $S^i = S_1^i + S_2^i + \cdots + S_n^i$, and encrypts these random values using the public key in the public key set $\{pk_1, pk_2, \cdots, pk_n\}$ to get $\{pk_1(S_1^i), pk_2(S_2^i), \cdots, pk_n(S_n^i)\}$, and upload it to the aggregation server. The server collates all encrypted data sets received and sends them to all clients in broadcast form. After receiving all encrypted datasets, client i uses its own private key sk_i to decrypt the ciphertext set $\{pk_i(S_i^1), pk_i(S_i^2), \cdots, pk_i(S_i^n)\}$(except $pk_i(S_i^i)$) encrypted by other clients using public key pk_i. As a result of this decryption process, client i successfully obtains the data $S_i^1, S_i^2, \cdots, S_i^n$(except S_i^i), after which it is calculated to obtain $U^i = \sum_{k=1}^n S_i^k$. At the same time, we randomly select a client to generate a set of public and private keys $\{PK, SK\}$ using pailliar and encrypt it with the public key in the public key set and upload it to the server. Next, the server sends it to all clients in the form of broadcast. After receiving it, clients decrypt it to obtain a set of shared key pairs $\{PK, SK\}$.

Algorithm 1 Our Design for Secure Federated Learning

```
Input: N parties $P = \{P_1, P_2, \cdots, P_n\}$; Each party has private key $sk_i$, correspond public key $pk_i$;All parties share a pair of keys $\{PK, SK\}$.
Output: A global model $W_{global}$
Initialize $W$.
for each party $P_i$ do
    Upload $pk_i$ to the cloud server.
    Random $S_1^i, S_2^i, \cdots, S_n^i$
    Compute $S^i = S_1^i + S_2^i + \cdots + S_n^i$
    for each $j \in N, j \neq i$ do
        Download $pk_j$ from the cloud server.
        Encrypt $pk_j(S_j^i)$ and Upload to the cloud server.

        Download $pk_i(S_i^j)$ from the cloud server.
        $S_i^j \leftarrow pk_i(S_i^j)$
    end for
    Compute $U^i = \sum_{k=1}^n S_i^k$
end for
for each round $t = 1, 2, \cdots$ do
    for each party $P_i$ do
        $W_t^i \leftarrow ClientUpdate$(i,$W_{t-1}$)
    end for
    $W_t \leftarrow \frac{1}{|N|}(\sum_{i=1}^n W_t^i - \sum_{i=1}^n U^i)$
end for
$W \leftarrow W_t$
return W
ClientUpdate($i, W$):
Split $P_i$ dataset into batches $\mathfrak{B}$.
for each local epoch $e$ from 1 to $E$ do
    for each batch $\beta$ in $\mathfrak{B}$ do
        $W \leftarrow W - \eta \cdot \bigtriangledown L(W;\beta)$. // $\eta$ is the learning rate.
    end for
end for
$W \leftarrow W + S^i$
return W
```

- **Interference value sum phase.** Each client encrypts U^i using the public key PK and uplodes $PK(U^i)$ to the server. Due to the additive homomorphism of pailliar, the server aggregates the uploaded values and can obtain $PK(\sum_{i=1}^n U^i)$. Clients downloads $PK(\sum_{i=1}^{n} U^i)$ from the server and decrypts it to get $\sum_{i=1}^{n} U^i$.

- **Secure summation phase.** In this phase, the data uploaded by the client is disturbed to prevent the adversary from stealing the uploaded local and global parameters. Details are as follows.

  1) In the initial phase, a client is randomly selected to generate the initial parameters W_0 and send its encrypted value $PK(W_0)$ to the server, and the remaining parties download the encrypted value from the server and decrypt to obtain the initial parameters;

  2) Party P_i trains the model locally and gets the trained gradient G_i, then multiplies the gradient by the learning rate α plus the disturbance S^i to get $I_i = \alpha \cdot G_i + S^i$;

  3) Each party sends I_i to the server separately. The server sums up all the parameters received to get $I = \sum_{i=1}^{n}(\alpha \cdot G_i + S^i)$. Finally, the participant downloads the updated parameter I from the server and subtracts the sum of the initial random values $\sum_{i=1}^{n} U^i$ to obtain $R = \sum_{i=1}^{n}(\alpha \cdot G_i + S^i) - \sum_{i=1}^{n} U^i = \sum_{i=1}^{n} \alpha \cdot G_i$. Then, the global parameters are updated to obtain a new round of model parameters $W_{new} = W_{old} - \sum_{i=1}^{n} \alpha \cdot G_i$, and perform a new round of deep learning model training until the model converges.

Our scheme achieves a high privacy security without losing the accuracy of the model, which not only ensures the privacy security of local parameters and global parameters in the process of model training, but also does not leak the private data of honest parties even if n-2 of the n parties collude with the server.

4.2 Correctness Analysis

Theorem 1. *When all parties are semi-honest, the final output R obtained by this protocol is the sum of the product of the learning rate α and the gradient G_i of all parties.*

Proof. For each $i \in \{1, 2, \cdots, n\}$, party i has a pair of values $\{S^i, U^i\}$ and $S^i = S_1^i + S_2^i + \cdots + S_n^i$, $U^i = S_i^1 + S_i^2 + \cdots + S_i^n$. So that, we have:

$$\begin{aligned}\sum_{i=1}^{n} S^i &= \sum_{i=1}^{n}(S_1^i + S_2^i + \cdots + S_n^i) \\ &= \sum_{i=1}^{n} S_1^i + \sum_{i=1}^{n} S_2^i + \cdots + \sum_{i=1}^{n} S_n^i \\ &= U^1 + U^2 + \cdots + U^n \\ &= \sum_{i=1}^{n} U^i\end{aligned}$$

So that, it is easy to derive the following equation.

$$\begin{aligned}R &= \sum_{i=1}^{n}(\alpha \cdot G_i + S^i) - \sum_{i=1}^{n} U^i \\ &= \sum_{i=1}^{n} \alpha \cdot G_i + \sum_{i=1}^{n} S^i - \sum_{i=1}^{n} U^i \\ &= \sum_{i=1}^{n} \alpha \cdot G_i + \sum_{i=1}^{n}\sum_{k=1}^{n} S_k^i - \sum_{i=1}^{n}\sum_{k=1}^{n} S_i^k \\ &= \sum_{i=1}^{n} \alpha \cdot G_i\end{aligned}$$

Finally, the output R is the sum of the product of the learning rate α and the gradient G_i of all participants, so the correctness is proved. □

4.3 Security Analysis

This paper assumes that servers and participants are semi-honest, i.e., they faithfully execute the algorithm and the protocol process, but they may retain all intermediate results and try to derive information beyond the result from them.

There are two possible leaks in the protocol: the first is the honest party gradient leak, caused by the semi-honest party colluding with the server; the second is the global parameter leak, caused by the server obtaining the decryption key. The security analysis of the two cases is as follows:

Theorem 2. *If the server colludes with clients, even if n-2 clients are corrupt, the adversary cannot infer any training data or model updates for the parties involved.*

Proof. Suppose that $\{C_1, C_2\}$ is the honest client and $\{C_3, C_4, \cdots, C_n\}$ is the set of adversaries composed of clients. Since the server is corrupted, the adversary knows parameters $I_1 = \alpha \cdot G_1 + S^1, U^1$ uploaded by the honest server C_1. However, the adversary cannot infer the values of S_1^1 and S_1^2 from $U^1 - S_1^3 - S_1^4 - \cdots - S_1^n = S_1^1 + S_1^2$, so the value of S^1 cannot be obtained. The gradient information of the honest party will not be stolen by the adversary. Since the attacker has no information to exploit, he cannot infer the identity of the member and the training data, completing the proof. □

Theorem 3. *If the participants do not colluded with the server, the semi-honest server cannot obtain the global parameters of the deep learning model training and the data set information of the participants.*

Proof. From Theorem 2, we prove that the adversary cannot infer the identity and training data of the honest participants. During the actual execution, suppose the colluder shares the encrypted public and private keys of the secure summation phase with the server, the server can obtain the global parameters, but the colluder does not obtain any useful information. Therefore, it is not worthwhile for the participant to collude with the server. In summary, we believe that the participants do not share the public and private keys with the server, the secure computation is always performed in the ciphertext state, which is transparent to the server, and the global parameters are not leaked to the server, completing the proof. □

5 Experiments

5.1 Setup

The open source datasets MNIST and CIFAR10 are selected as datasets. MNIST is a dataset of images of handwritten digits. It contains 60,000 training images and 10,000 test images. In particular, the resolution of each image is 28×28. The CIFAR10 dataset consists of 32×32 color images of 10 classes, with 6000 images per class. There are 50000 training images and 10000 test images. We will split the data by client ID, so that different clients have different sub-datasets that do not intersect with each other.

The programming language used in the experiment is python. The PyTorch is used to train the model. The neural network model used is the CNN, whose structure mainly consists of a fully connected layer, a pooling layer, a ReLU function, and a convolutional layer. This paper uses different numbers of clients to better measure the effectiveness of the scheme.

5.2 Accuracy Evaluation

We evaluate three cases: the plaintext federated learning with no privacy of model updates(plain-FL), the federated learning with paillier encryption model parameters(paillier-FL), and our secure protocol. In this paper, we measure the performance of the model by testing the accuracy of the model on a test set. For the MNIST and CIFAR10 datasets, we randomly shuffle the training examples and distribute them evenly across clients. The batch size is 100, the learning rate is 0.005, and each local training iteration is set to 3. The number of clients in the experiment is 5, 10, and 20, respectively.

Table 2: The peak accuracy of different schemes

		Accuracy		
Clients	**Dataset**	plain-FL	our system	paillier-FL
MNIST	5	99.67%	99.59%	99.36%
	10	99.58%	99.71%	99.26%
	20	99.71%	99.65%	99.54%
CIFAR10	5	60.62%	60.67%	60.01%
	10	60.54%	60.53%	59.96%
	20	60.91%	60.64%	59.92%

The results of the experiment are shown in Figure 3. The accuracy of our system remains almost the same as that of plain-FL, both on the MNIST dataset and on the CIFAR10 dataset. The highest peak value of our system is 99.67% on the MNIST dataset and 60.42% on the CIFAR10 dataset. The specific results based on different number of clients are shown in Table 2. Our scheme maintains a similar accuracy to plain-FL for different numbers of clients, and the accuracy curves of our scheme on the MNIST and CIFAR10 datasets match overall, with almost negligible differences.

5.3 Time Evaluation

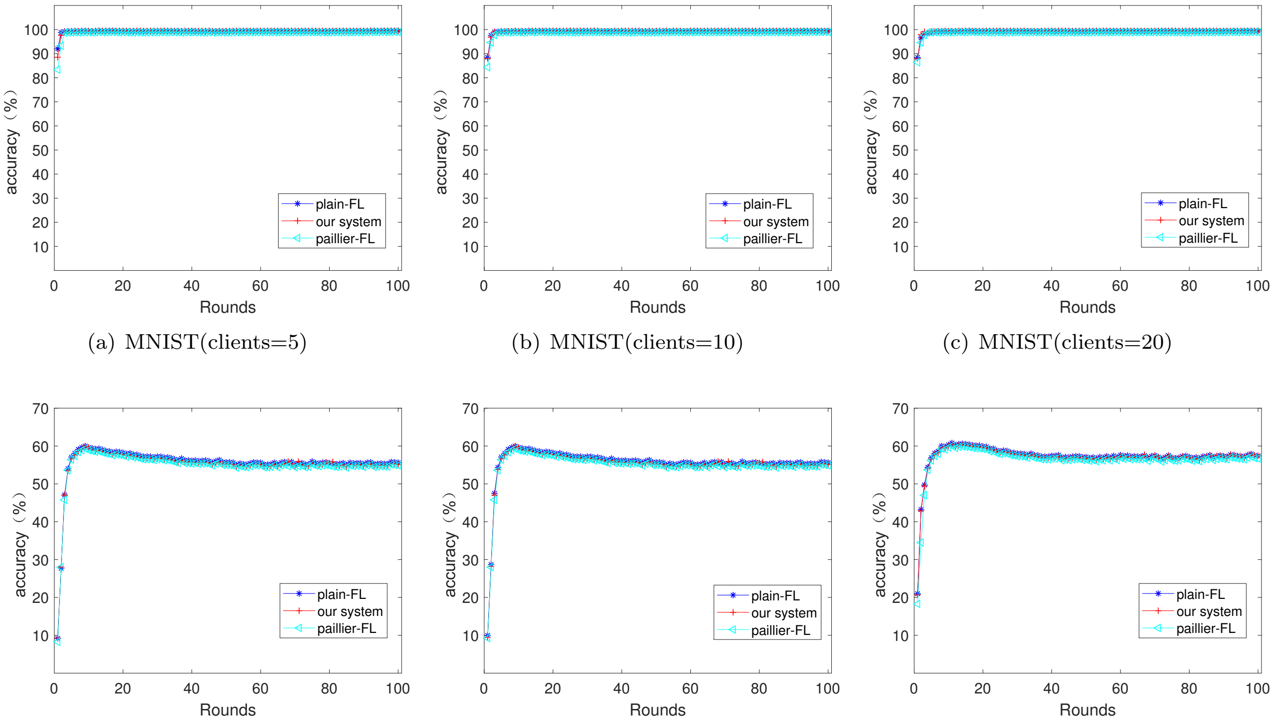
When homomorphic encryption schemes are used throughout to protect the privacy of federal learning, a significant time overhead is incurred because a large number of model parameters must be encrypted and computed all at once. Our scheme does not increase the time overhead of each round of iterations because it only uses the encryption algorithm in the initialization phase. We ran the experiments for 100 iterations, and present the results of three schemes averaged in Figure 4. Our design is more efficient than paillier-FL on the MNIST dataset (up to 49×) and on the CIFAR10 dataset (up to 108×). The aggregation efficiency is greatly improved.

6 Conclusions

Federated learning is a collaborative approach that enables clients to train machine learning models without disclosing sensitive data locally. However, several studies have shown that during the learning process, whether local information, such as uploaded models or gradients, or global information, such as aggregated models, is still vulnerable to exploitation by adversaries, thereby compromising user privacy. In this paper, we propose a privacy-preserving federated learning aggregation scheme that utilizes parameter obfuscation and restoration to protect the local privacy of semi-honest participants against collusion attacks, even if n-2 participants collude with the server. The scheme ensures that no global parameters are leaked to the aggregation server, and we prove that it does not compromise the accuracy of the federated learning model.

Acknowledgments

This work is supported by the National Natural Science Foundations of China (Grant No.61972225), the Natural Science Basic Research Program Funded by Shaanxi Province (Program No.2021JM-453, 2023JC-YB-826), Key Industry Innovation Chain in Shaanxi Province (No.2020ZDLGY07-05).



(a) MNIST(clients=5)

(b) MNIST(clients=10)

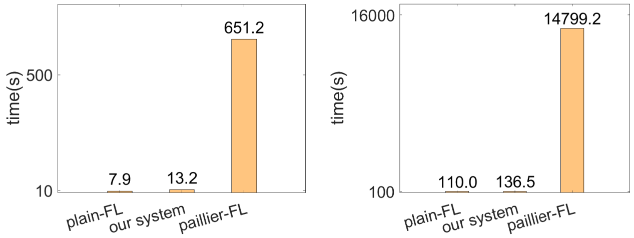
(c) MNIST(clients=20)

(d) CIFAR(clients=5)

(e) CIFAR(clients=10)

(f) CIFAR(clients=20)

Figure 3: The change curve of model accuracy as the number of iterations increases



(a) MNIST

(b) CIFAR

Figure 4: Comparison of the time consumption of the average round of iterations

References

[1] M. Abadi, A. Chu, I. Goodfellow, H. B. McMahan, I. Mironov, K. Talwar, and L. Zhang, “Deep learning with differential privacy,” in *Proceedings of the 2016 ACM SIGSAC conference on computer and communications security*, 2016, pp. 308–318.

[2] Y. Aono, T. Hayashi, L. Wang, S. Moriai *et al.*, “Privacy-preserving deep learning via additively homomorphic encryption,” *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 5, pp. 1333–1345, 2017.

[3] S. Banabilah, M. Aloqaily, E. Alsayed, N. Malik, and Y. Jararweh, “Federated learning review: Fundamentals, enabling technologies, and future applications,” *Information processing & management*, vol. 59, no. 6, p. 103061, 2022.

[4] A. Bellet, R. Guerraoui, M. Taziki, and M. Tommasi, “Fast and differentially private algorithms for decentralized collaborative machine learning,” Ph.D. dissertation, INRIA Lille, 2017.

[5] S. A. Bello, L. O. Oyedele, O. O. Akinade, M. Bilal, J. M. D. Delgado, L. A. Akanbi, A. O. Ajayi, and H. A. Owolabi, “Cloud computing in construction industry: Use cases, benefits and challenges,” *Automation in Construction*, vol. 122, p. 103441, 2021.

[6] K. Bonawitz, H. Eichner, W. Grieskamp, D. Huba, A. Ingerman, V. Ivanov, C. Kiddon, J. Konečnỳ, S. Mazzocchi, B. McMahan *et al.*, “Towards federated learning at scale: System design,” *Proceedings of machine learning and systems*, vol. 1, pp. 374–388, 2019.

[7] K. Bonawitz, V. Ivanov, B. Kreuter, A. Marcedone, H. B. McMahan, S. Patel, D. Ramage, A. Segal, and K. Seth, “Practical secure aggregation for privacy-preserving machine learning,” in *proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, 2017, pp. 1175–1191.

[8] P. S. Chung, C. W. Liu, M. S. Hwang, “A study of attribute-based proxy re-encryption scheme in cloud environments”, *International Journal of Network Security*, vol. 16, no. 1, pp. 1-13, 2014.

[9] S. Dong, P. Wang, and K. Abbas, “A survey on deep learning and its applications,” *Computer Science Review*, vol. 40, p. 100379, 2021.

[10] Y. Dong, X. Chen, L. Shen, and D. Wang, "Eastfly: Efficient and secure ternary federated learning," *Computers & Security*, vol. 94, p. 101824, 2020.

[11] A. H. Gandomi, F. Chen, and L. Abualigah, "Machine learning technologies for big data analytics," p. 421, 2022.

[12] T.-T. Gao, H. Li, and S.-L. Yin, "Adaptive convolutional neural network-based information fusion for facial expression recognition," *International Journal of Electronics and Information Engineering*, vol. 13, no. 1, pp. 17–23, 2021.

[13] J. Geiping, H. Bauermeister, H. Dröge, and M. Moeller, "Inverting gradients-how easy is it to break privacy in federated learning?" *Advances in Neural Information Processing Systems*, vol. 33, pp. 16 937–16 947, 2020.

[14] J. Geng, Y. Mou, Q. Li, F. Li, O. Beyan, S. Decker, and C. Rong, "Improved gradient inversion attacks and defenses in federated learning," *IEEE Transactions on Big Data*, 2023.

[15] B. Hitaj, G. Ateniese, and F. Perez-Cruz, "Deep models under the gan: information leakage from collaborative deep learning," in *Proceedings of the 2017 ACM SIGSAC conference on computer and communications security*, 2017, pp. 603–618.

[16] K. Hu, Y. Li, M. Xia, J. Wu, M. Lu, S. Zhang, and L. Weng, "Federated learning: a distributed shared machine learning method," *Complexity*, vol. 2021, pp. 1–20, 2021.

[17] W. P. Hu, C. B. Lin, J. T. Wu, C. Y. Yang, M. S, Hwang, "Research on privacy and security of federated learning in intelligent plant factory systems", *International Journal of Network Security*, vol. 25, no. 2, pp. 377-384, 2023.

[18] C. W. Liu, W. F. Hsien, C. C. Yang, M. S. Hwang, "A survey of attribute-based access control with user revocation in cloud data storage", *International Journal of Network Security*, vol. 18, no. 5, pp. 900-916, 2016.

[19] C. W. Liu, W. F. Hsien, C. C. Yang, and M. S. Hwang, "A survey of public auditing for shared data storage with user revocation in cloud computing", *International Journal of Network Security*, vol. 18, no. 4, pp. 650-666, 2016.

[20] J. Liu, J. Huang, Y. Zhou, X. Li, S. Ji, H. Xiong, and D. Dou, "From distributed machine learning to federated learning: A survey," *Knowledge and Information Systems*, vol. 64, no. 4, pp. 885–917, 2022.

[21] H. B. McMahan, E. Moore, D. Ramage, and B. A. y Arcas, "Federated learning of deep networks using model averaging," *arXiv preprint arXiv:1602.05629*, vol. 2, 2016.

[22] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Artificial intelligence and statistics*. PMLR, 2017, pp. 1273–1282.

[23] H. B. McMahan, D. Ramage, K. Talwar, and L. Zhang, "Learning differentially private recurrent language models," *arXiv preprint arXiv:1710.06963*, 2017.

[24] L. Melis, C. Song, E. De Cristofaro, and V. Shmatikov, "Inference attacks against collaborative learning," *arXiv preprint arXiv:1805.04049*, vol. 13, 2018.

[25] T. Orekondy, S. J. Oh, B. Schiele, and M. Fritz, "Understanding and controlling user linkability in decentralized learning," *arXiv preprint arXiv:1805.05838*, 2018.

[26] A. Pyrgelis, C. Troncoso, and E. De Cristofaro, "Knock knock, who's there? membership inference on aggregate location data," *arXiv preprint arXiv:1708.06145*, 2017.

[27] Y. Shengxing and C. Zhong, "Efficient secure federated learning aggregation framework based on homomorphic encryption." *Journal on Communication/Tongxin Xuebao*, vol. 44, no. 1, 2023.

[28] R. Shokri and V. Shmatikov, "Privacy-preserving deep learning," in *Proceedings of the 22nd ACM SIGSAC conference on computer and communications security*, 2015, pp. 1310–1321.

[29] R. Shokri, M. Stronati, C. Song, and V. Shmatikov, "Membership inference attacks against machine learning models," in *2017 IEEE symposium on security and privacy (SP)*. IEEE, 2017, pp. 3–18.

[30] C. Song, T. Ristenpart, and V. Shmatikov, "Machine learning models that remember too much," in *Proceedings of the 2017 ACM SIGSAC Conference on computer and communications security*, 2017, pp. 587–601.

[31] P. Veličković, N. D. Lane, S. Bhattacharya, A. Chieh, O. Bellahsen, and M. Vegreville, "Scaling health analytics to millions without compromising privacy using deep distributed behavior models," in *Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare*, 2017, pp. 92–100.

[32] L. Yan, X. Wang, and S. Yin, "Campus garbage image classification algorithm based on new attention mechanism," *International Journal of Electronics and Information Engineering*, vol. 13, no. 4, pp. 131–141, 2021.

[33] L. Yin, J. Feng, H. Xun, Z. Sun, and X. Cheng, "A privacy-preserving federated learning for multiparty data sharing in social iots," *IEEE Transactions on Network Science and Engineering*, vol. 8, no. 3, pp. 2706–2718, 2021.

[34] C. Zhang, S. Li, J. Xia, W. Wang, F. Yan, and Y. Liu, "Batchcrypt: Efficient homomorphic encryption for cross-silo federated learning," in *Proceedings of the 2020 USENIX Annual Technical Conference (USENIX ATC 2020)*, 2020.

[35] C. Zhang, Y. Xie, H. Bai, B. Yu, W. Li, and Y. Gao, "A survey on federated learning," *Knowledge-Based Systems*, vol. 216, p. 106775, 2021.

[36] X. Zhang, A. Fu, H. Wang, C. Zhou, and Z. Chen, "A privacy-preserving and verifiable federated learning scheme," in *IEEE International Conference on Communications (ICC'20)*, pp. 1–6, 2020.

Biography

Xiangxiang Ma is a master student of the School of Computer Science, Xi'an Polytechnic University, China. His current research interests is designing privacy-preserving protocols based on Federated learning.

Jian Chen is a master student of the School of Computer Science, Xi'an Polytechnic University, China. His current research interests is Privacy protection and task allocation of crowdsensing.

Linming Gong is a lecturer in the School of Computer Science, Xi'an Polytechnic University, China, and a member of International Association for Computing Machinery (ACM). He received the Ph.D. degree from the School of Computer Science, Shaanxi Normal University, Xi'an, China, in 2017. His current research interests include applied cryptography, secure multiparty computation, computer and network security, mobile and wireless communication security, and privacy-preserving data mining.

Daoshun Wang received the B.S. degree from the Department of Mathematics, Lanzhou University, Lanzhou, China, in 1987, and the Ph.D. degree from the Department of Mathematics, Sichuan University, Chengdu, Sichuan, China, in 2001. He is currently an Associate Professor in the Department of Computer Science and Technology, Tsinghua University, Beijing, China. His research interests include visual cryptography, digital watermarking, and label anti-counterfeiting.

Differential Privacy Publishing of Location-based Statistical Data Based on Nonuniform Quadtree

Yan Yan, Yue Zhang, and Xingying Qian

(Corresponding author: Yue Zhang)

School of Computer and Communication, Lanzhou University of Technology
No.287 Langongping Road, Qilihe District, Lanzhou 730050, China
Email: 212081201001@lut.edu.cn

(Received Mar. 27, 2023; Revised and Accepted Nov. 22, 2023; First Online Dec. 24, 2023)

Abstract

Differential privacy publishing of location-based statistical data can reduce various risks caused by privacy leakage while keeping the statistical characteristics of the data. A nonuniform quadtree structure-based privacy protection method is proposed to improve the problems of the unreasonable structure and the low efficiency of the traditional statistical partition and publishing methods. Firstly, a reasonable spatial decomposition structure is constructed by nonuniform quadtree partition and depth-first traversal according to data distribution characteristics. Secondly, the corresponding differential privacy budget allocation and adjustment method ensures that all partitioned regions have the same differential privacy protection strength. Finally, neighbors with the same attributes in the nonuniform quadtree partition structure are merged, and Laplace noise is incorporated into the statistical results to form the published data. Experimental comparison of the real-world datasets proves the advantages of the envisaged method in improving the querying accuracy of the published data and the execution efficiency of the publishing algorithm.

Keywords: Differential Privacy; Location Privacy; Privacy Budget Allocation; Privacy Spatial Decomposition; Statistical Publishing of Location-based Data

1 Introduction

Location-based data are continuously generated and provided by terminals, devices, sensors, vehicles, and personnel from various industries such as Mobile Internet, Internet of Vehicles, Internet of Things, digital healthcare, finance, consumption, communication, etc. According to a survey, more than 80% of real-world data is related to geographic location [9]. Location-based statistical analysis and mining of big data have been widely used in many fields, such as socio-demographic surveys, public health research, urban traffic and road planning, social opinion analysis, business model investigation and adjustment, agricultural yield and disease prediction, bioinformation analysis, and many other fields.

However, the release of location-based big data statistical information not only brings convenient services but also causes many problems, such as data abuse, personal privacy disclosure, and trade secrets violation. Location-based data are quite sensitive which can reflect the specific location of devices, the degree of system association, and personal privacy. Improper release or reverse inference analysis of location-based big data may expose the specific locations of key devices and nodes as well as the roles they play in the system, or even threaten the physical security and communication security of related devices and nodes [1, 2]. It may lead to the leakage of users' home addresses, living habits, health status, economic conditions, social relationships, and other private information [3, 15], which endangers the legitimate rights and interests of users and the safety of life and property. Therefore, solving the privacy protection problem in the process of location-based big data statistical publishing has become the most urgent task for the development of the location-related big data industry.

Differential privacy model [5, 6] is currently a hot topic of research in data publishing privacy protection techniques, which do not require prior assumptions about the background knowledge that attackers may have. By adding random perturbations to the published data, the attacker cannot identify whether a certain tuple is in the original data in the statistical sense, no matter what background knowledge he has. The release of location-based big data statistical information can be used to query the number of other users within a certain range, the number of available means of transportation, traffic conditions, etc., or it can be used to improve the quality of location-based services through further analysis and consolidation [18]. The statistical partition and publication process decomposes and indexes the two-dimensional space based on a specific index structure and publishes statistics number of location points within the indexed area, which reduced the risk of leakage of the user's real location. The privacy protection effect of location-based statistical publishing results can be further improved by adding differential privacy noise to the statistical values of the real location points in the indexed region.

Most existing statistical partition and publication methods adopt a top-down spatial decomposition process. Grid or tree index structures are usually constructed and a similar partition strategy is iteratively executed for each sub-region, resulting in unreasonable partition structure, increased noise error and uniform assumption error, and the loss of query accuracy of published statistics. To address the above problems, this paper proposes a new partition method and the corresponding differential privacy budget allocation and adjustment scheme, which effectively improves the availability of published data and the efficiency of the publishing algorithm.

The rest of the paper is organized as follows: Section 2 reviews the state-of-the-art pertinent to private spatial decomposition methods. Section 3 defines the differential privacy model for publishing location-based statistical data. Section 4 analyzes the problems of traditional privacy spatial decomposition and proposes a nonuniform quadtree partitioning algorithm. Section 5 details the differential privacy budget allocation and statistical release methods. Finally, Section 6 reports a set of empirical studies, whereas Section 7 concludes the paper and lists the limitations of this study and the work to be done in the future.

2 Related Work

To protect the privacy of location information, various solutions have been proposed by researchers, such as location anonymization [14], suppression [10], encryption [8], perturbation [24], etc. Differential privacy partitioning and publishing have been popular location-based statistical information protection methods recently. Privacy spatial decomposition is carried out and random noise is incorporated into the statistical results to achieve differential privacy protection of the published data. The uniform grid (UG) algorithm, proposed by Qardaji *et al.* [12] decomposes the two-dimensional space into grids of uniform size. Although the structure is simple and efficient, the availability of published results is low. Therefore, the authors propose the adaptive grid (AG) algorithm to perform mesh partition for the first layer of grids, which improves the query accuracy of statistical publishing results. Zhou *et al.* [25] propose a three-level standard deviation circle radius adaptive grid decomposition method (SDCAG) based on Bernoulli random sampling technique. The exponential mechanism and high-pass filtering technology filter out grid cells smaller than a predefined threshold at the second level of the partitioning structure. The grid cells larger than the threshold are further divided, while the adjacent grid cells smaller than the threshold are merged to form coarse-grained grid cells. The spatial decomposition algorithm proposed by Rodríguez *et al.* [13] extracts the optimal spatial decomposition with the help of statistical information about the spatial distribution of particles so that each cell included spatially uniformly distributed particles. The location privacy protection scheme proposed by Wei *et al.* [17] uses an adaptive three-level grid decomposition (ATGD) algorithm and a differential privacy-based adaptive complete pyramid grid (DPACPG) algorithm to split the location information in the mobile crowdsourcing service into multi-level grids containing noise. Yan *et al.* [21] propose a big data statistical information publishing method that combines grid partitioning and clustering. Firstly, the two-dimensional space is divided into uniform grids to calculate big data statistical information. Then, the uniformity of the non-empty grid is judged, and the uniformly distributed grids are graded using the wavelet decomposition method. Finally, the neighborhood similarity grids are clustered, and Laplace noise is incorporated within the same density level, which better improved the availability of statistical publishing results.

Tree-based partition structure has better hierarchical inclusion characteristics and can provide more convenient spatial query services. Cormode *et al.* [4] use the data-independent complete quadtree to divide two-dimensional space and design the geometric privacy budget allocation method, which increases the accuracy of range counting queries to a certain extent. Yang *et al.* [22] improve the traditional quadtree partition structure and propose a density-based partitioning (DBP) algorithm, which recursively divides the region into four sub-units according to the maximum density difference. Some partition methods combine the grid structure with the tree structure to realize a hybrid partition structure. Zhang *et al.* [23] propose a grid-based quadtree range query (GT-R) response method. The grid is used to divide the value domain of user data uniformly to produce cell regions of equal size, while the quadtree structure is used to index all cell regions. The hierarchical differential privacy hybrid decomposition (HDPHD) strategy proposed by Yan *et al.* [20] firstly implements adaptive grid partition according to the density distribution of the dataset. Then, two density thresholds are set to classify the grids into three types. The grids of different density types are partitioned by the adaptive grid method and heuristic quadtree decomposition algorithm respectively, which effectively improves the availability of published data. In reference [19], the authors propose an unbalanced quadtree partition (UBQP) method based on the condition of regional uniformity, which traverses the subtrees according to the depth-first strategy and adaptively performs quadtree iterative partitioning according to the uniformity condition.

3 Definition and Characteristics of Differential Privacy

The differential privacy model is a privacy protection method based on data perturbation. It has a rigorous mathematical definition and can quantitatively analyze the degree of privacy protection. The main principle is to protect users' privacy by adding random noise to the statistical results of the original data and removing individual characteristics on the premise of preserving the statistical characteristics of the dataset. The differential privacy model does not need to establish special attack assumptions because the attacker cannot determine whether a certain record exists in the original dataset based on different output results [16]. As a research hotspot of privacy protection technology, the differential privacy model has a natural match with the location-based big data statistical publishing application. This is because of the large scale and dynamic change of location-related big data, which makes the impact of adding or deleting certain data in the dataset very small on the whole, which is very consistent with the connotation of the definition of differential privacy.

Definition 1. *ϵ-differential privacy [6]: For the sibling datasets T_1 and T_2 (T_1 and T_2 have the same attributes and differ by only one tuple) and any output $S \subseteq Range(M)$, if the probability that an algorithm M obtains the same output result on T_1 and T_2 satisfies the following inequality, then the algorithm M is said to satisfy ϵ-differential privacy.*

$$P_r[M(T_1) \in S] \leq e^{\epsilon} \times P_r[M(T_2) \in S] \quad (1)$$

In Definition 1, ϵ is the privacy budget. The smaller it is, the greater the strength of privacy protection provided by algorithm M; conversely, the smaller the strength of privacy protection. According to the definition of differential privacy, it is clear that even if an attacker gets all the other data except a specific target record, it is still impossible to infer whether the target record exists in the original data, thus achieving privacy protection for user data.

Definition 2. *Sensitivity [7]: Given a query mapping function f, its sensitivity Δf is defined as the maximum $L1$ parametric distance between the outputs of that query mapping function on the sibling datasets T_1 and T_2:*

$$\Delta f = max_{T_1,T_2} \parallel f(T_1) - f(T_2) \parallel_1 \quad (2)$$

For location-based big data statistical publishing applications, the sensitivity of statistical publishing results is $\Delta f = 1$.

Definition 3. *Laplace mechanism [7]: The Laplace noise mechanism achieves differential privacy protection by adding a small amount of independent noise to the output result of a query mapping function f (f: $T \rightarrow R^d$, that is: a mapping relationship between a dataset T and a d-dimensional space). $f(T)$ is used to represent the result obtained by the query mapping function f applied to the original dataset T, then the query result returned by the Laplace mechanism can be expressed as $M(T) = f(T) + (Y_1, Y_2, ..., Y_d)$. Where, Y_i is an independent continuous random variable with the same distribution, and its probability density function is:*

$$Lap(x \mid b) = \frac{1}{2b} e^{-\frac{|x|}{b}} \quad (3)$$

It can be observed from formula (3) that the incorporated independent noise follows the Laplace distribution with the mean of 0 and the scale parameter of b. Combined with the definition of sensitivity, ϵ-differential privacy is satisfied when the scale parameter $b = \frac{\Delta f}{\epsilon}$.

Theorem 1. *Serial combination property [11]: A set of random algorithms $\{M_1, M_2, ..., M_n\}$ acting on the same dataset T, where M_i ($1 \leq i \leq n$) satisfies ϵ_i-differential privacy on the dataset T, then the set of random algorithms $\{M_1, M_2, ..., M_n\}$ as a whole can achieve $\sum_{i=1}^{n} \epsilon_i$-differential privacy for the dataset T.*

Theorem 2. *Parallel combination property [11]: If the dataset T can be divided into multiple independent and disjoint subsets $\{T_1, T_2, ..., T_n\}$, a set of randomized algorithms $\{M_1, M_2, ..., M_n\}$ acting on the above data subset respectively, where M_i ($1 \leq i \leq n$) satisfies ϵ_i-differential privacy for the data subset T_i, then the set of random algorithms $\{M_1, M_2, ..., M_n\}$ can achieve $max\{\epsilon_i\}$-differential privacy for the dataset T.*

4 Privacy Spatial Decomposition Based on Nonuniform Quadtree

Statistical partition and publication of big data based on differential privacy protection have a certain difference between the published results and the original statistical results, which mainly come from two aspects. One is the noise error introduced by the differential privacy model. For an arbitrary spatial query range Q, with $C(Q)$ and $C(Q^*)$ denoting the original data statistics and the release result statistics within the query range Q respectively, the error between them is called the noise error, which can be calculated by the following formula:

$$NE(Q) = \mid C(Q) - C(Q^*) \mid \quad (4)$$

Another source of the difference between published results and the original statistic is the uniform assumption error. The partition and distribution methods for two-dimensional spaces tend to assume that the data is uniformly distributed within the smallest granularity of partition areas. q_i is used to represent the partition area intersecting with the query area Q, and r_i is the proportion of the area intersecting with the query area Q and each partition area. The error within the query area Q caused by the uniformity assumption estimate is called the uniform assumption error and can be calculated by the following formula:

$$UE(Q) = \mid \sum_{i=1}^{m} r_i \cdot C(q_i) - C(Q) \mid \quad (5)$$

Location-related big data is spatially correlated and regionally heterogeneous. For example, areas with relatively complete infrastructure or developed traffic roads have a denser distribution of users, so a large number of location-related big data will be generated. On the contrary, the distribution of users is relatively sparse, and the amount of data is relatively small. The reasonableness of the division release structure not only affects the statistical release error of location-based big data but also directly affects the time efficiency of the release process. The traditional quadtree partitioning structure recursively divides the entire region into four sub-regions of equal size until a set partitioning stopping condition is met (as shown in Figure 1). This method always divides from the center of the region without considering the real distribution of the data. In the process of division, a large number of null nodes (i.e., blocks with a statistical value of 0) are likely to be generated. When the differential privacy noise disturbance is added to these regions later, a large noise error will be introduced, leading to the decline of statistical characteristics of published results. In addition, the division process requires recursive access, uniformity judgment, and further partitioning of individual sub-nodes of the quadtree (in Figure 1, lines of different colors are used to mark the division boundaries of different levels of the quadtree), which consumes a long execution time.

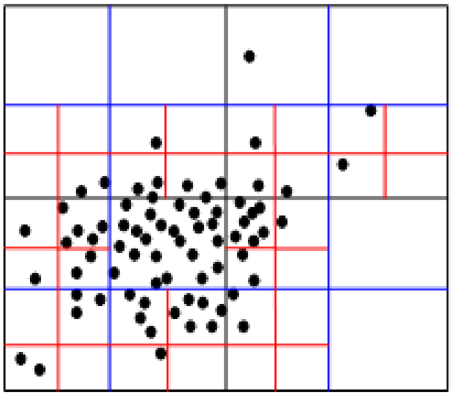


Figure 1: Standard quadtree iterative partition diagram

In order to solve the problems of publishing error and time efficiency caused by the above unreasonable division structure, this paper proposes a spatial decomposition algorithm based on nonuniform quadtree structure (as shown in Figure 2). The method adaptively adjusts the segmentation position of the quadtree according to the distribution characteristics of data, instead of simply dividing the quadtree in half, so that the spatially adjacent regions with the same distribution characteristics maintain better connectivity and avoid a large number of empty nodes caused by over-partitioning. By setting a uniformity judgment condition and the partitioning stopping conditions and traversing each sub-region through the depth-first principle, the partition depth can be determined adaptively according to the distribution characteristics of data.

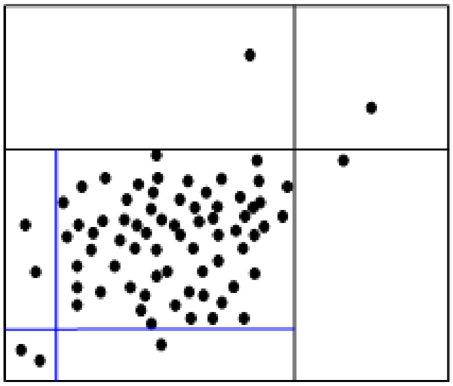


Figure 2: Nonuniform quadtree iterative partition diagram

The nonuniform quadtree space partition method proposed in this paper sets the following partitioning stop conditions:

1) If the data in the region is uniformly distributed, the division is stopped to avoid the accumulation of noise error.

2) If the statistical value in the region is less than the set threshold n, stop the division to avoid generating too many empty nodes.

3) If the area covered by the region is less than the set threshold r, stop dividing to avoid the region being too small.

Definition 4. *Regional uniformity: For a region S with density $Den(S)$, $Den(D_1)$, $Den(D_2)$, ..., $Den(D_i)$ are used to represent the sub-region density after partitioning in multiple directions such as horizontal, vertical, diagonal, etc. Construct the row vector $V = \{Den(D_1), Den(D_2), ..., Den(D_i)\}$, then the distribution uniformity of the region S can be represented by the variance $Var(V)$ of the vector V. If the following Formula (6) is satisfied, then the region S can be considered as uniformly distributed.*

$$\left|\log_{10}(Var(V)) - \log_{10}(\frac{Den(S)}{i})^2\right| \leq \theta \tag{6}$$

Where i is the number of sub-regions after multi-directional segmentation and θ is the threshold value.

Definition 5. *Four-connected regions: Regions are considered to be four-connected if they are located above, below, left, or right of other regions adjacent to them.*

Definition 6. *Dense region: For a region S with statistical value $count(S)$, which is divided into m_x and m_y intervals according to the X-axis direction and Y-axis direction respectively, the region with statistical value $count(S)$ satisfying the following inequality is called a dense region.*

$$Den > \frac{count(S)}{m_x \cdot m_y} \tag{7}$$

Algorithm 1 delineates the detailed implementation process of the proposed spatial decomposition algorithm based on nonuniform quadtree (SDNQ).

Algorithm 1 Spatial decomposition algorithm based on nonuniform quadtree (SDNQ)

```
1: Input: dataset $T$, uniformity judgment threshold $\theta$,
   empty node threshold $n$, minimum node threshold $r$
2: Output: the maximum partition depth of nonuniform
   quadtree $h$, partition structure matrix $P$
3: $h = 0$, $P = \Phi$
4: $current\ node \leftarrow T$
5: if the number of data points in the current node $\leq n$
   or the size of the current node $\leq r$ then
6:     $P = P \bigcup current\ node$
7: else if regional uniformity within the current node
   $\leq \theta$ then
8:     $P = P \bigcup current\ node$
9: else
10:    find the boundary of the most four-connected regions based on the dense regions
11:    the current node is divided into four sub-regions by
       nonuniform quadtree, $T = \{ S_1, S_2, S_3, S_4\}$
12:    $h = h + 1$, $P = P \bigcup \{ S_1, S_2, S_3, S_4\}$
13:    for $i \in \{1, 2, 3, 4\}$ do
14:        $current\ node \leftarrow S_i$
15:        go to step 5
16:    end for
17: end if
18: return $h$, $P$
```

5 Differential Privacy Budget Allocation and Publishing Protection

In order to realize the privacy protection of the published results of the statistical partition of big data with the help of the differential privacy model, it is necessary to combine the allocation of the differential privacy budget with the spatial division structure and add Laplace noise to the statistical results of each region. Considering that the nonuniform quadtree partitioning structure proposed in this paper retains the hierarchical relationship of the quadtree on the one hand, and not all regions have the same partitioning depth on the other hand, the following privacy budget allocation scheme is adopted. Firstly, according to the maximum division depth h of the nonuniform quadtree, the geometric series differential privacy budget allocation is carried out for the nodes of each layer [4], that is, the differential privacy budget of $\epsilon_i = \frac{\epsilon(\sqrt[3]{2}-1)2^{\frac{h-i}{3}}}{2^{\frac{h+1}{3}}-1}$ is allocated for the nodes at level i ($i = 0$ corresponds to the root node, $i = h$ corresponds to the leaf node). Then, the privacy budget of each node is adjusted according to the partition structure matrix P of the nonuniform quadtree. For the non-leaf layer nodes that satisfy the partitioning stop conditions, their privacy budget can be adjusted according to formula (8):

$$\epsilon_i^* = \sum_{j=i}^{h+1} \epsilon_j \tag{8}$$

Where ϵ_j denotes the geometric privacy budget obtained for a node with division depth j and ϵ_i^* denotes the adjusted privacy budget for that node. The adjustment method of the privacy budget is shown in Figure 3, where the black nodes indicate the regions that satisfy the partitioning stop conditions and do not continue to be divided. According to the privacy budget adjustment strategy, $\epsilon_2^* = \sum_{j=2}^{4} \epsilon_j = \epsilon_2 + \epsilon_3 + \epsilon_4$ and $\epsilon_3^* = \sum_{j=3}^{4} \epsilon_j = \epsilon_3 + \epsilon_4$. This adjustment strategy ensures that each path from the root node to the leaf nodes in a nonuniform quadtree structure satisfies the differential privacy of $\sum_{i=1}^{h+1} \epsilon_i = \epsilon$.

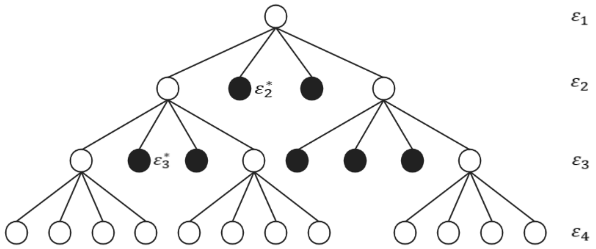


Figure 3: Nonuniform quadtree differential privacy budget adjustment diagram

According to the analysis in Section 4, when the data distribution is relatively uniform, merging spatially adjacent uniform regions can reduce the error introduced by adding Laplace noise; when the data distribution is not uniform, merging the spatially adjacent empty regions can play the same role. Inspired by this, this paper merges regions with the same attributes in the nonuniform quadtree partitioning structure so as to reduce the error caused by noise interference and improve the availability of the published data. Algorithm 2 describes the specific process of privacy protection for big data statistical release based on the nonuniform quadtree partition structure.

Corollary 1. *The published results of statistical partition of big data obtained by Algorithm 2 satisfy ϵ-differential privacy protection.*

Proof. The location-related big data statistical partitioning release mainly provides range counting query services to users, and for an arbitrary query range Q submitted by users, there are the following four situations:

1) The query range Q is completely contained within the region covered by a sub-region. According to the partitioning and merging method of the quadtree, Q may be included within the initial region or within the region formed by merging the subtrees. For the former case, it is known that the region is divided

Algorithm 2 Differential privacy publishing algorithm based on SDNQ

```
Input: dataset $T$, maximum partition depth $h$, partition structure matrix $P$, privacy budget $\epsilon$, sensitivity $S$
Output: Publish dataset $T^*$
algorithm 1 is invoked to obtain the nonuniform quadtree partition structure $P$ and the maximum partition depth $h$
assign geometric privacy budget $\epsilon_i = \frac{\epsilon(\sqrt[3]{2}-1)2^{\frac{h-i}{3}}}{2^{\frac{h+1}{3}}-1}$ to nodes of each layer $(0 \leq i \leq h)$
According to Formula (8), the privacy budget is adjusted for the non-leaf layer nodes in the partition structure $P$ that satisfy the partitioning stop conditions, and the differential privacy budget value $\{\epsilon_1, \epsilon_2, ..., \epsilon_{T_m}\}$ of each region is obtained
for $T_i \in T$ do
    $par \leftarrow$ the parent node of the current region
    $child \leftarrow$ the child node of the current region
    if do not need to merge then
        $N\ (T_i) \leftarrow$ region count value
        $Nx(T_i) \leftarrow N(T_i) + Laplace(\frac{S}{\epsilon_{T_i}})$
    else
        merge regions with the same attributes and calculate the number of regions m
        $sum\_N \leftarrow$ the sum of the combined values
        $sum\_Noise \leftarrow sum\_N + Laplace(\frac{S}{\epsilon_{T_i}})$
        $Nx(T_i) = \frac{sum\_Noise}{m}$
    end if
end for
for $T_i \in T$ do
    if $T_i$ is a leaf node then
        $X(T_i) = Nx(T_i)$
    else
        $X(T_i) = \frac{4^h-4^{h-1}}{4^h-1} \cdot Nx(T_i) + \frac{4^{h-1}-1}{4^h-1} \cdot \sum_{p\in child(T_i)} X(p)$
    end if
end for
for $T_i \in T$ do
    if $T_i$ is a leaf node then
        $NX(T_i) = X(T_i)$
    else
        $NX(T_i) = X(T_i) + \frac{1}{4} \cdot \left[NX(par(T_i)) - \sum_{w\in par(T_i)} X(w)\right]$
    end if
end for
Return $T^* = \{NX(T_1), NX(T_2), ..., NX(T_i)\}$
```

without merging, so Algorithm 2 adds Laplace noise with differential privacy budget of ϵ to it. For the latter case, Algorithm 2 adds Laplace noise with differential privacy budget of ϵ to the merged regions. According to the parallel combination property of the differential privacy model described in Theorem 2, each sub-region within the merged regions satisfies differential privacy protection of $max\{\epsilon_i\} = \epsilon$.

2) The query range Q contains a ($a \geq 1$) complete sub-regions in the division release structure. At this point, each region satisfies the first case described above. Therefore, according to the parallel combination property of the differential privacy model described in Theorem 2, the a complete regions contained in the query range Q, all provide differential privacy protection with $max\{\epsilon_i\} = \epsilon$.

3) The query range Q intersects b ($b \geq 1$) sub-regions. In this case, each of the intersecting regions satisfies case (1), therefore providing differential privacy protection of strength ϵ.

4) The region covered by the query range Q contains a ($a \geq 1$) complete sub-regions and intersects b ($b \geq 1$) sub-regions. For the complete sub-regions in the query range Q, each sub-region can provide ϵ-differential privacy protection similar to that of the case (2). For the sub-regions intersecting the query range Q, the intersecting regions conform to case (3), and therefore, can also provide differential privacy protection with the strength of ϵ.

In summary, Algorithm 2 can provide ϵ-differential privacy protection for the published results of the statistical partition of big data. □

6 Experimentation and Analysis

To comprehensively evaluate and analyze the proposed statistical partition differential privacy publishing algorithm for big data based on nonuniform quadtree (SDNQ), the proposed method is compared with some related methods in recent years from the aspects of the availability of big data statistical publishing results, the operation efficiency of the publishing privacy protection algorithm, and the impact of partition granularity. These include but are not limited to, the adaptive three-level grid decomposition (ATGD) algorithm [17], the density-based partitioning (DBP) algorithm [22], the hierarchical differential privacy hybrid decomposition (HDPHD) strategy [20] and the unbalanced quadtree partition (UBQP) [19] method.

The experimental data are *Checkin*, a user check-in location information provided by Gowalla, *Landmark*, a location information set of 48 states in the United States, provided by infochimps, and *NewYork*, a ride record dataset provided by the NewYork Taxi Regulatory Commission. The experimental environment is Inter(R) Core (TM) i5-7300HQ CPU @ 2.50GHz, 12GB memory, Windows 10 operating system, and the algorithm program was written using MATLAB R2016a software.

Table 1: Experimental dataset and query range information

Parameter	Landmark	Checkin	NewYork
Data amount	870051	1000000	10996214
Coverage range (longitude×latitude)	57°×24°	353°×143°	1.19°×0.99°
Query range (longitude×latitude)	1.25°×0.625°	6°×3°	0.025°×0.02°
	2.5°×1.25°	12°×6°	0.05°×0.04°
	5°×2.5°	24°×12°	0.1°×0.08°
	10°×5°	48°×24°	0.2°×0.16°
	20°×10°	96°×48°	0.4°×0.32°
	40°×20°	192°×96°	0.8°×0.64°

6.1 Availability Analysis of the Location-related Statistical Release of Big Data

According to the application characteristics of location-related big data, the accuracy of the range count query service is used to measure the availability of location-related big data statistical publishing results. Range count queries of different sizes are performed for the published results, and the relative error of the proposed method is compared with the results of ATGD, DBP, HD-PHD, and UBQP.

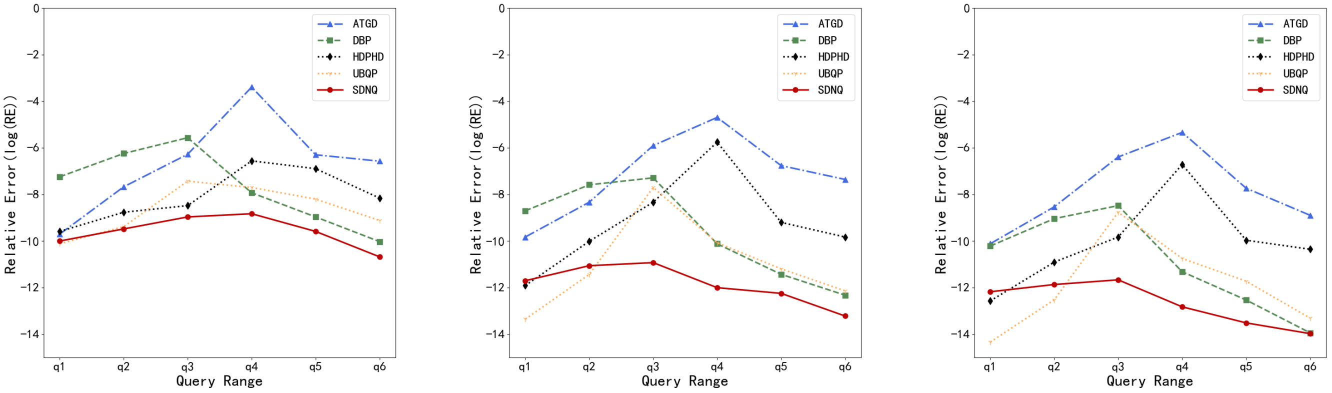
During the experiments, the original parameter settings of the corresponding algorithms are referred to and combined with the size of the dataset. The minimum partition area is set to 3°×1.5° for the *Landmark* dataset, 10°×5° for the *Checkin* dataset and 0.05°×0.025° for the *NewYork* dataset. The sensitivity $S = 1$ of range count query results is caused by adding differential privacy noise. The characteristics of various types of datasets and the size setting of the range count query region are shown in Table 1. The differential privacy model adds Laplace noise with privacy budget $\epsilon = 0.1$, $\epsilon = 0.5$, and $\epsilon = 1$, respectively, and each type of query region is randomly generated 1000 times to determine the average value of the relative error. The relative error is defined as shown in formula (9), where q represents the query range submitted by the user, $C(q)$ is the query statistics obtained in the corresponding range on the original big dataset, and $C^*(q)$ is the query statistics obtained in the corresponding range on the released big dataset. To prevent the denominator from being zero, set $\rho = 0.001 \mid T \mid$, with $\mid T \mid$ denoting the size of the dataset.

$$RE(q) = \frac{|C^*(q) - C(q)|}{max\left\{C(q), \rho\right\}} \tag{9}$$

Figures 4 to 6 depict the comparative results of the relative error of the various algorithms on different datasets in logarithmic coordinates. On the same experimental dataset, the relative error of each algorithm decreases as the privacy budget ϵ increases. The reason is that the added Laplace noise value decreases with the increase of privacy budget ϵ, and the error between the published data and the real statistical value also decreases. On the same dataset, the relative error tends to increase and then decrease as the query range increases from small to large. This is because when the query range is small (e.g., $q1$ and $q2$), the regions included only involve regions close to the leaf nodes in the partition structure, adding less noise interference and making the query error relatively small. When the query range is large (e.g., $q5$ and $q6$), the regions included involve regions close to the root node in the partition structure, and the corresponding query results only need to accumulate the statistics of a few regions, which also introduce less noise error, making the overall query error relatively small.

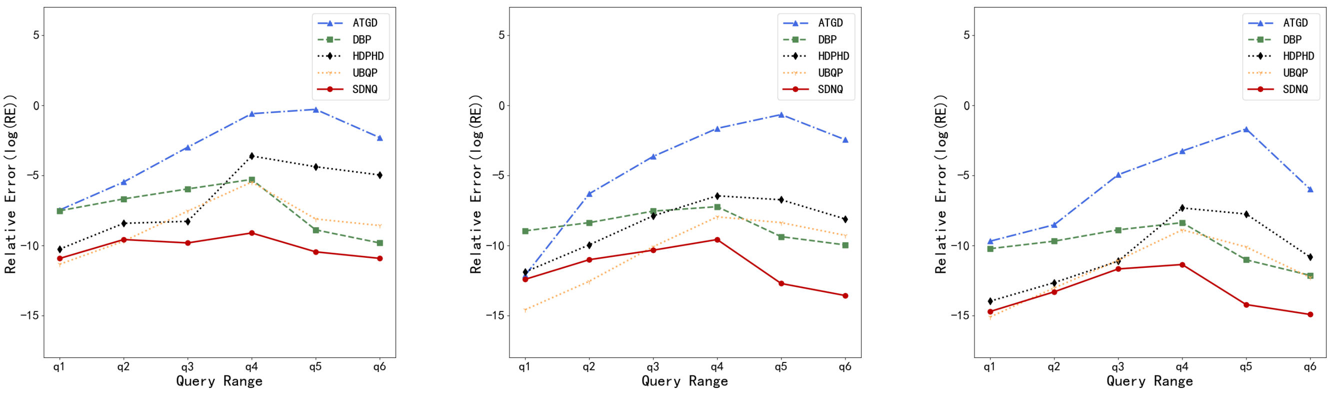
From the perspective of the relative error of range count query of different datasets, the relative error of the *NewYork* dataset with concentrated and dense location-related information distribution is the smallest, followed by that of the *Checkin* dataset, and that of the *Landmark* dataset is the largest. The main reason for this phenomenon is that when the dataset is distributed widely and low-density, many empty nodes are generated in the top-down partitioning process, resulting in excessive noise additions and accumulating high noise error. In contrast, when the dataset is centrally and uniformly distributed, the uniformity assumption estimate can apportion the added noise error to each subunit well, thus reducing the overall error.

The relative error of various partition and release algorithms are compared on the same dataset. The ATGD algorithm performs a three-level adaptive grid partition in the region with dense data distribution, which reduces the uniform assumption error, but cannot overcome the noise error introduced by empty nodes in the region with sparse data distribution. The DBP algorithm divides according to the data density and considers the non-uniformity of data distribution. However, in regions with sparse data distribution, a large number of empty nodes are generated, which introduces more noise error. The HDPHD algorithm uses adaptive mesh partitioning or heuristic quadtree partitioning for different density types respectively, and the UBQP algorithm performs heuristic



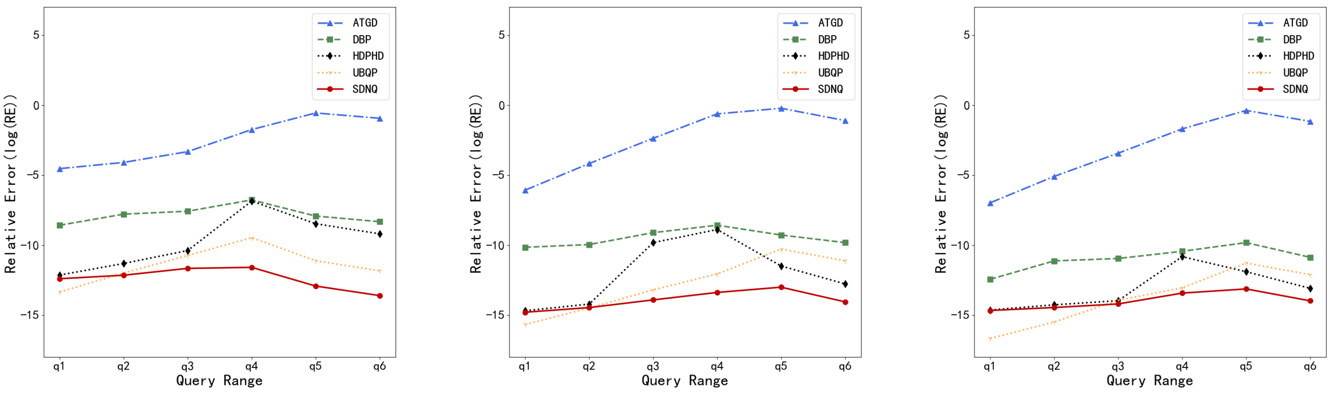
(a) *Landmark dataset*, $\epsilon = 0.1$ (b) *Landmark dataset*, $\epsilon = 0.5$ (c) *Landmark dataset*, $\epsilon = 1$

Figure 4: Comparison of querying accuracy on Landmark dataset



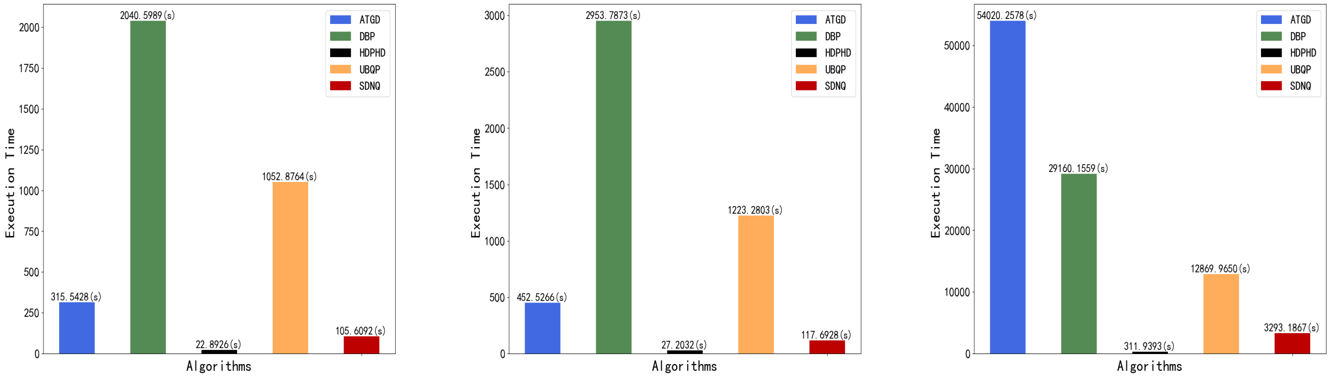
(a) *Chencin dataset*, $\epsilon = 0.1$ (b) *Chencin dataset*, $\epsilon = 0.5$ (c) *Chencin dataset*, $\epsilon = 1$

Figure 5: Comparison of querying accuracy on Chencin dataset



(a) *NewYork dataset*, $\epsilon = 0.1$ (b) *NewYork dataset*, $\epsilon = 0.5$ (c) *NewYork dataset*, $\epsilon = 1$

Figure 6: Comparison of querying accuracy on NewYork dataset



(a) *Landmark dataset*, $\epsilon = 0.5$

(b) *Checkin dataset*, $\epsilon = 0.5$

(c) *NewYork dataset*, $\epsilon = 0.5$

Figure 7: Comparison of operating efficiency of various algorithms under different datasets

quadtree partition according to the uniformity of data distribution, which to some extent compensates for the uniform assumption error caused by blind partition and the noise error caused by empty nodes. The SDNQ method proposed in this paper adaptively determines the division location and merges the regions with the same attributes to reduce the noise error and the uniform assumption error. Therefore, it can basically obtain better query accuracy than other algorithms under different datasets and different privacy budgets.

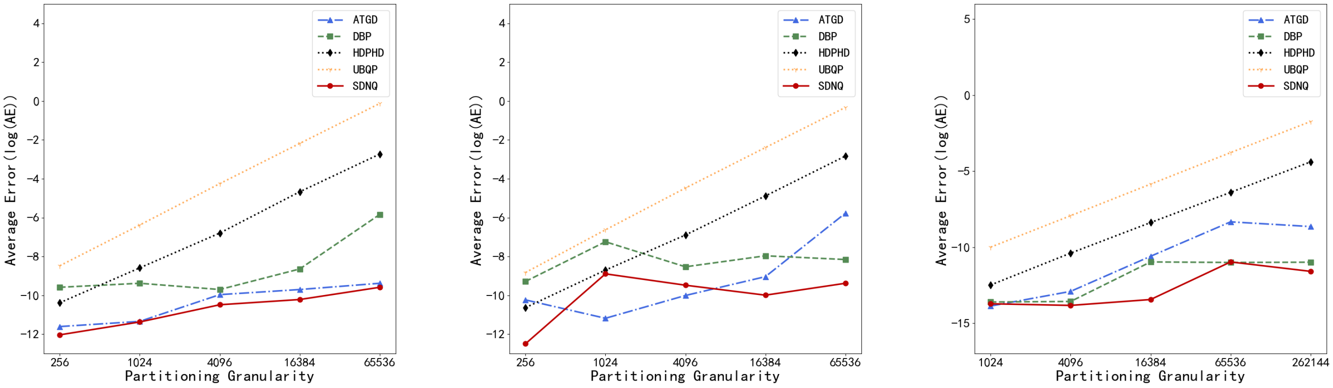
6.2 Efficiency Analysis of Privacy Protection Publishing Algorithm

In order to verify the operation efficiency of the nonuniform quadtree partitioning differential privacy publishing algorithm proposed in this paper, the running time of this algorithm is compared with the results of ATGD, DBP, HDPHD, and UBQP algorithms. Since the differential privacy budget strength has no obvious effect on the running time of the statistical publishing algorithm, this section takes $\epsilon = 0.5$ as an example to conduct an experimental comparison on three real large datasets of different sizes. Figure 7 depicts the execution time of various privacy-preserving algorithms on real datasets. From the experimental results, it can be seen that the overall running time of the various algorithms increases as the size of the dataset increases. The ATGD algorithm needs to perform three partitions combined with the specific distribution of the data, and the partition granularity is larger in the area with denser data distribution, so it takes longer time on the *NewYork* dataset. The DBP algorithm requires a large number of repeated comparisons to find the best division position, and the overall execution time is higher than the other algorithms. The tree-based iterative process of the UBQP algorithm uses depth-first traversal, which is affected by the size of the dataset, and the overall execution time is long. Since the dataset is evenly distributed over dense areas, the HDPHD algorithm avoids unnecessary heuristic quadtree partitioning and therefore consumes less run time than other algorithms. The SDNQ algorithm proposed in this paper can avoid unnecessary divisions in the region where the data distribution is sparse or dense, saving a lot of time.

Table 2: Time complexity comparison of partitioning algorithms

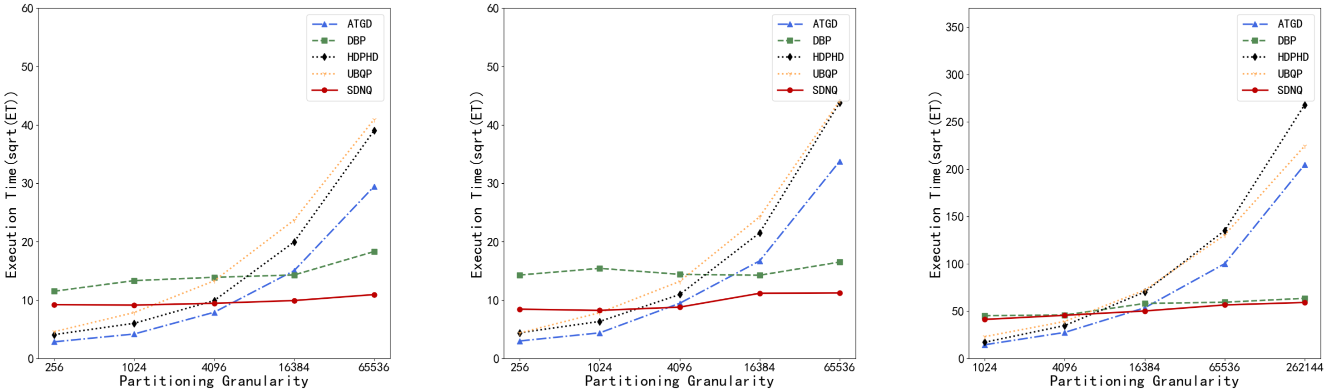
Algorithm	Time Complexity
ATGD	$O(3n)$
DBP	$O(hn)$
HDPHD	$O(n^2)$
UBQP	$O(hn)$
SDNQ	$O(hn)$

Table 2 compares the time complexity of the above partition and publishing algorithms. For a large dataset containing n records, the ATGD algorithm requires three scans of the input data to form a three-level adaptive grid division structure, and the index time complexity is $O(3n)$. The HDPHD algorithm adopts heuristic quadtree partition or adaptive grid partition respectively for different types of first-layer grids, so the overall time complexity is $O(n + n^2) \approx O(n^2)$. The partitioning process of DBP, UBQP, and SDNQ algorithms is related to the specific distribution of the dataset and the partitioning depth h of quadtree. In the worst case, the time complexity of these three algorithms is $O(hn)$ for a location-related dataset with n records and the h layers of division depth.



(a) *Landmark dataset*, $\epsilon = 0.5$ (b) *Checkin dataset*, $\epsilon = 0.5$ (c) *NewYork dataset*, $\epsilon = 0.5$

Figure 8: Partitioning granularity vs. average error



(a) *Landmark dataset*, $\epsilon = 0.5$ (b) *Checkin dataset*, $\epsilon = 0.5$ (c) *NewYork dataset*, $\epsilon = 0.5$

Figure 9: Partitioning granularity vs. execution time

6.3 Impact of Partition Granularity on Overall Performance

According to the analysis in Section 4, the traditional privacy space decomposition method is susceptible to the granularity of segmentation, resulting in over-partitioning or under-partitioning, which not only damages the availability of the published data but also may slow down the efficiency of the overall data publishing algorithm. This section analyses the relationship between partitioning granularity setting, accuracy of publishing results, and efficiency of publishing algorithm. Since the strength of the privacy budget has no significant effect on the setting of partitioning granularity, this section analyses the average error in publishing results due to different granularity settings, using $\epsilon = 0.5$ as an example. To be fair, the initial partition granularity of all comparison algorithms is basically the same. The average error of published results is defined as follows:

$$AE = \frac{\sum_{i=1}^{N} |C_i^* - C_i|}{|T|} \tag{10}$$

Where N denotes the initial partition granularity of the dataset, C_i denotes the original statistics within the region, C_i^* denotes the published statistics corresponding to the region, and $| T |$ is the size of the dataset.

Figure 8 depicts the logarithm results of the average error for all the algorithms at different partition granularity. It is easy to see that the average error of all the partitioning algorithms basically increases with increasing partitioning granularity. The main reason is that when the partition granularity is small, the coverage of a single region is broader and the unevenness of the data distribution within the region may lead to more nonuniform assumption error in the published statistics. With the increase of partition granularity, the coverage of each region will decrease, and the non-uniform assumption error in the region will decrease, but the number of empty regions will increase, so the total noise error will increase. The SDNQ algorithm proposed in this paper has the lowest average error on almost all experimental datasets because the proposed method combines the distribution characteristics of the data well when selecting the partition location and avoids the increase of error caused by the phenomenon of over-partition and under-partition. At the same time, the proposed method merges regions with the same properties in the partition structure, which further reduced the error accumulation in uniform regions and empty regions. The above results demonstrate the superiority of the SDNQ algorithm in terms of the availability of statistical results.

Figure 9 illustrates the execution time versus partition granularity for various partition and publishing algorithms on different datasets in square root coordinates. Intuitively, the execution time increases with the size of the dataset and the partition granularity. As the partition granularity increases, more recursive operations and uniformity judgments are generated; therefore, the execution time of all algorithms increases accordingly. The ATGD, UBQP, and HDPHD algorithms consume less time when the partition granularity is small, but the execution efficiency is significantly reduced when the partition granularity is large. With the increase of partition granularity, the proposed algorithm is more efficient than other algorithms. It can be predicted that the advantages of the proposed algorithm will be more obvious in the case of larger size and more complex distribution of location-based datasets.

7 Conclusions

With the further development of big data technology and the wide popularity of intelligent mobile terminals, various location-based services have been closely related to the work and life of users, and privacy protection issues have also attracted widespread attention. The spatial decomposition and differential privacy protection model can be used to realize the statistical release of location-related information, which may avoid risks caused by privacy leakage under the premise of ensuring data availability. In order to further improve the availability and execution efficiency of statistical partitioning and publishing, this paper proposes a privacy protection partitioning and publishing method based on nonuniform quadtree. The regional partition position is adaptively determined and the partition depth is adjusted according to the distribution characteristics of the dataset. The corresponding differential privacy budget allocation and adjustment method is designed to realize the differential privacy protection of statistical partition published data. The noise error and uniform assumption error in the published results are reduced by merging the adjacent regions with the same attributes. The feasibility and effectiveness of the proposed method are demonstrated by experiments and analysis on big real-location datasets.

However, the research still has some limitations. Firstly, most of the existing spatial partition methods for location-related big data are based on a grid or tree structure, while the areas where humans actually work and live are primarily closely related to the distribution of infrastructure and cannot be well represented by a grid or tree structure. Therefore, it is one of our future works to design more detailed and flexible partition methods combined with efficient indexing structures that are conducive to improving the availability of published results. Secondly, this paper only focuses on the statistical partitioning and publishing of static location-based big data while in real life, the location of users continuously and randomly changes over time. The dynamic change in the location of a large number of users makes the previous spatial division structure of statistical publishing time inapplicable to the next publishing time. Therefore, it will be a direction for further research to design statistical partitioning and publishing methods that combine the dynamic changes of location-based big data and enhance the accuracy and practicality of publishing statistics under the premise of ensuring user privacy. Finally, this paper discusses the statistical publishing method of location-based big data based on the centralized differential privacy model. The premise is that the user's location is collected and statistically published by a trusted third party. However, there is no completely reliable platform for practical applications. Problems such as technical failures, superuser leaks, and hacker attacks make the users' original locations in danger. Therefore, the process of privacy protection of user location is transferred to the user terminal to realize local differential privacy location protection has become a hot research topic in recent years.

Acknowledgments

This study is supported by the National Nature Science Foundation of China (No. 62361036, No. 61762059), and the Nature Science Foundation of Gansu Province (No. 22JR5RA279).

References

[1] R. Agarwal and M. Hussain, "Generic framework for privacy preservation in cyber-physical systems," in *Progress in Advanced Computing and Intelligent Engineering: Proceedings of ICACIE 2019, Volume 1*, pp. 257–266, Bhubaneswar, India, December 2021.

[2] R. R. Cao and Q. X. Han, "Research on security threats and key technologies of internet of things," *Cyberspace Security*, vol. 11, no. 11, pp. 70–75, 2020.

[3] V. Chang, Y. Q. Mou, and Q. A. Xu, "The ethical issues of location-based services on big data and iot," in *Modern Industrial IoT, Big Data and Supply Chain: Proceedings of the IIoTBDSC 2020*, pp. 195–

205, Macao, Special Administration Region (SAR) of China, September 2021.

[4] G. Cormode, C. Procopiuc, D. Srivastava, E. T. Shen, and T. Yu, "Differentially private spatial decompositions," in *2012 IEEE 28th International Conference on Data Engineering*, pp. 20–31, Arlington, VA, USA, April 2012.

[5] C. Dwork, "Differential privacy," in *Automata, Languages and Programming: 33rd International Colloquium, ICALP 2006*, pp. 1–12, Venice, Italy, July 2006.

[6] C. Dwork, "Differential privacy: A survey of results," in *Theory and Applications of Models of Computation: 5th International Conference, TAMC 2008*, pp. 1–19, Xi'an, China, April 2008.

[7] C. Dwork, F. McSherry, K. Nissim, and A. Smith, "Calibrating noise to sensitivity in private data analysis," in *Theory of Cryptography: Third Theory of Cryptography Conference, TCC 2006*, pp. 265–284, New York, NY, USA, March 2006.

[8] J. Kaur, A. Agrawal, and R. A. Khan, "Encryfuscation: A model for preserving data and location privacy in fog based iot scenario," *Journal of King Saud University-Computer and Information Sciences*, vol. 34, no. 9, pp. 6808–6817, 2022.

[9] D. R. Li, Z. F. Shao, W. B. Yu, et al., "Public epidemic prevention and control services based on spatiotemporal big data makes cities smarter," *Geomatics and Information Science of Wuhan University*, vol. 45, no. 4, pp. 475–487, 2020.

[10] C. Y. Lin, "Suppression techniques for privacy-preserving trajectory data publishing," *Knowledge-Based Systems*, vol. 206, p. 106354, 2020.

[11] F. D. McSherry, "Privacy integrated queries: an extensible platform for privacy-preserving data analysis," in *Proceedings of the 2009 ACM SIGMOD International Conference on Management of data*, pp. 19–30, Providence Rhode Island USA, June-July 2009.

[12] W. Qardaji, W. N. Yang, and N. H. Li, "Differentially private grids for geospatial data," in *2013 IEEE 29th international conference on data engineering (ICDE)*, pp. 757–768, Brisbane, Australia, April 2013.

[13] K. M. Rodríguez, M. Bossy, R. Maftei, S. Shekarforush, and C. Henry, "New spatial decomposition method for accurate, mesh-independent agglomeration predictions in particle-laden flows," *Applied Mathematical Modelling*, vol. 90, pp. 582–614, 2021.

[14] X. Y. Shen, L. C. Wang, Q. Q. Pei, Y. Liu, and M. M. Li, "Location privacy-preserving in online taxi-hailing services," *Peer-to-Peer Networking and Applications*, vol. 14, pp. 69–81, 2021.

[15] R. S. A. Usmani, I. A. T. Hashem, T. R. Pillai, A. Saeed, and A. M. Abdullahi, "Geographic information system and big spatial data: A review and challenges," *International Journal of Enterprise Information Systems (IJEIS)*, vol. 16, no. 4, pp. 101–145, 2020.

[16] G. Y. Wang and G. Yang, "Survey on frequency estimation algorithms with local differential privacy," *Software Guide*, vol. 20, no. 1, pp. 226–233, 2021.

[17] J. H. Wei, Y. P. Lin, X. Yao, and J. Zhang, "Differential privacy-based location protection in spatial crowdsourcing," *IEEE Transactions on Services Computing*, vol. 15, no. 1, pp. 45–58, 2019.

[18] Y. Yan, Y. M. Cong, M. Adnan, and Q. Z. Sheng, "Statistics release and privacy protection method of location big data based on deep learning," *Journal on Communications*, vol. 43, no. 1, pp. 203–216, 2022.

[19] Y. Yan, X. Gao, M. Adnan, T. Feng, and P. S. Xie, "Differential private spatial decomposition and location publishing based on unbalanced quadtree partition algorithm," *IEEE Access*, vol. 8, no. 1, pp. 104775–104787, 2020.

[20] Y. Yan, X. H. Hao, and L. X. Zhang, "Hierarchical differential privacy hybrid decomposition algorithm for location big data," *Cluster Computing*, vol. 22, no. 4, pp. 9269–9280, 2019.

[21] Y. Yan, Z. C. Sun, A. Mahmood, F. Xu, Z.Y. Dong, and Q. Z. Sheng, "Achieving differential privacy publishing of location-based statistical data using grid clustering," *ISPRS International Journal of Geo-Information*, vol. 11, no. 7, p. 404, 2022.

[22] M. M. Yang, T. Q. Zhu, Y. Xiang, and W. L. Zhou, "Density-based location preservation for mobile crowdsensing with differential privacy," *Ieee Access*, vol. 6, pp. 14779–14789, 2018.

[23] X. J. Zhang, N. Fu, and X. F. Meng, "Towards spatial range queries under local differential privacy," *J. Comput. Res. Dev.*, vol. 57, no. 4, p. 847, 2020.

[24] Y. Zhao, D. Yuan, J. T. Du, and J. J. Chen, "Geo-ellipse-indistinguishability: community-aware location privacy protection for directional distribution," *IEEE Transactions on Knowledge and Data Engineering*, 2022.

[25] G. Q. Zhou, X. L. Tang, and S. Qin, "Adaptive grid decomposition algorithm based on standard deviation circle radius," *International Journal of Performability Engineering*, vol. 15, no. 8, p. 2145, 2019.

Biography

Yan Yan received the Ph.D. degree in control theory and control engineering from Lanzhou University of Technology, China. She is currently an Associate Professor at School of Computer and Communication, Lanzhou University of Technology. Her research interests include information security, privacy preserving technology, and differential privacy. She is a member of the IEEE and the China Computer Federation.

Yue Zhang is currently a master student of the School of Computer and Communication, Lanzhou University of Technology, China. She received the B.Eng. degree from North University of China in 2021. Her research interests include privacy preserving data publishing and information security.

Xingying Qian is currently a master student of the School of Computer and Communication, Lanzhou University of Technology, China. She received the B.Eng. degree from Tianjing University of Technology in 2020. Her research interests include machine learning and privacy, information security and differential privacy.

Analysis of One Lightweight Authentication and Matrix-based Key Agreement Scheme for Healthcare in Fog Computing

Lihua Liu and Yingqing Jia
(Corresponding author: Lihua Liu)

Department of Mathematics, Shanghai Maritime University
Haigang Ave 1550, Shanghai 201306, China
Email: liulh@shmtu.edu.cn

(Received Jan. 29, 2023; Revised and Accepted Nov. 26, 2023; First Online Dec. 24, 2023)

Abstract

We find the scheme [Peer Peer Netw. Appl., 12(4), 924–933, 2019] flawed, because: (1) both the doctor and the patient cannot finish their computations due to the false session-key extraction formula; (2) the involved 1-2-OT (oblivious transfer) protocol is not specified, and the doctor cannot decide which two target messages will be sent; (3) the preset shared key is only used for authentication, but it is also generally used for session-key extraction, which means the scheme can be significantly simplified; (4) each sensor needs to encrypt its identity number using RSA, but the RSA computation (2048-bit modulus) is usually considered overweight for a body sensor.

Keywords: Fog Computing; Healthcare; Key Agreement; Oblivious Transfer; Pairing

1 Introduction

In 2015, Cisco partnered with Microsoft, Intel, et al. to form the OpenFog Consortium. Its primary goals were to both promote and standardize fog computing. It is a decentralized infrastructure in which data, storage, computation and applications are located somewhere between the data source and the cloud. In a foggy environment, intelligence is at the local area network, and data is transmitted from endpoints to a fog gateway, where it's then transmitted to sources for processing.

In 2017, Khan *et al.* [11] discussed the problem of fog computing security. Kharel *et al.* [12] proposed an architecture for smart health monitoring system based on fog computing. After that, Aazam and *et al.* [1,18,22,24] surveyed fog computing architecture, evaluation, and future research directions. Lin *et al.* [1, 14, 16, 26] explored the deployment of fog computing systems at logistics centers in Industry 4.0 and other applications. Ma *et al.* [17] proposed a client-side deduplication scheme wiht ownership management in fog computing. Pan *et al.* [10, 20, 23] designed a secure smart card-based password authentication scheme, and studied malware detection and classification based on artificial intelligence. In 2021, Ali *et al.* [2] presented a clogging resistant secure authentication scheme for fog computing services. Battula *et al.* [3, 4, 15] presented a generic stochastic model for resource availability in fog computing environments.

Recently, Shen *et al.* [25] have presented a lightweight authentication and matrix-based key agreement scheme for healthcare in fog computing. Though the scheme is interesting, we find it is flawed because some computations are falsely specified. We also find the preset shared keys are only used for authentication. As we know, the shared keys can be directly used for keyed hash function to extract session keys. We would like to stress that the involved RSA computation is quite overweight for a body sensor, which is generally considered to be outsourced.

2 Review of the Key Agreement Scheme for Fog Computing

The scheme involves four entities: body sensors, the patient, the target doctor, and the cloud server. The patient uses sensors to measure medical data and sends those abnormal data to the target doctor, who evaluates his diseases and communicates with the patient. It consists of five phases: system initialization, SP lightweight authentication (between the last sensor and the patient), PD authentication (between the patient and the doctor), key agreement (between the patient and the doctor), and healthcare data uploading and downloading.

Let G be an additive group with a prime order q and a generator g, $e(\cdot, \cdot)$ be a bilinear map, η_1, η_2 be security parameters, H_1, H_2, H_3 be three hash functions. The preset key ξ_i is shared by the patient and his last body sensor. The preset key λ_i is shared by the patient and the target doctor. The scheme can be depicted as follows (Table 1).

Table 1: Shen *et al.*'s key agreement scheme

Sensor $\{\xi_i\}$	Patient $\{\xi_i, \lambda_i\}$	Doctor $\{\lambda_i\}$	Server
The sensors compute $d = RSA(id_{1,1}) \Vert \cdots \Vert RSA(id_{1,t})$. The last sensor decrypts d to get $d' = id_{1,1}$ $\Vert id_{1,2} \Vert \cdots \Vert id_{1,t}$, and	$\xleftarrow{\text{query}}$	Let $\bar{M}$ be the file to describe which doctor belongs to which field. The target doctor picks $a, b \in Z_q$ to set	
generates $R_i \in Z_q$, T_i. It computes $U_i = H_1(\xi_i \Vert T_i)$, $D_i = H_1(R_i \Vert id_{1,1}$ $\Vert id_{1,2} \Vert \cdots \Vert id_{1,t} \Vert T_i)$, $M_1 = U_i \oplus D_i$, $V_1 = h_1(D_i \Vert f_t \Vert T_i)$, where f_t is the connection time.	$U'_i = H_1(\xi_i \Vert T_i)$, $D'_i = U'_i \oplus M_1$, $V'_1 = H_1(D'_i \Vert f_t \Vert T_i)$. If $V'_1 = V_1$, pick T'_i, $R'_i \in Z_q$, compute $M_2 = (T'_i \Vert R'_i) \oplus U'_i$, $V_2 = H_1(p_i \Vert T'_i \Vert R'_i)$.	$SK = (x, y) = (\eta_2 a, \eta_2 b)$, $PK = (X, Y) = (g^{\eta_2 a}, g^{\eta_2 b})$. Let $\tilde{M}$ be the target doctor's description (=d_i). Pick $a \in G$ to compute	
$\xrightarrow{M_1 \Vert V_1 \Vert T_i}$	where p_i represents the i-th patient. $\xleftarrow{M_2 \Vert V_2}$	$A = aH_3(d_i)\eta_1$, $B = A^y$, $C = A^x B^{xH_3(\bar{M})}$.	
The last sensor computes $M'_2 = U_i \oplus M_2$, $V'_2 = H_1(p_i \Vert M'_2)$. Check that $V_2 = V'_2$.		For timestamp T_d, compute $M_5 = H_2(d_i \Vert \lambda_i \Vert T_d) \oplus H_2(\tilde{M})$, $M_6 = M_5 \oplus B$, $M_7 = M_5 \oplus A$, $M = M_5 \Vert M_6 \Vert M_7 \Vert H_3(\bar{M})$,	
	$M_8 = H_2(d_i \Vert \lambda_i \Vert T_d)$, $M'_5 = M_8 \oplus M_5$.	$V_3 = H_2(e(C, g))$, $V_4 = H_2(e(A, Y))$.	
	If $M'_5 = H_2(\tilde{M})$, compute $M'_6 = M_8 \oplus B$, $M'_7 = M_8 \oplus A$, $V'_3 = H_2(e(A, X)e(B, X)^{H_3(\bar{M})})$, $V'_4 = H_2(e(B, g))$.	$\xleftarrow{M, V_3, V_4, T_d}$	
	Check that $V'_3 = V_3, V'_4 = V_4$. Compute $m_{1,1} = H_2(V'_3)$, $m_{1,2} = H_3(\tilde{M})$, $m_{i,j} = H_3(p_i)$.	Compute $n_{1,1} = H_3(V_4)$, $n_{i,j} = H_3(d_i)$. $D = \begin{pmatrix} n_{1,1} & \cdots & n_{i,1} \\ \vdots & \ddots & \vdots \\ n_{1,j} & \cdots & n_{i,j} \end{pmatrix}$	
	$P = \begin{pmatrix} m_{1,1} & \cdots & m_{i,1} \\ \vdots & \ddots & \vdots \\ m_{1,j} & \cdots & m_{i,j} \end{pmatrix}$ Let $X_1 = \begin{pmatrix} m_{1,1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ m_{1,j} & \cdots & 0 \end{pmatrix}$ $\cdots, X_e = \begin{pmatrix} 0 & \cdots & m_{i,1} \\ \vdots & \ddots & \vdots \\ 0 & \cdots & m_{i,j} \end{pmatrix}$ s.t., $P = X_1 + X_2 + \cdots + X_e$.		
	For each $j = 1, \cdots, e$, pick $i \in [1, p]$, generate the matrixes $M_1, \cdots, M_p$, where $M_i = X_j$ (i is only known to the patient). $\xrightarrow{M_1, \cdots, M_p}$	For each $i = 1, \cdots, e$, compute $N_i = M_i D_i - R_j$, where R_j is a random matrix, $D_1 = (n_{1,1}, \cdots, n_{1,j})$, $\cdots, D_i = (n_{i,1}, \cdots, n_{i,j})$.	
	Recover $Q_i = X_j D_i - R_j$.	$\xleftarrow[\text{1-2-OT}]{N_i}$	
	Compute the session-key $R_a = H_2(\sum_{i,j=1}^{e}(X_j D_i - R_j))$. Use the key to transfer the healthcare data.		
	$\xrightarrow{\text{upload the encrypted healthcare data to the cloud server}}$		
		Download the healthcare data and decrypt it with the session-key.	$\xleftarrow{\text{healthcare data}}$

3 Analysis

◇ *It fails to keep consistency.* For example (see §4.3):

> The doctor generates randomly $a, b \in Z_q^*$, and computes $SK = (x, y) = (\eta_2 a, \eta_2 b), X = g^{\eta_2 x}$ and $Y = g^{\eta_2 y}$, where public keys are $PK = (X, Y)$. Let $a \in G$, compute $A = aH_3(d_i)\eta_1, B = A^y, C = A^x B^{xH_3(\bar{M})}$.

It falsely requires either $a \in Z_q^*$ or $a \in G$. In fact, the group G is derived from an elliptic curve and used for construction of the bilinear map $e(\cdot, \cdot)$.

By the latter specification, we have

$$\begin{aligned} V_3 &= H_2(e(C, g)) = H_2(e(A^x B^{xH_3(\bar{M})}, g)), \\ V_3' &= H_2(e(A, X)e(B, X)^{H_3(\bar{M})}) \\ &= H_2(e(A^{\eta_2 x}, g)e(B^{\eta_2 x H_3(\bar{M})}, g)) \\ &= H_2(e(A^{\eta_2 x} B^{\eta_2 x H_3(\bar{M})}, g)) \\ &\neq V_3. \\ V_4 &= H_2(e(A, Y)) = H_2(e(A, g)^{\eta_2 y}), \\ V_4' &= H_2(e(B, g)) = H_2(e(A^y, g)) = H_2(e(A, g)^y) \\ &\neq V_4. \end{aligned}$$

To revise, it could specify that

> The doctor randomly generates $x, y \in Z_q^*$, and computes $SK = (x, y), X = g^x, Y = g^y$, where public keys are $PK = (X, Y)$. Let $a \in Z_q^*$, compute $A = g^a, B = A^y, C = A^x B^{xH_3(\bar{M})}$.

In fact, the computations related to A require only that A is a random element in the group G. Due to the bilinear property, we have

$$\begin{aligned} V_3' &= H_2(e(A, X)e(B, X)^{H_3(\bar{M})}) \\ &= H_2(e(A^x, g)e(B^{xH_3(\bar{M})}, g)) \\ &= H_2(e(A^x B^{xH_3(\bar{M})}, g)) = H_2(e(C, g)) = V_3. \\ V_4' &= H_2(e(B, g)) = H_2(e(A^y, g)) \\ &= H_2(e(A, g^y)) = H_2(e(A, Y)) = V_4. \end{aligned}$$

In the same section (§4.3), it writes: "compute $M_6' = M_8 \oplus B, M_7' = M_8 \oplus A$." Clearly, it is also falsely specified, because B and A are not directly transferred by the doctor. It should be revised as "compute $B = M_6' \oplus M_8, A = M_7' \oplus M_8$", because M_6, M_7 are directly concatenated in M, and M is simply transferred. Note that the patient needs to use the preset shared key λ_i to get M_8.

By the way, it falsely specifies that G is an additive group. Actually, G should be a multiplicative group because of the latter notations, such as $X = g^{\eta_2 x}$.

◇ *The doctor and the patient cannot finish their computations in the key agreement phase*, because the session-key extraction formula is falsely specified. As we see, the resulting session key between the patient and the target doctor is

$$R_a = H_2\left(\sum_{i,j=1}^{e}(X_j D_i - R_j)\right).$$

The formula depends on the indexes i and j. It emphasizes that "the number i is a secret number that only is known to the patient, the doctor shouldn't determine which X_j is M_i." That is to say, the doctor cannot finish the computation because the true index i is not known to him.

Likewise, the patient cannot finish the computation because the R_j is a random matrix picked by the doctor and not known to the patient. Even if the formula is revised as

$$R_a = H_2\left(\sum_{i=1}^{e} Q_i\right),$$

we find the patient cannot finish the computation because of the next shortcoming.

◇ *The involved 1-2-OT protocol is not specified.* We find the doctor cannot decide which two target messages will be sent. Even if the padding messages are specified, the receiver (patient) cannot recover all Q_i's, due to the obliviousness of these transfers.

As we know, oblivious transfer (OT) was introduced by Rabin [21], which can be used to design secure multi-party computation schemes. In the basic model, the sender has a secret m and wants to have the receiver obtain m with a 50:50 chance, while the sender cannot decide whether m was retrieved. 1-out-of-2 oblivious transfer, introduced by Even, Goldreich and Lempel [8], is a generalization of the original OT. In the model, the sender has two secrets m_1 and m_2 and wants to have the receiver obtain only one of them at the receiver's choice, while the sender cannot decide which secret was retrieved.

We want to stress that the paradigm of oblivious transfer [19] is more complicated than the general key agreement. Intuitively, it is a bad choice to design a key agreement protocol by means of oblivious transfer (see the discussions, [5]).

◇ *The preset shared key between the patient and the target doctor is only used for authentication.* Actually, it is also generally used for keyed hash function to extract session-keys. That is to say, the scheme can be greatly simplified. We refer to [9, 13] for the applications of keyed hash function.

◇ *The RSA computation is quite overweight for a body sensor.* As we see, each sensor needs to encrypt its identity number using RSA. But the RSA computation (2048-bit modulus) is generally considered to be hard for a body sensor, and to be outsourced [6, 7].

4 Conclusion

We show that Shen *et al.*'s key agreement scheme has some falsely specified computations. We want to stress that it is not a good choice to design a lightweight key agreement scheme by means of oblivious transfer, and the RSA computation is somewhat overweight for a body sensor.

Acknowledgment

We thank the National Natural Science Foundation of China (61411146001). We are grateful to the reviewers for their valuable suggestions.

References

[1] M. Aazam, S. Zeadally, and K. Harras, "Fog computing architecture, evaluation, and future research directions," *IEEE Commun. Mag.*, vol. 56, no. 5, pp. 46–52, 2018.

[2] Z. Ali and *et al.*, "A clogging resistant secure authentication scheme for fog computing services," *Comput. Networks*, vol. 185, p. 107731, 2021.

[3] K. Anggriani, N. Wu, and M. S. Hwang, "Research on data hiding schemes for AMBTC compressed images," *Int. J. Netw. Secur.*, vol. 24, no. 6, pp. 1114–1123, 2022.

[4] S. Battula and *et al.*, "A generic stochastic model for resource availability in fog computing environments," *IEEE Trans. Parallel Distributed Syst.*, vol. 32, no. 4, pp. 960–974, 2021.

[5] Z. Cao and L. Liu, "Improvement of greenhohenberger adaptive oblivious transfer: A review," *Int. J. Netw. Secur.*, vol. 17, no. 4, pp. 454–462, 2015.

[6] G. Crescenzo and *et al.*, "Computing multiple exponentiations in discrete log and RSA groups: From batch verification to batch delegation," in *2017 IEEE Conference on Communications and Network Security, CNS*, pp. 531–539. IEEE, 2017.

[7] G. Crescenzo and *et al.*, "Secure delegation to a single malicious server: Exponentiation in RSA-type groups," in *7th IEEE Conference on Communications and Network Security, CNS*, pp. 1–9. IEEE, 2019.

[8] S. Even, O. Goldreich, and A. Lempel, "A randomized protocol for signing contracts," *Communications of the ACM*, vol. 28, no. 6, pp. 637–647, 1985.

[9] P. Gope and B. Sikdar, "Privacy-aware authenticated key agreement scheme for secure smart grid communication," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 3953–3962, 2019.

[10] L. Huang, C. Chang, and M. S. Hwang, "Research on malware detection and classification based on artificial intelligence," *Int. J. Netw. Secur.*, vol. 22, no. 5, pp. 717–727, 2020.

[11] S. Khan, S. Parkinson, and Y. Qin, "Fog computing security: a review of current applications and security solutions," *J. Cloud Comput.*, vol. 6, p. 19, 2017.

[12] J. Kharel, H. Reda, and S. Shin, "An architecture for smart health monitoring system based on fog computing," *J. Commun.*, vol. 12, no. 4, pp. 228–233, 2017.

[13] P. Kumar and *et al.*, "Lightweight authentication and key agreement for smart metering in smart energy networks," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 4349–4359, 2019.

[14] C. Lin and J. Yang, "Cost-efficient deployment of fog computing systems at logistics centers in industry 4.0," *IEEE Trans. Ind. Informatics*, vol. 14, no. 10, pp. 4603–4611, 2018.

[15] Y. C. Lu and M. S. Hwang, "A cryptographic key generation scheme without a trusted third party for access control in multilevel wireless sensor networks," *Int. J. Netw. Secur.*, vol. 24, no. 5, pp. 959–964, 2022.

[16] N. Luong and *et al.*, "A machine-learning-based auction for resource trading in fog computing," *IEEE Commun. Mag.*, vol. 58, no. 3, pp. 82–88, 2020.

[17] H. Ma, G. Tian, and L. Zhang, "Anti-leakage client-side deduplication with ownership management in fog computing," *Int. J. Netw. Secur.*, vol. 22, no. 1, pp. 24–35, 2020.

[18] C. Mouradian and *et al.*, "A comprehensive survey on fog computing: State-of-the-art and research challenges," *IEEE Commun. Surv. Tutorials*, vol. 20, no. 1, pp. 416–464, 2018.

[19] M. Naor and B. Pinkas, "Computationally secure oblivious transfer," *Journal of Cryptology*, vol. 18, no. 1, pp. 1–35, 2005.

[20] H. Pan, H. Yang, and M. Hwang, "An enhanced secure smart card-based password authentication scheme," *Int. J. Netw. Secur.*, vol. 22, no. 2, pp. 358–363, 2020.

[21] M. Rabin, *How to exchange secrets by oblivious transfer, Technical Report TR-81.* 1981.

[22] A. Rahmani and *et al.*, "Exploiting smart e-health gateways at the edge of healthcare Internet-of-Things: A fog computing approach," *Future Gener. Comput. Syst.*, vol. 78, pp. 641–658, 2018.

[23] J. Saadatmandan and A. Rahimi, "A secure authenticated key agreement protocol for application at digital certificat," *Int. J. Netw. Secur.*, vol. 22, no. 2, pp. 250–256, 2020.

[24] O. Salman and *et al.*, "Iot survey: An SDN and fog computing perspective," *Comput. Networks*, vol. 143, pp. 221–246, 2018.

[25] J. Shen and *et al.*, "Lightweight authentication and matrix-based key agreement scheme for healthcare in fog computing," *Peer Peer Netw. Appl.*, vol. 12, no. 4, pp. 924–933, 2019.

[26] F. Tseng and *et al.*, "A lightweight autoscaling mechanism for fog computing in industrial applications," *IEEE Trans. Ind. Informatics*, vol. 14, no. 10, pp. 4529–4537, 2018.

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.

Yingqing Jia is currently pursuing her master degree from Department of Mathematics at Shanghai Maritime University. Her research interests include information theory and applied mathematics.

Research on Information Images in Digital Media Using an Improved Digital Watermarking Algorithm

Ying Sun, Guofeng Ma, and Shumei Zhao
(Corresponding author: Ying Sun)

Zhengzhou Railway Vocational and Technical College, China
No. 81 yard, Qianjin Road, Zhongyuan District, Zhengzhou, Henan 450000, China
Email: sying03@outlook.com

(Received Oct. 3, 2022; Revised and Accepted Dec. 1, 2023; First Online Dec. 24, 2023)

Abstract

The information conveyed by digital media is susceptible to attacks during transmission, making it essential to safeguard it using digital watermarking algorithms. To enhance the security of the transmission process, this paper briefly analyzed digital watermarking algorithms. Subsequently, an enhanced digital watermarking algorithm was developed. The combination of discrete wavelet transform and singular value decomposition was optimized for the size of the embedding factor using the whale optimization algorithm (WOA). Experimental tests were conducted using Lena images. The results demonstrated that under a 10% JPEG lossy compression attack, the algorithm achieved a peak signal-to-noise ratio (PSNR) of 30.125 dB and a normalized correlation (NC) coefficient of 0.789. When subjected to a 0.05 salt-and-pepper noise attack, the algorithm yielded a PSNR of 18.654 and an NC coefficient of 0.925. Moreover, even when exposed to shear attacks at various locations, the algorithm maintained an NC coefficient above 0.8, albeit with a noticeable decline in PSNR. This improved algorithm had the highest NC coefficient in a comparative analysis with other algorithms. These results demonstrate the reliability of the improved algorithm, making it a viable option for securing digital media information images in real-world applications.

Keywords: Digital Media; Digital Watermarking Algorithm; Information Image; Whale Optimization Algorithm

1 Introduction

The continuous advancement and evolution of computer technology have significantly expanded the applications of digital media information, catering to people's increasing demands for accessing and sharing information. However, while digital media enhances convenience in our lives, the issue of unauthorized parties leaking, tampering with, or destroying digital media information is growing more severe. To ensure the security of digital media information, various information hiding techniques have emerged [11], and one of them is digital watermarking algorithms [3,4]. These algorithms can invisibly embed a special symbol or image into digital media, which can be used to authenticate the authenticity and integrity of the information [2,20].

Kumar *et al.* [13] employed the bat algorithm to optimize the semi-blind watermarking scheme and found it had good perceptual quality and robustness through experiments. Yamni *et al.* [26] introduced a method for audio watermarking that utilizes a combination of mixed linear-nonlinear coupled mapping lattice chaotic system, discrete tchebichef moment transform, and discrete wavelet transform (DWT). The approach achieved superior robustness and payload capacity in empirical tests.

Shivdeep *et al.* [22] developed a color watermark embedding method that modifies pixel values in the carrier image based on the resemblance between the carrier image and the watermark and found a high throughput through experiments. Mousavi *et al.* [17] designed an approach that utilized discrete cosine transform (DCT) and employed a genetic algorithm to extract the optimal graph structure. They found that the method was more robust. Although many digital watermarking algorithms have been applied to digital media information images, there remain challenges, particularly in terms of resistance to external attacks. Thus, continuous and in-depth research on digital watermarking algorithms is essential to enhance their applicability in digital media. This paper introduces an improved watermarking algorithm for digital media information images, offering a novel approach to enhancing information transmission security and providing theoretical support for further advancements in digital watermarking algorithms.

2 Improved Digital Watermarking Algorithm

2.1 Digital Watermarking Algorithm

Digital watermarking algorithms applied to digital media information images mainly include the following two types [10, 24, 25].

Space domain algorithm

It directly modifies the original data to embed the watermarking information, including the least significant bit algorithm [15], patchwork algorithm [21], etc. This type of algorithm is easy to execute and possesses a good invisibility, but in the face of the attack, it is extremely easy to be affected.

Transform domain algorithm

It refers to the method of processing the carrier image by superimposing watermark information on the transform domain, including DCT [6], DWT [1], etc. This type of algorithm has higher security compared to the spatial domain algorithms, making it more extensively used.

In the face of digital watermarking algorithms, the main evaluation methods are shown below.

Invisibility

The watermark embedded by the digital watermarking algorithm should not interfere with the visualization of the information image, i.e., it should have good concealment and invisibility, which is generally measured by the peak signal-to-noise ratio (PSNR) [23]:

$$PSNR = 10 \log \frac{D^2 MN}{\sum_{x=1}^{M} \sum_{y=1}^{N} (I(x,y) - I_W(x,y))^2}$$

where M and N are image sizes, D is the peak value of the signal, and $I(x,y)$ and $I_W(x,y)$ are images before and after the incorporation of the watermark.

Robustness

The information carried in a digital watermark should be reliably and correctly identified even under attack, which is generally measured using the normalized correlation (NC) coefficient [5]:

$$NC(W,W') = \frac{\sum\sum W(i,j) \cdot W'(i,j)}{\sqrt{\sum\sum W^2(i,j)}\sqrt{\sum\sum W'^2(i,j)}}$$

where $W(i,j)$ refers to the original watermark and $W'(i,j)$ refers to the extracted watermark.

2.2 Discrete Wavelet Transform

Wavelet transform has good applications in image processing [12], speech analysis [18], and other fields. In the digital watermarking algorithm, the commonly used is DWT [1], and its principle is to do discretization of the wavelet transform value. The information image of the digital media is divided into vertical, horizontal, low-frequency, and diagonal bands after the DWT processing. Generally, L is used to indicate the low-pass filter, and H is used to indicate the high-pass filter. As an example, Figure 1 illustrates the schematic diagram of a three-layer DWT.

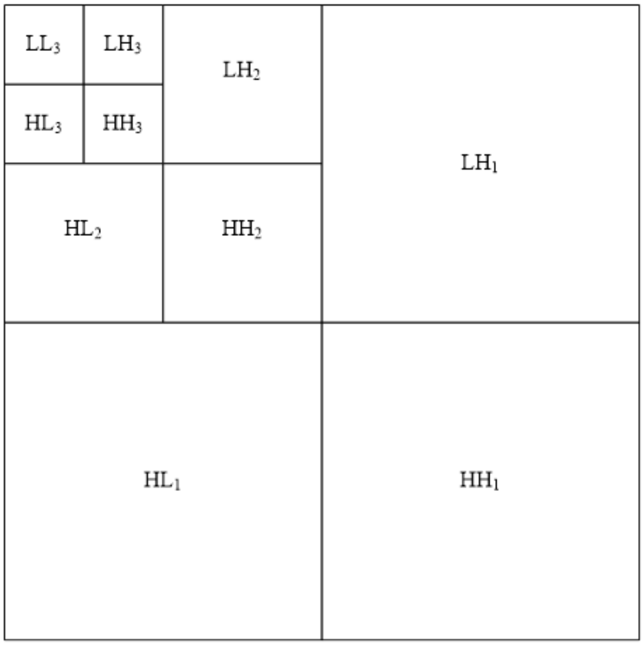


Figure 1: Schematic diagram of three-layer DWT decomposition of information image

As shown in Figure 1, LL3 is a low-frequency subband, LH1, LH2, and LH3 represent horizontal high-frequency subbands, HL1, HL2, and HL3 represent vertical high-frequency subbands, and HH1, HH2, and HH3 represent diagonal high-frequency subbands. The low-frequency sub-band is the frequency band closest to the initial image, and embedding watermarks in these areas can achieve greater robustness.

2.3 Singular Value Decomposition

Singular value decomposition (SVD) has a very wide range of applications in the fields of pattern recognition [7] and signal analysis [9]. For information image A with a size of $m \times n$, let $A \in R^{m \times n}$, then there are orthogonal matrices:

$$\begin{aligned} U &= [u_1, u_2, \cdots, u_m] \in R^{m \times n}, \\ V &= [v_1, v_2, \cdots, v_m] \in R^{m \times n}. \end{aligned}$$

SVD is performed on information image A:

$$A = USV^T = \sum_{i=1}^{N} \delta_i u_i v_i^T,$$

where S is the diagonal matrix, δ_i is the singular value after decomposition, u_i and v_i are the left and right singular vectors of δ_i, $Au_i = \delta_i v_i$, $Av_i = \delta_i u_i$.

The advantage of applying SVD in digital watermarking algorithms is that SVD has excellent stability [27]. Even if the image receives slight perturbation, it will not be affected too much, and U and V have good resistance to geometric attacks.

2.4 Whale Optimization Algorithm

In the SVD-based digital watermarking algorithm, the quality of the information image and its resistance to attacks depend on the choice of embedding factors. However, existing algorithms often rely on quantitative embedding factors that may not be well-suited to every image. To make the algorithm more usable, this paper introduces the application of the whale optimization algorithm (WOA) for the selection of embedding factors. WOA is a whale-inspired optimization algorithm [16], which has been applied in data prediction [19] and parameter optimization [8]. In the hunting process, the whale will move towards the most favorable position in the current to encircle the prey. The position update formula of the whale can be written as:

$$\begin{aligned} D &= |C \cdot X^*(t) - X(t)|, \\ X(t+1) &= X^*(t) - A \cdot D, \end{aligned}$$

where D refers to the distance between whale and prey, $X^*(t)$ is the most optimal position for the whale at present, $X(t)$ refers to the whale position currently, and A and C are convergence factor and swing factor, respectively. The calculation formulas of A and C are:

$$\begin{aligned} A &= 2ar - a, \\ C &= 2r. \end{aligned}$$

In the above equations, a decreases linearly from 2 to 0 during iteration, and r is a random number in [0,1]. The whale then herds its prey by bubble nets, at which point the position update equation can be written as:

$$X(t+1) = D \cdot e^{bl} \cdot \cos(2\pi l) + X^*(t).$$

It is assumed that the whale chooses its attack way with a 50% probability. It can be expressed as:

$$X(t+1) = \begin{cases} X^*(t) - A \cdot D, & p < 0.5 \\ D \cdot e^{bl} \cdot \cos(2\pi l) + X^*(t) & p \geq 0.5 \end{cases}$$

where b is a constant, which is 1 by default, and l is a random number in [-1, 1]. In addition, whales also have the ability to randomly search for optimization in order to avoid getting stuck in local optima. The position update formula can be written as:

$$\begin{aligned} X(t+1) &= Xrand(t) - A \cdot D, \\ D &= |C \cdot Xrand(t) - X(t)|, \end{aligned}$$

where $Xrand(t)$ is the position of one random whale.

The WOA is applied to determine the embedding factor, and the process is illustrated in Figure 2.

As shown in Figure 1, the whale population is initialized first. Then, through the watermark embedding algorithm, the PSNR values of the initial and watermarked images are obtained, and different attacks are applied to the watermarked image. After generating the attacked watermarked image, the watermark is extracted, and the NC coefficient is calculated. The fitness function of the WOA is:

$$objective = PSNR(A, A') + \phi \times [NC(W, W') + \sum_{i=1}^{n} NC(W, W_i')],$$

where $PSNR(A, A')$ refers to the PSNR between the initial image and the one with watermark, $NC(W, W')$ refers to the NC coefficient of the initial and extracted watermarks, $NC(W, W_i')$ is the NC coefficient of the initial and extracted watermarks after attacking, and $\phi = 10$ is the weighting factor.

2.5 Improved Digital Watermarking Algorithm

2.5.1 Watermark Preprocessing

In order to prevent watermarks from being destroyed, it is generally necessary to do pre-processing before embedding, i.e., information encryption. The Arnold transform [14] is used, and it is expressed as:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} (\bmod N), \tag{1}$$

where (x, y) and (x', y') are the coordinates of the pixel of the watermarked image before and after scrambling, and N is the image order.

For recovery, Arnold contrary transformation can be used:

$$\begin{pmatrix} x'' \\ y'' \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} (\bmod N), \tag{2}$$

where (x'', y'') is the coordinates of the pixel after contrary transformation.

2.5.2 Watermark Embedding Process

1) Suppose there is a $M \times M$ carrier image called A. Three-layer DWT decomposition is performed on it to obtain low-frequency subband LL3.

2) Suppose there is a $N \times N$ watermark image called W. Encrypted watermark W' is obtained after Arnold transformation.

3) SVD is applied on LL3 to get singular value S:

$$[U, S, V] = SVD(LL_3).$$

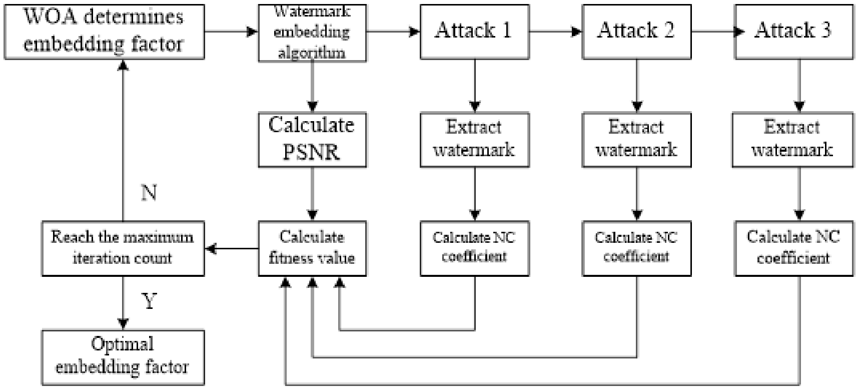


Figure 2: Flowchart of WOA for determining embedding factor

4) Singular value S_W is obtained by applying SVD on W':

$$[U, S, V] = SVD(W').$$

5) Low-frequency components are embedded with watermarks following the addition rule:

$$S' = S + \rho \times S_W,$$

where ρ is the embedding factor, obtained by the WOA.

6) Inverse discrete wavelet transform (IDWT) is performed on the subbands containing watermarks to get the watermarked image.

2.5.3 Watermark Extraction Process

1) A three-layer DWT decomposition is performed on the original and watermarked images to obtain LL3 and LL3'.

2) SVD is applied on LL3 and LL3' to obtain singular values S and S':

$$\begin{cases} [U, S, V] = SVD(LL3) \\ [U', S', V'] = SVD(LL3') \end{cases}$$

3) Singular value S'_W of the watermark is calculated:

$$S'_W = \frac{S' - S}{\rho}$$

4) The watermark image is reconstructed:

$$W' = U_W \times S'_W \times V_W^T.$$

5) Decryption is performed by Arnold contrary transformation to get the watermark image.

3 Experimental Results and Analysis

3.1 Experimental Setup

In the standard test image library, Lena was selected as the carrier image (Figure 3), and a 64 × 64 binary image was used as the watermark image (Figure 4).

Figure 3: Original carrier image

The experimental environment was the Windows 10 operating system and the MATLAB 2015b platform. DWT was implemented using the db1 basis belonging to Daubechies type. For watermark encryption, the scrambling number of Arnold transform was set as 30. In the WOA, the population size was set at 10, and the maximum iteration count was 100. The performance of the enhanced watermarking algorithm was evaluated based on the PSNR and NC coefficient.

3.2 Analysis of Results

In the absence of any attack, the improved algorithm achieved a PSNR of 38.124 dB and a NC coefficient of 1.000. Then, an analysis was conducted on the algorithm

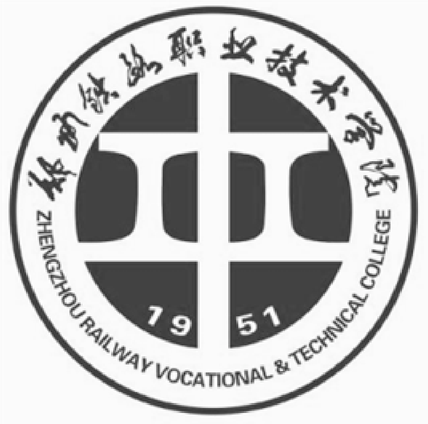


Figure 4: A watermark image

performance under different attacks. Firstly, under various levels of JPEG compression attack, the performance the improved digital watermarking algorithm performs is shown in Table 1.

Table 1: Performance against JPEG compression attack

Compression factor/%	PSNR/dB	NC coefficient
90	36.232	1.000
80	36.125	0.999
70	35.114	0.995
60	34.658	0.992
50	32.215	0.987
30	31.741	0.961
10	30.125	0.789

Table 1 reveals that when subjected to various levels of JPEG lossy compression attacks, the PSNR and NC coefficient exhibited varying degrees of reduction. JPEG lossy compression is commonly employed in data size reduction during information preservation and transmission. Therefore, a digital watermarking algorithm must possess resilience against this type of attack. As observed in the results, a smaller compression factor signified more substantial compression, resulting in a lower PSNR and an NC coefficient. For instance, when the compression factor was set at 10%, the improved algorithm yielded a PSNR of 30.125 dB, representing a decrease of 20.98% compared to the non-attacked scenario, and an NC coefficient of 0.789, indicating a 21.1% reduction compared to the unattacked condition. Despite the noticeable decrease, the algorithm remained practical for real-world use.

The performance of the improved digital watermarking algorithm under salt-and-pepper noise attack of different intensities is presented in Table 2.

As per the data in Table 2, the algorithm exhibited a noteworthy decrease in PSNR and NC coefficient with increasing noise intensity. At an intensity of 0.05, the PSNR was 18.654 dB, and the NC coefficient was 0.925. These values were significantly lower compared to the non-attacked scenario; however, the image quality remained at a recognizable level, demonstrating the algorithm's effectiveness in mitigating salt-and-pepper noise attacks.

Table 2: Resistance to pretzel noise attack

Intensity	PSNR/dB	NC coefficient
0.001	34.897	0.992
0.003	30.658	0.964
0.005	28.121	0.987
0.01	25.784	0.974
0.05	18.654	0.925

The performance of the improved digital watermarking algorithm under cropping attack at different locations (cropping size: 1/4) is shown in Table 3.

Table 3: Cropping attack resistance

Location	PSNR/dB	NC coefficient
Upper left	11.565	0.854
Lower left	9.872	0.905
Upper right	11.957	0.844
Lower right	12.112	0.912
Middle	11.067	0.846

Cropping attacks indeed result in image loss and are considered menacing. As shown in Table 3, the algorithm was most susceptible to cropping attacks compared to the previous two types of attacks. The PSNR experienced a substantial decline, approximately 10 dB, due to image loss. However, the NC coefficient remained high, indicating that the extracted watermark image was not significantly distorted.

To further determine the performance of the improved algorithm, experiments were conducted on the same carrier image and the same embedding factor to compare the improved algorithm with the methods in literature [29] and literature [28]. The NC coefficients of different methods under different attacks are displayed in Table 3.

According to the data shown in Table 4, the improved algorithm had the highest NC coefficient, i.e., 1.000, under non-attacked conditions. Under a 60% JPEG compression attack, the NC coefficients for the methods described in literature [29] and literature [28] were 0.909 and 0.999. These findings demonstrated that the performance of the approach described in literature [29] was relatively weak when facing JPEG compression attacks. Under a 0.02 salt-and-pepper noise attack, both the improved algorithm and the method used in literature [29] attained an NC coefficient of 0.999, while the method used in lit-

Table 4: Performance comparison with other digital watermarking algorithms

	Literature [29]	Literature [28]	The improved algorithm
Unattacked	0.991	0.985	1.000
60% JPEG compression	0.909	0.999	0.999
0.02 pepper noise	0.999	0.963	0.999
Gaussian low-pass filtering	0.999	0.930	0.999
Scaling attack	0.802	0.817	0.985
Cropping attack	0.856	0.935	0.999

erature [28] achieved a value of 0.963. These results suggested that the method used in literature [28] was somewhat less resilient against salt-and-pepper noise attacks. In the context of scaling and cropping attacks, both the methods used in literature [29] and literature [28] exhibited a significant decrease in NC coefficients. However, the improved algorithm maintained an NC coefficient of 0.985 under the scaling attack and 0.999 under the cropping attack. A comprehensive analysis demonstrated that the improved algorithm consistently maintained a high NC value when confronted with different types of attacks, ensuring clear extraction of watermarks. These findings confirmed the robustness of the improved algorithm.

4 Conclusion

In this paper, an improved digital watermarking algorithm was developed for digital media information images, and its effectiveness was validated through extensive experimental analysis. The experiments demonstrated that the improved digital watermarking algorithm consistently maintained a high PSNR and NC coefficient when subjected to various types of attacks. Moreover, it outperformed the other algorithms, affirming the reliability and robustness of this algorithm. Consequently, this algorithm can be applied to real digital media information images to enhance the security of information transmission.

References

[1] H. Barouqa, A. Al-Haj, “Watermarking e-government document images using the discrete wavelets transform and Schur decomposition,” in *7th International Conference on Information Management (ICIM’21)*, pp. 102-106, 2021.

[2] C. C. Chang, K. F. Hwang, M. S. Hwang, “Robust authentication scheme for protecting copyrights of images and graphics”, *IEE Proceedings-Vision, Image and Signal Processing*, vol. 149, no. 1, pp. 43-50, 2002.

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

[4] C. C. Chang, K. F. Hwang, M. S. Hwang, “A block based digital watermarks for copy protection of images,” in *Asia-Pacific Conference on Communications*, vol. 2, pp. 977-980, 1999.

[5] Y. Chen, H. Wan, M. Zou, “An unsupervised unimodal registration method based on Wasserstein Gan,” *Journal of Southern Medical University*, vol. 41, no. 9, pp. 1366-1373, 2021.

[6] W. Diaztary, D. Atmajaya, F. Umar, Purnawansyah, Harlinda, S. M. Abdullah, “Tiny encryption algorithm on discrete cosine transform watermarking,” in *East Indonesia Conference on Computer and Information Technology*, 2021.

[7] J. Gai, Y. Hu, “Research on fault diagnosis based on singular value decomposition and fuzzy neural network,” *Shock and Vibration*, vol. 2018, no. pt.3, pp. 1-7, 2018.

[8] Q. Guo, L. Gao, X. Chu, H. Sun, “Parameter identification for static var compensator model using sensitivity analysis and improved whale optimization algorithm,” *CSEE Journal of Power and Energy Systems*, vol. 8, no. 2, pp. 535-547, 2022.

[9] Y. Hu, J. Huang, Z. Tian, S. Pan, “Ground microseismic data denoising based on single-channel singular value decomposition and amplitude ratio,” *Geophysical Prospecting for Petroleum*, vol. 58, no. 1, pp. 43-52, 62, 2019.

[10] M. S. Hwang, K. F. Hwang, C. C. Chang, “A time-stamping protocol for digital watermarking”, *Applied Mathematics and Computation*, vol. 169, pp. 1276–1284, 2005.

[11] M. M. Iqbal, U. Khadam, K. J. Han, J. Han, S. Jabbar, “A robust digital watermarking algorithm for text document copyright protection based on feature coding,” in *15th International Wireless Communications & Mobile Computing Conference (IWCMC’19)*, pp. 1940-1945, 2019.

[12] S. P. Jakhar, A. Nandal, A. Dhaka, B. Jiang, L. Zhou, V. N. Mishra, “Erratum: Fractal feature based image resolution enhancement using wavelet–fractal transformation in gradient domain,” *Journal of Circuits, Systems and Computers*, vol. 32, no. 2, pp. 1-26, 2023.

[13] C. Kumar, P. Sivananthamaitrey, P. R. Kumar, “An optimized semi-blind watermarking technique

for digital images using bat algorithm," in *International Conference on Computing, Communication and Power Technology (IC3P'22)*, pp. 167-171, 2022.

[14] L. Kumar, K. U. Singh, "A secure image watermarking scheme based on DWT, SVD and Arnold transform," *IOP Conference Series: Materials Science and Engineering*, vol. 1099, no. 1, pp. 1-13, 2021.

[15] R. R. A. Lubis, S. M. Hardi, M. Zarlis, I. Jaya, J. T. Tarigan, "Analysis on combination of watermarking algorithm: Modified least significant bit algorithm with least significant bit+1," *Journal of Physics: Conference Series*, vol. 1235, no. 1, pp. 1-6, 2019.

[16] A. Mohammadbeigi, A. Maroosi, M. Hemmati, "Optimal chiller loading for energy conservation using a hybrid whale optimization algorithm based on population membrane systems," *International Journal of Modelling & Simulation*, vol. 42, no. 1/2, pp. 101-116, 2022.

[17] M. R. Mousavi, A. Naghsh, "Robust digital image watermarking method using graph-based transform (GBT) and genetic algorithm," *Journal of Intelligent Procedures in Electrical Technology*, vol. 10, no. 39, pp. 13-22, 2019.

[18] D. S. Putra, Y. U. W. Weru, Fardiansyah, Fitriady, Fakhruddin, Z. Yahya, "Pattern recognition of facial electromyography (FEMG) signal for aceh language speech using Nave Bayes and learning vector quantization (LVQ)," *IOP Conference Series: Materials Science and Engineering*, vol. 1062, pp. 1-8, 2021.

[19] M. A. A. E. S. Rostum, H. M. M. Moustafa, I. E. S. Ziedan, A. A. Zamel, "A combined effective time series model based on clustering and whale optimization algorithm for forecasting smart meters electricity consumption," *International Journal for Computation and Mathematics in Electrical and Electronic Engineering*, vol. 41, no. 1, pp. 209-237, 2022.

[20] A. V. Sidorenko, I. V. Shakinko, "The digital watermarking algorithm using discrete chaotic maps," *System Analysis and Applied Information Science*, no. 2, pp. 72-76, 2020.

[21] C. Shi, H. Wang, X. Li, "Learned dictionaries-based watermarking for speech authentication," in *IEEE 5th International Conference on Cloud Computing and Big Data Analytics (ICCCBDA'20)*, pp. 504-513, 2020.

[22] Shivdeep, S. Ghosh, H. Rahaman, "A new digital color image watermarking algorithm with its FPGA and ASIC implementation," in *International Symposium on Devices, Circuits and Systems (ISDCS'20)*, pp. 1-6, 2020.

[23] M. Wang, S. Qi, Y. Wu, Y. Sun, R. Chang, H. Pang, W. Qian, "CE-NC-VesselSegNet: supervised by contrast-enhanced CT images but utilized to segment pulmonary vessels from non-contrast-enhanced CT images," *Biomedical Signal Processing and Control*, vol. 82, pp. 1-11, 2023.

[24] C. C. Wu, S. J. Kao, W. C. Kuo, M. S. Hwang, "A robust-fragile watermarking scheme for image authentication," in *3rd International Conference on Innovative Computing Information and Control*, pp. 176, 2008.

[25] N. I. Wu, C. M. Wang, C. S. Tsai, M. S. Hwang, "A certificate-based wtermarking scheme for coloured images", The Image Science Journal, vol. 56, no. 6, pp. 326-332, 2008.

[26] M. Yamni, H. Karmouni, M. Sayyouri, H. Qjidaa, "Efficient watermarking algorithm for digital audio/speech signal," *Digital Signal Processing*, vol. 120, pp. 1-13, 2022.

[27] X. Yang, S. Wang, W. Xu, J. Qiao, C. Yu, C. Fernandez, "Fuzzy adaptive singular value decomposition cubature Kalman filtering algorithm for lithium-ion battery state-of-charge estimation," *International Journal of Circuit Theory and Applications*, vol. 50, no. 2, pp. 614-632, 2022.

[28] P. Zheng, Y. Zhang, "A robust image watermarking scheme in hybrid transform domains resisting to rotation attacks," *Multimedia Tools and Applications*, vol. 79, no. 7, pp. 18343-18365, 2020.

[29] N. Zhou, W. M. Xia Hou, R. H. Wen, W. P. Zou, "Imperceptible digital watermarking scheme in multiple transform domains," *Multimedia Tools and Applications*, vol. 77, no. 23, pp. 30251-30267, 2018.

Biography

Sun Ying, born in Changyuan County, Henan province, graduated from the University of Electronic Science and Technology in software engineering with a master's degree. She is an excellent lecturer and is now working at Zhengzhou Railway Vocational and Technical College, China. She is interested in digital media technology and virtual reality technology applications.

Guofeng Ma, born in Kaifeng, Henan Province, graduated from Huazhong University of Science and Technology with a master's degree in engineering. He is an associate professor at Zhengzhou Railway Vocational and Technical College, China. He is interested in computer applications.

Huacheng Xie is currently working in the School of Computer and Information Technology, Xinyang Normal University. His current research interests include network and system security.

Research on Resource Scheduling Strategy Based on Energy Consumption in Datacenter

Jianhua He and Dongli Wu
(Corresponding author: Jianhua He)

School of Computer Science and Engineering, Sichuan University of Science & Engineering
Cuiping District, Yibin City, Sichuan 644000, China
Email: hejianhua@suse.edu.cn

(Received Mar. 27, 2023; Revised and Accepted Nov. 22, 2023; First Online Dec. 24, 2023)

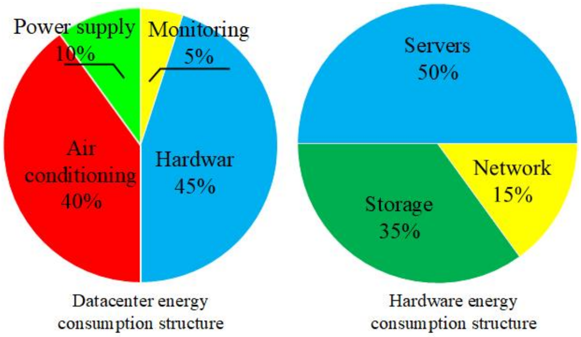
Abstract

Current energy optimization algorithms are processed based on the data center load but do not fully account for the additional overhead in migration. This paper proposes a scheduling method, LFMOS (Load Forecasting and Multi-Object Searching), that combines load identification and migration implementation. First, the load forecasting model can classify cloud computing tasks in advance according to the load status of the data center. Then, according to the prediction results, the migration program can select VMs (virtual machines) and routes that have less impact on tasks to migrate according to the prediction results to ensure QoS and reduce energy consumption. The multi-target search algorithm for the network can also effectively ensure the network state of the data center, avoid network overload, and affect the stable operation of the data center. The proposed method is verified by real workload through the CloudSim platform. Finally, the experimental results show that compared with other methods, this method significantly reduces energy consumption, improves service quality, and reduces the frequency of host shutdowns.

Keywords: Energy Consumption Optimization; Load Forecasting; Multi-objective Optimization

1 Introduction

As one of the most effective resource service models in recent years, cloud computing centrally manages and uniformly schedules all software and hardware resources serving users, provides a mature business model for all kinds of enterprises, and greatly reduces the difficulty of enterprises to access the Internet. According to a research report on the size of the cloud computing market, the global cloud computing industry will grow at a compound annual growth rate of 15.7% and reach USD 15,549.4 billion in 2030 [9]. This shows that the development of the cloud computing industry is unprecedentedly strong [4, 11, 18, 19].



Datacenter energy consumption structure

Hardware energy consumption structure

Figure 1: Energy consumption structure in datacenter

As the cloud computing market grows, the size of datacenter continues to expand, and so does consumption. Studies show that management, equipment operation, and air conditioning account for 30% of the total energy consumption of datacenters worldwide [30]. Most hardware resource workloads are only 10%-50%, but low-load resources still consume 50%-60% of the energy of the full-load hardware, with an overall energy consumption of 70% [26]. According to a report from IEA (International Energy Agency), global datacenter electricity consumption in 2021 was 220-320TWh, or accounting for about 0.9-1.3 % of global final electricity demand, and datacenters and data transmission networks are responsible for nearly 1% of energy-related GHG (Greenhouse gases) emissions [12]. In the composition of datacenter energy consumption, its hardware consumption and air conditioning system account for 85% of the total energy consumption, and the physical machine (PM) consumes half of the power on the hardware [29]. The proportion of energy consumption of each part of the datacenter and hardware is shown Figure 1. A reasonable scheduling strategy can improve resource utilization, balance services and energy consumption, and has a significant economic value for increasing the running time of hardware resources and reducing the operating costs of enterprises. Many studies have shown that computing tasks are not fully loaded most of the time. To reduce energy consumption, load balancing, or host shutdown maintenance, many scholars have studied the energy consumption in datacenters. Ding *et al.* [6] proposed a Q-learning algorithm based on the queuing model for centralized schedule tasks. This method assigns tasks according to task priority and dynamic policies. At the same time, it rewards policies that can minimize task allocation to improve policy weight and achieve better scheduling. According to analyzing the edge computing queuing system, Guo *et al.* [10] proposed a delay-based workload allocation, and then used Lyapunov drift-plus-penalty theory to optimize the energy consumption. Devaraj *et al.* [5] proposed to use firefly algorithm to minimize the search space and use multi-objective particle swarm optimization to identify the enhanced response. Considering input errors, Stavrinides Et al. [27] proposed an energy-efficient, QoS-aware and cost-effective scheduling strategy by using per-core Dynamic Voltage and Frequency Scaling (DVFS) on the underlying heterogeneous multi-core processors. Haghighi *et al.* [2] proposed a resource management hybrid technique that uses an improved k-means clustering by genetic algorithm to map tasks and dynamically integrate, minimizing the number of virtual machine migrations. Jena *et al.* [13] proposed a dynamic load balancing algorithm that uses hybrid process to adjust the speed of particle swarm optimization and Q-learning to produce optimal action. The above methods show that dynamic task scheduling from the various hardware indicators of DC can effectively reduce energy consumption and ensure QoS, but the specific implementation depends on the operating status of the DC, and the migration effect varies greatly under different working conditions of the datacenter. To address the above programs, based on the hardware load and network structure of the datacenter, a load forecast and multi-object search algorithm (LFMOS) is proposed to reduce the energy consumption of the datacenter. This method uses the predictive analysis load and biological population algorithm to explore the migration algorithm of the path.

The remaining content is organized as follows: Section 2 gives a brief review of forecasting methods and multi-objective optimization algorithms, and then introduces the improvements in this paper. Section 3 introduces the filtering algorithm and then describes how to implement the prediction algorithm and the path search algorithm. Comprehensive experiments will be performed in Section 4 to illustrate the performance of LFMOS, and section 5 makes a conclusion.

2 Related Work

Migration quality is affected by several impacts, such as current tasks, datacenter structure, and PM status. Traditional migration research aims to solve load-balancing problems, and the impact on datacenter physical equipment and other tasks during migration has not been studied much. The algorithm proposed in this paper mainly relies on load- partitioning and load forecasting algorithms. Based on the rules of task load changes, the load types of PMs and VMs are accurately identified to search paths in complex network environments, and find fast and short-distance migration routes for migration operations. Then a fast and stable migration method is completed to reduce the impact on other tasks and solve the problem that is difficult to fully use multiple computing resources.

To accurately describe the above methods, here are some descriptions of the datacenter parameters.

The CPU, memory, and bandwidth status of PM in the datacenter is described as Equation (1)-Equation (3):

$$P_{CPU} = \sum_{i=1}^{n} V_{iCPU}/Total_{CPU}. \tag{1}$$

$$P_{memy} = \sum_{i=1}^{n} V_{imemy}/Total_{memy}. \tag{2}$$

$$P_{band} = \sum_{i=1}^{n} V_{iband}/Total_{band}. \tag{3}$$

where n is the total number of VM; V_i is the i_{th} VM, and the subscript indicates the resource type, i.e., the i-th of all running VMs in this PM. $Total$ is all resources that PM can provide. P is the utilization rate of current resources, whose range is from 0 to 1. When performing a migration, a higher workload would be more likely to affect other tasks. As such, it is necessary to select low-impact tasks for migration.

2.1 Load Forecasting Algorithm

Basic Forecasting Method. In-depth researches on the forecasting method of time series workload have been carried out. Yadav *et al.* [28] used a time-series long short-term memory (LSTM) model to forecast cloud workload, but only studied the problem in the overall load dimension. Ruan *et al.* [25] used a time-series LSTM model to forecast the storage load of datacenters. Kumar *et al.* [16] forecasted the overall load of the datacenter. Kumar *et al.* [15] used the error prevention score o improve the forecasting accuracy. The above results shows that the time series load of the datacenter can be effectively predicted, but these methods all use relevant neural networks to handle the problem, which needs a large amount of data, and will cause additional overhead to the datacenter. Li *et al.* [17] used the prediction results of GM (1, 1) to train the LSTM neural network model. Madhi *et al.* [20] used the integral reconstructed background values to optimize the GM (1, 1) prediction ability. Compared with the original method, this method can effectively improve the prediction ability. The above methods show that the effect of using GM (1, 1) to handle time series problems is optimal.

The resource requirements of tasks are scattered and disordered. A gray forecasting model is used here to predict the workload of time series and improve the forecast accuracy by reducing the prediction size of the original model and increasing the variables. Combined with task demand and hardware workload prediction, taking the memory modification frequency, network throughput and CPU utilization as the main parameters to classify VMs and PM after data noise filtering, the variation of hardware workload and computing tasks are both modeled in this method. The principle is as follows:

The original load sequence $x^{(0)}(k) \geq 0, (k = 1, 2, \cdots, n)$ has k observations, and $X^{(1)} = (x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(n))$ is the first accumulation sequence of $x^{(0)}$. The form of differential equation for the GM(1,1) predictive model is as Equation (4):

$$ {}^{dx^{(1)}}/_{dt} + ax^{(1)} = b \tag{4} $$

where a and b are the parameters to be solved.

$\hat{a} = [a, b]^T$ is a vector of parameters to be solved, and it can be calculated by the Equation (5):

$$ \hat{a} = (B^T B)^{-1} B^T Y \tag{5} $$

where $Y = [x^{(0)}(2), x^{(0)}(3), \cdots, x^{(0)}(n)]^T$, $B = [-Z^{(1)}(2), -Z^{(1)}(3), \cdots, -Z^{(1)}(n)]^T$, $z^{(1)}$(k)= $\frac{1}{2}$($x^{(1)}$(k)+$x^{(1)}$(k-1)), $k = 1, 2, 3, \cdots, n$.

The gray forecasting model can be obtained by solving Equation (6):

$$ \hat{x}^{(1)}(k+1) = (x^{(1)}(1) - \frac{b}{a})e^{-ak} + \frac{b}{a}, k = 0, 1, \cdots, n \tag{6} $$

Error Correction Method. Suppose the load original sequence is $X^{(1)} = (x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(n))$, and the predicted sequence is $\hat{X}^{(1)} = (\hat{x}^{(1)}(1), \hat{x}^{(1)}(2), \cdots, \hat{x}^{(1)}(n))$, the error between the predicted result and the real result is $q^{(0)}(k) = x^{(1)}(k) - \hat{x}^{(1)}(k)$. When $k = 1, 2, \cdots, n$, the residual sequence is $Q^{(1)} = (q^{(0)}(1), q^{(0)}(2), \cdots, q^{(0)}(n))$. Substituting the residual sequence into the above gray prediction model, the time response sequence of the residual is Equation (7):

$$ \hat{q}^{(1)}(t) = (q^{(0)}(1) - b'/a')e^{-a'k} + b'/a' \tag{7} $$

Taking the differential of Equation (7), the modified model of the residual can be obtained as Equation (8):

$$ \hat{q}^{(0)}(k'+1) = -a'(q^{(0)}(1) - b'/a')e^{-a'k} \tag{8} $$

Substituting Equation (8) into Eq.6 to obtain the load forecasting model with error correction, the formula is as Equation (9):

$$ \begin{aligned} \hat{x}^{(1)}(k+1) = (x^{(1)}(1) - b/a)e^{-ak} + b/a + \\ \lambda(-a')(q^{(0)}(1) - b'/a')e^{-a'k} \end{aligned} \tag{9} $$

where $\lambda = \begin{cases} 1 & k \geq i \\ 0 & k < i \end{cases}$.

Target Host Search Method. The datacenter uses a mesh structure to connect PMs and uses virtualization technology to allocate m VMs in n PMs, and the relationship between VMs and PMs is shown in Figure 2. The VM placement is an N-P problem that needs to meet the computing resource requirements of the task, such as smaller energy consumption, robustness, and load balancing requirements. Different computing tasks have different requirements for computing resources. This part proposes a host and route search algorithm to improve resource utilization, reduce migration time, reduce datacenter energy consumption, and improve datacenter performance.

Based on the above descriptions, the collection of datacenter VMs can be expressed as $VM = \{VM_1, VM_2, \cdots, VM_n\}$, and each VM has three attributes: CPU, memory, and bandwidth. The relationship between VMs and Hosts is described by a matrix R, shown as Equation (10):

$$ R = PM^T VM = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix} \tag{10} $$

where a_{ij} is the relationship between VM_i and $host_j$, with two values 0 and 1. $a_{ij} = 1$ and $\sum_{j=1}^{m} a_{ij} = 1$ i.e., means at any time, only one PM serves VM_i.

Improved Aco Search Algorithm. The biological population algorithm is used to solve complex solution space problems that are difficult to solve with traditional algorithms. Gamal *et al.* [8] proposed a hybrid metaheuristics technique method to demonstrate the effect of biological population algorithm in dynamic cloud computing environments. Neelima *et al.* [22] proposed an algorithm that combines the dragonfly algorithm and the firefly algorithm, considering the efficiency, cost and effectiveness to achieve the advantages of scheduling. Princess *et al.* [1] analyzed the factors affecting the load and proposed a method that combines the advantages of Harries Hawks Optimization and Pigeon-inspired Optimization Algorithm. This method can effectively improve the efficiency of the search algorithm and quickly balance load. Negi *et al.* [23] proposed to classify the load and use the optimization method of multi-objective ranking preference to implement dynamic load balancing based on artificial neural network, which can achieve dynamic balance between load and energy consumption. Mapetu *et al.* [21] proposed an efficient particle swarm optimization algorithm that uses constraint functions to optimize the time complexity and reduce the calculation. Ebadifard *et al.* [7] studied the influence of DC network on scheduling, and proposed to reduce migration overhead by predicting CPU and I/O, and the risk of overload of limited additional computing requests is reduced by virtual machine autonomy. The problem to be solved is not only searching for possible routes, but also exploring the carrying capacity of the path for migration. In this way, an improved ACO is used to solve the pathfinding problem in datacenters. In the basic ACO, ants can only obtain simple information, but not perceive the amount of specific resources. Here the ant search strategy is optimized to increase task rewards and drive ants to search target

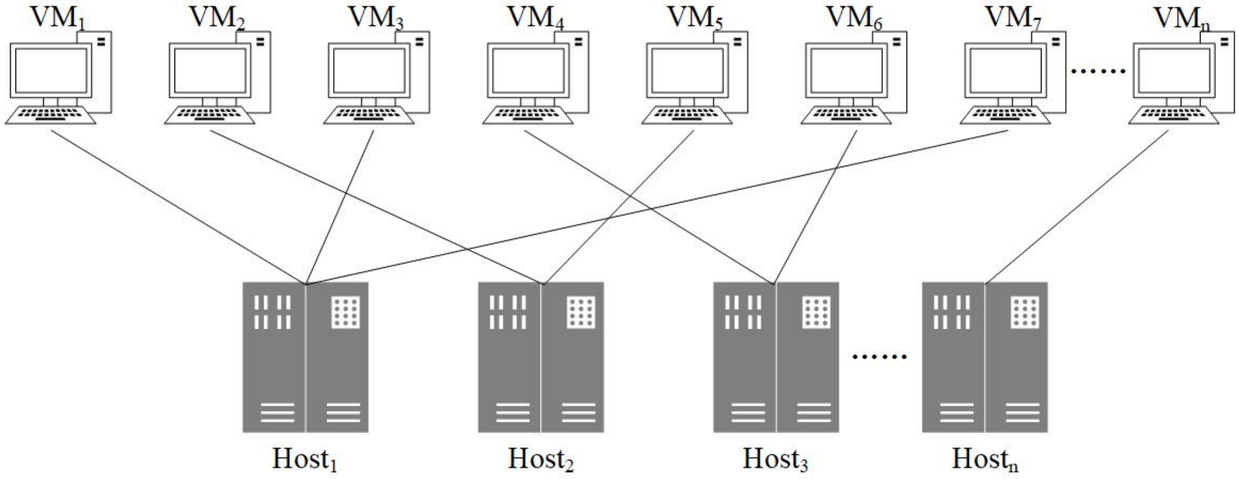


Figure 2: Deployment relationship between VMs and Hosts

host with the resources needed for the migration VM. After the ant searches, it will leave pheromones on the path, and the target PM will deduct the corresponding computing resources to achieve resource coupling. After multiple rounds of searching, the route with the highest pheromone in the network is the effective migration route of the VM to migrate from current PM to target PM. To avoid the influence of high local pheromone concentration on problem-solving, this algorithm adopts the double random method of ant placement and line search to expand the search range.

Target Host Selection. The resources taken away by ants are reprehensive. Taking CPU resources as an example, the resource consumed by VM relocation is as Equation (11):

$$D_{ij} = (CPU_{host_j} + CPU_{VM_i})/Tcpu_{host} \tag{11}$$

where $Tcpu_{host}$ is the total resource provided by PM that cannot be changed. CPU_{host_j} is the resource that has been provided when the system preparing for migration. CPU_{VM_i} is the resource that the PM needs to provide to the migrated VM. D_{ij} is the ratio of the resources required by VM to the total resources that the host can provide, which means the cost of the resources provided by the PM, and it is also called the heuristic factor for target selection.

Pheromone Volatilization and Routes Select. During the operation of the datacenter, network load is important to the quality of service (QoS) and VM migration. In a datacenter network, routing table information cannot avoid the impact of migration on cloud computing tasks. In this part, the author models the network load as a temperature that affects pheromone volatilization. Pheromones will volatilize fast at high temperatures, and the possibility of high-load lines being selected by ants will be low. The pheromone volatility model is as Equation (12) and Equation (13):

$$\lambda_{ij} = 1 - Band_{pre}/Band_{total} \tag{12}$$

$$\tau_{ij} = \lambda_{ij}(1-\rho)\tau_{ij} \tag{13}$$

where λ_{ij} is the temperature coefficient of the route from routei to $route_j$. τ_{ij} is basic pheromone volatility coefficientthat is affected by the basic volatility coefficient and λ_{ij}.

Equation (12) and Equation (13) show that if the path load is too heavy, the resources required for migration are also large, and the possibility of ants choosing this path is low. That is, through setting the temperature coefficient, the probability of line selection can be dynamically adjusted to avoid the migration of network lines with weak service capabilities. There are multiple lines connecting any two routes in the LAN and each route is connected to multiple routes, and its structure is shown in Figure 3. The available bandwidth of each link and the load to be migrated has different impacts on the original task, and it is necessary to search for a reasonable migration line to reduce the impact and improve efficiency. The probability of an ant choosing routej at routei is defined asP_{ij} , and the probability of route selection is determined by the pheromone concentration and the node heuristic factor, described as Equation (14):

$$P_{ij} = {\tau_{ij}}^{\alpha} \times {D_{ij}}^{-\beta} \Big/ \sum_{i=0}^{n}\sum_{j=0}^{m} {\tau_{ij}}^{\alpha} \times {D_{ij}}^{-\beta} \tag{14}$$

where α and β are the weights of the pheromone and heuristic factor respectively relative to the search. Ants are more likely to choose the routes with large P_{ij} for exploration.

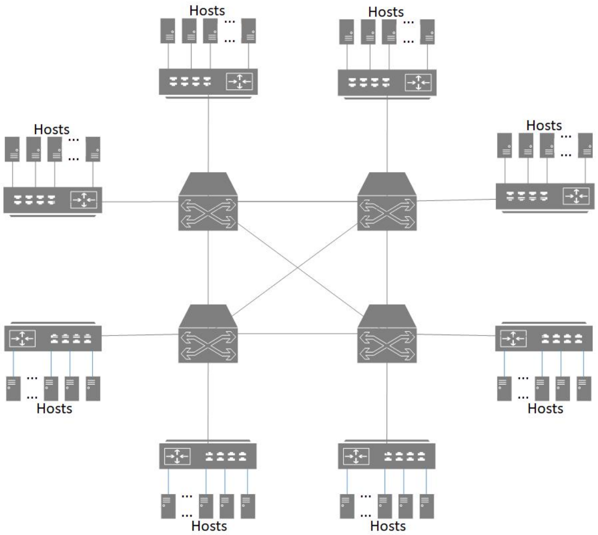


Figure 3: Datacenter LAN diagram

3 Implementation of Migration Algorithm

3.1 Load Noise Filtering

Although the original data can fully reflect the workload change of the datacenter, small and frequent workload changes will increase the calculation and reduce the effect. Percival *et al.* [24] proposed a filtering method to reduce noise and enhance signal in the time-series model. Jiang *et al.* [14] studied the performance of using SG filter in time-series. The SG filter algorithm is a weighted simulation of smoothing time series datathat widely used in data smoothing and noise reduction. When processing noise, SG filter algorithm can also retain the curve state of the original sequence and is not limited by the size of the dataset. Chen *et al.* [3] used SG filtering to reduce the load volume of data in datacenter using SG filtering, which can achieve efficient and complex predictions. The SG filtering algorithm is used to filter workload noise and reduce the amount of data to be processed. The SG filter model is as Equation (15):

$$Y_j^* = \sum_{i=-m}^{i=m} C_i Y_{j+1}/N \tag{15}$$

where Y_j^* is the fitting result of time series;C_i is the filter coefficient of the i_{th} data value; Y is the original sequence; N is the number of convolutions; j is the coefficient of the original sequence data set, and m is the size of filter window. The comparison between the original workload data and the filtered data is shown in Figure 4. The above experimental results show that compared with the original load, SG filtering algorithm can effectively reduce the range of data fluctuation. Compared with the Kalman filtering, SG filtering can preserve the trend of data change.

3.2 Load Forecasting Method Implementation

Method.

Input: Real load sequence after denoising.

Output: Load the prediction result LP_{result}.

Step 1: Obtain the current real load sequenceLS, and define $n-2$ subsequencessLS_i.

Step 2: Divide the sequence LS into $n-2$ subsequences $sLS_i = (x_{n-i+1}, x_{n-i+2}, \cdots, x_n)$ according to the single-step size, and $sLS_3 = (x_{n-2},\ x_{n-1},\ x_n)$; $\cdots$; $sLS_i = (x_{n-i+1},\ x_{n-\mathrm{i}+2}, \cdots, x_n); \cdots; sLS_n = (x_1, x_2, \cdots, x_n)$.

Step 3: The model is used to predict the sub-workload sequence; the result is $LPR_i$and the final result is$LPR = (LPR_1, LPR_2, \cdots, LPR_n)$.

Step 4: Set weight (W) for all predicted values. W is the proportion of the LPR in the whole, where$w_i = LPR_i/\sum_{j=3}^{n} LPR_j$. w_i represents the weight of the predicted values of the subsequence with length i.

Step 5: Fit the predicted value of the subsequence with the weight to obtain the final predicted valueLP_{result}, $LP_{result} = W \cdot LPR$.

The load forecasting process is shown in Figure 5.

3.3 Prediction-Based Migrated Host Selection Method

The algorithm first divides several load queues based on the load characteristics, to obtain 10 cycles of VM load data through the VM monitor for data screening. Then, the load of the next cycle is predicted by the improved gray GM (1, 1) prediction algorithm with step size and weight respectively based on the CPU, Memory, and Bandwidth. According to the predicted load, the load type is judged according to the corresponding thresholds and weights. If the prediction result shows that VMs in PM will be overloaded, the VM with the smallest load in the PM will be repeatedly marked to the migration queue for migration, until the computing resources provided by the PM can satisfy the stable operation of the overloaded VM. After the migration is completed, the VM monitor continues to obtain new load data to calculate the residuals of the actual load and the predicted load, and correct the prediction of the new round of load. The algorithm uses continuously generated real data and residual correction to ensure the accuracy and stability of the prediction algorithm.

3.4 Path Search Algorithm

The improved search algorithm has parameters, such as ants, pheromone concentration, node status, and line temperature. The path node represents the attempt at the

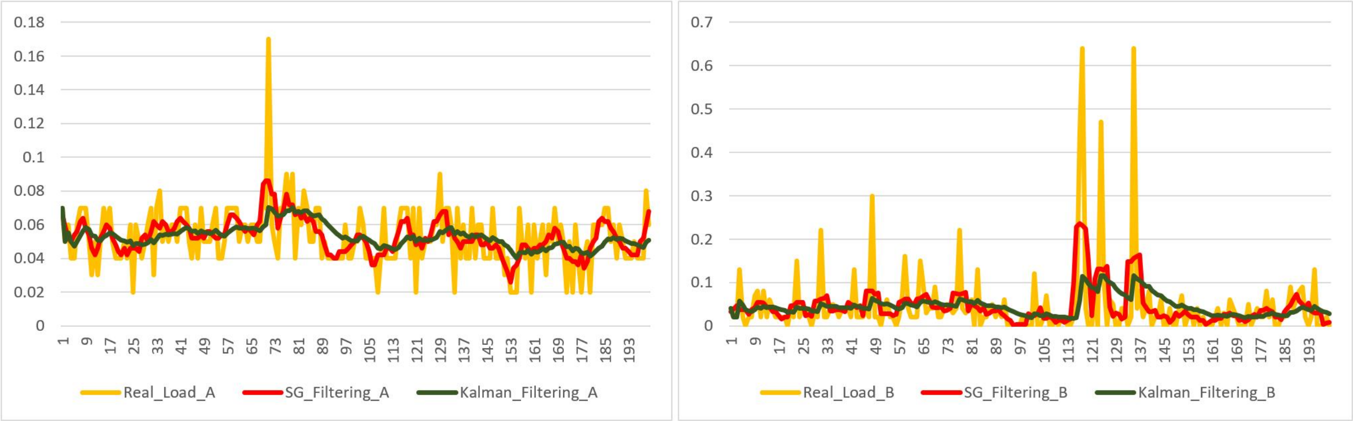


Figure 4: Filtering result of workload

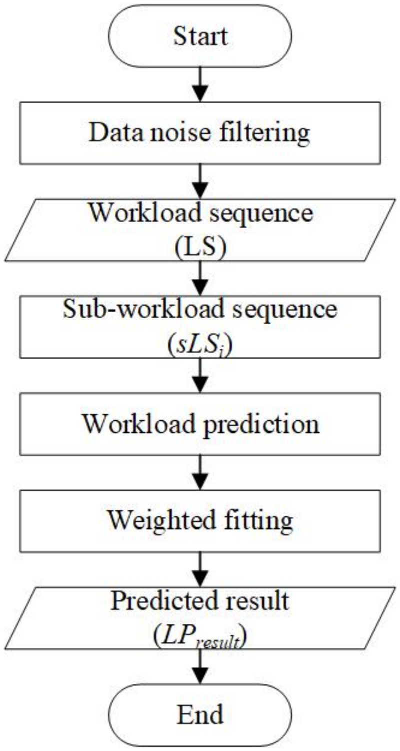


Figure 5: Energy consumption structure in datacenter

optimal solution; the pheromone concentration represents the optimal degree of the current solution, and the line temperature represents the load degree of the current network link. The search algorithm steps are as follows:

Step 1: Population initialization: set the size of ant colony, search range of problem-solving space, number of searches, and pheromone concentration. Ants are used to move randomly in the path of the problem space, searching for the solution to the problem. When the ants walk in the path, they will leave pheromones for other ants for referring, so that the multiple ants can search together.

Step 2: Acquisition of line ambient temperature: Obtain all network load conditions between two nodes from the VMM, and calculate the line temperature based on the load. The pheromone volatilization speed is dynamically set according to the temperature parameter.

Step 3: Random route selection: put individual ants in the route and drive the ants to find food. When selecting node lines, random perturbation is used to drive ants for selecting new lines randomly, and departmental ants are randomly placed in the lines to drive ants for selecting new nodes and expanding the search range. The local optimal problem caused only by ant pheromones is avoided and a better solution is obtained.

Step 4: Find food: Ants find resources and take them away, leaving pheromones in the route for other ants for route reference. The concentration of pheromones volatilizes over time, preventing outdated information from affecting the decision-making of ants.

Step 5: Drive the ants to complete all search tasks. After this iteration, the line with the most pheromone is the optimal path.

Step 6: Repeat Steps 2 to 5 until the set termination condition is reached, and output the migration strategy.

Step 7: Put the free-load PM to sleep after executing the migration task.

The forecasting process of the load forecasting model is shown in Figure 6.

4 Experiment

4.1 Datasets

When cloud computing services are used, the provider will configure a VM on the PM for normal use, and the datacenter will provide relevant configuration information after obtaining the configuration of the VM. To test the effectiveness of the algorithm, this paper uses the CloudSim platform for virtual simulation of the algorithm. To accurately simulate the task load in a real environment, the experiment uses the real data provided by CloudSim as the load input, which is collected by PlanetLab from virtual machines around the world at five-minute intervals.

4.2 Evaluation Indicators

To verify the effectiveness of the LFMOS model, the energy consumption, number of PM shutdowns, and number of VM migrations, migration execution time, and services level agreement are used for the scheduling performance of the model. Energy consumption indicates the energy consumed by the PM to operate all tasks. The number of host shutdowns indicates the number of dormant PMs due to energy consumption during the scheduling process. The number of VM migrations indicates the total number of VMs migrated in the overall scheduling. The migration time indicates the time elapsed from triggering VM migration to resuming operation. The SLA indicates the length of time during which the host has insufficient performance in tasks. The smaller the value of the five metrics, the more stable in datacenter, the higher the QoS, and the smaller the impact on cloud tasks with the proposed strategy. The smaller the value, the more stable the data center scheduling strategy.

4.3 Comparison Methods

To evaluate the validity of the proposed model, it is compared with Median Absolute Deviation (MAD), Inter Quartile Range (IQR), and Local Regression Robust (LRR), that proposed by CloudSim Lab.

4.4 Experimental Settings

Taking the real workload data provided by the platform as an example, the experimental simulation parameters are shown in Table 1 - Table 3

Table 1: PM performance settings

Parameters	Values
Number of PM	800
Number of CPU per PM	2
PM CPU performance	{1860, 2860}MIPs
PM memory	{4096, 4096}MIPs
PM bandwidth	1000Mbit/sMIPs

Table 2: VM performance settings

Parameters	Values
Number of VM	1033
Number of CPU per VM	1
VM CPU performance	{2500, 2000, 1000, 500}MIPs
VM memory	{870, 1740, 1740, 613}MIPs
VM bandwidth	100Mbit/sMIPs

Table 3: Task settings

Parameters	Values
Number of cloud task	1033
Total executing time	216000000s
PScheduling interval time	300s

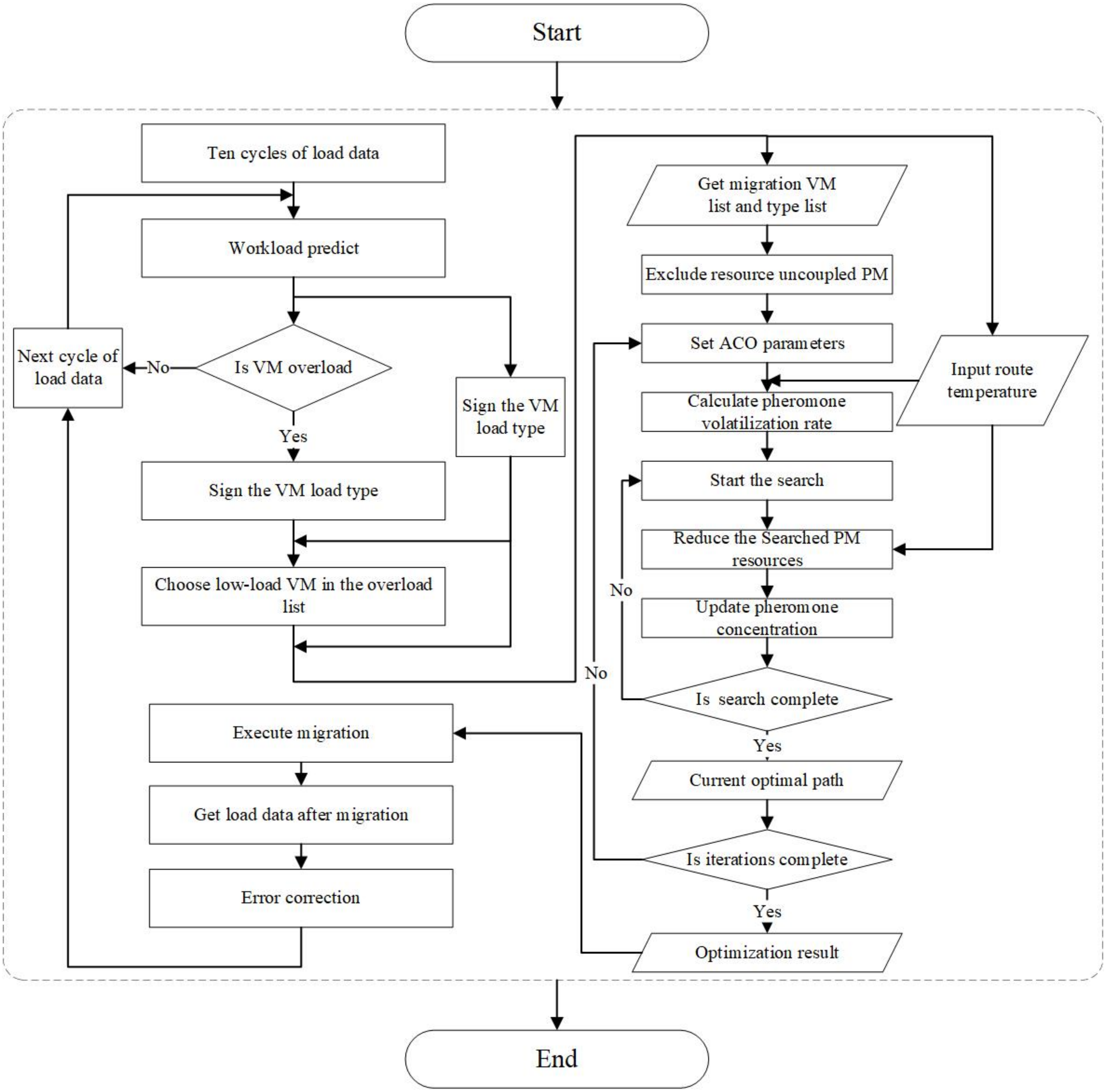


Figure 6: Migration flow chart

4.5 Experimental Results and Analysis

To demonstrate the effectiveness of the LFMOS model, experiments are conducted from five aspects to compare with the three methods proposed by CloudSim.

Energy Consumption. In terms of energy performance of four VM scheduling strategies at different thresholds, the LFMOS strategy identifies as much low-load PMs as possible, effectively integrating the fragmented loads and reducing the energy consumption of the datacenter. Compared with the MAD, IQR, and LRR strategies, the LFMOS can reduce energy consumption by about 55.6%, 55%, and 29.4%, respectively. The experimental results are shown in Figure 7.

Number of VM migrations. In terms of the number of VM migrations of four VM scheduling strategies at different thresholds, due to the stable load changes

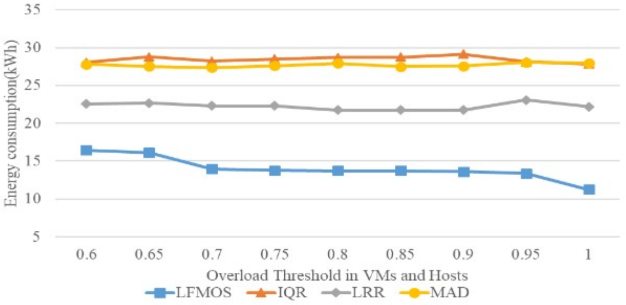


Figure 7: Simulation results of energy consumption

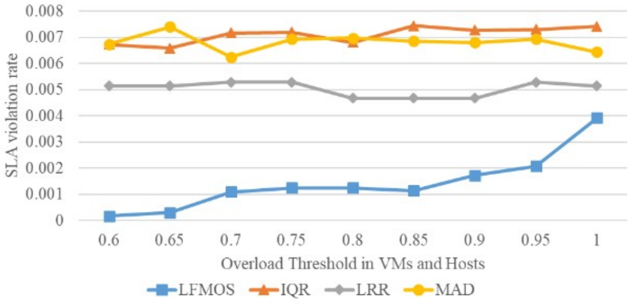


Figure 9: Simulation results of SLA violation rate

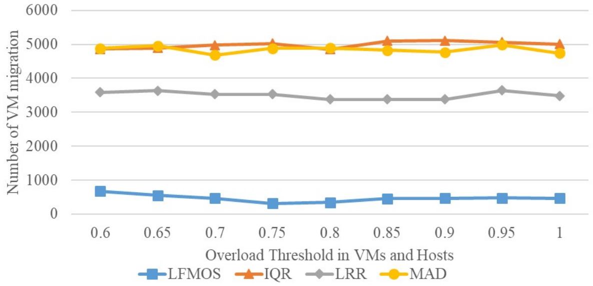


Figure 8: Simulation results of number of VM migrations

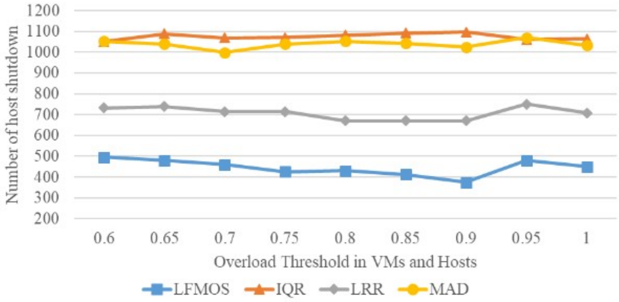


Figure 10: Number of host shutdowns

of the datacenter, the LFMOS strategy has a significant effect in reducing the number of VM migrations. Compared with the IQR, LRR, and MAD strategies, the LFMOS strategy can reduce the number of VM migrations by about 83.7% and 68.4%, respectively. When the load threshold is greater than 0.8, the migration times will increase due to the resource competition. The experimental results are shown in Figure 8.

SLA Violation Rate. In terms of the SLA violation rate of four VM scheduling strategies at different thresholds, the LFMOS method can effectively maintain a low default rate under low load conditions in the datacenter. Compared with IQR, LRR and MAD strategies, the LFMOS can reduce the datacenter SLA default rate by about 80%. However, as the host load threshold increases, the occupancy of different resources in the datacenter tends to be saturated, and the SLA default rate under the LFMOS policy continues to rise. The experimental results are shown in Figure 9.

Number of PM Shutdowns. In terms of the number of PM shutdowns aspects of four VM scheduling strategies ay different thresholds, the LFMOS achieves stable operation of data centers through load classification and resource coupling matching strategies. Compared with the IQR, LRR, and MAD strategies, the LFMOS strategy can reduce the number of host shutdowns by about 50% and 35%. The experimental conclusion is shown in Figure 10.

Migration Execution Time. In terms of the migration execution time of four VM scheduling strategies at different thresholds, the LFMOS can improve the efficiency of migration execution by utilizing a low-load network path for migration through an improved ACO algorithm. Compared with MAD, IQR, and LRR strategies, the LFMOS strategy can reduce the migration time, which is long under a high overload threshold. This is due to the high occupancy rate of each computing resource of the PM in a high overload state, and the relatively small number of resources used for migration,, resulting in a long migration time. The experimental results are shown in Figure 11.

5 Conclusion

According to the experiment results, when the overload threshold is 0.85, the performance and energy consumption of the datacenter are optimal. Compared with other selection strategies, the LFMOS strategy has better performance in energy consumption, number of migrations, SLA breach rate, and migration execution time. Experiments show that the LFMOS strategy can improve the resource efficiency and reduce energy consumption. At the same time, the path search strategy can achieve faster migration efficiency and ensure a low SLA breach rate. The

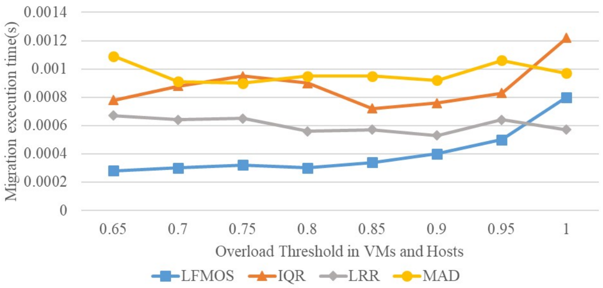


Figure 11: Simulation results of migration execution time

multi-target search algorithm of the network can also effectively ensure the network state of the data center, avoid network overload, and affect the stable operation of the data center. However, compared with other algorithms, the scheduling performance of the strategy proposed in this paper is mediocre when the datacenter load is heavy. In the future, it will focus on solving the resource scheduling problems in heavy load environments.

6 Acknowledgement

This work was supported in part by the Key Laboratory of Higher Education of Sichuan Province for Enterprise Informatization and Internet of Things under Grant 2021WYJ04, in part by Key Laboratory of Higher Education of Sichuan Province for Enterprise Informatization and Internet of Things under Grant 2021WYY03. And the authors would like to thank the editors and anonymous reviewers for their constructive comments and suggestions, which greatly helped improve the overall quality of the manuscript.

References

[1] P. P. G. Annie and A. S. Radhamani, "A hybrid meta-heuristic for optimal load balancing in cloud computing," *Journal of grid computing*, vol. 19, no. 2, p. 21, 2021.

[2] H. M. Askarizade, M. H. Maeen, and M. Haghparast, "An energy-efficient dynamic resource management approach based on clustering and metaheuristic algorithms in cloud computing iaas platforms: Energy efficient dynamic cloud resource management," *Wireless Personal Communications*, vol. 104, pp. 1367–1391, 2019.

[3] L. Chen, W. W. Zhang, and H. M. Ye, "Accurate workload prediction for edge data centers: Savitzky-golay filter, cnn and bilstm with attention mechanism," *Applied Intelligence*, vol. 52, no. 11, pp. 13027–13042, 2022.

[4] P. S. Chung, C. W. Liu, M. S. Hwang, "A study of attribute-based proxy re-encryption scheme in cloud environments", *International Journal of Network Security*, vol. 16, no. 1, pp. 1-13, 2014.

[5] A. F. S. Devaraj, M. Elhoseny, S. Dhanasekaran, E. L. Lydia, and K. Shankar, "Hybridization of firefly and improved multi-objective particle swarm optimization algorithm for energy efficient load balancing in cloud computing environments," *Journal of Parallel and Distributed Computing*, vol. 142, pp. 36–45, 2020.

[6] D. Ding, X. C. Fan, Y. H. Zhao andK. X. Kang, Q. Yin, and J. Zeng, "Q-learning based dynamic task scheduling for energy-efficient cloud computing," *Future Generation Computer Systems*, vol. 108, pp. 361–371, 2020.

[7] F. Ebadifard and S. M. Babamir, "Autonomic task scheduling algorithm for dynamic workloads through a load balancing technique for the cloud-computing environment," *Cluster Computing*, vol. 24, pp. 1075–1101, 2021.

[8] M. Gamal, R. Rizk, H. Mahdi, and B. E. Elnaghi, "Osmotic bio-inspired load balancing algorithm in cloud computing," *IEEE Access*, vol. 7, pp. 42735–42744, 2019.

[9] Grand View Research, "Cloud computing market size, share & trends analysis report by service (iaas, paas, saas), by deployment (public, private, hybrid), by enterprise size, by end use (bfsi, it & telecom, retail & consumer goods), by region, and segment forecasts, 2022-2030,". Tech. Rep. GVR-4-68038-210-5, 2022.

[10] M. Guo, L. Li, and Q. S. Guan, "Energy-efficient and delay-guaranteed workload allocation in iot-edge-cloud computing systems," *IEEE Access*, vol. 7, pp. 78685–78697, 2019.

[11] M. S. Hwang, T. H. Sun, C. C. Lee, "Achieving dynamic data guarantee and data confidentiality of public auditing in cloud storage service," *Journal of Circuits, Systems, and Computers*, vol. 26, no. 5, 2017.

[12] International Energy Agency, "Data centres and data transmission networks,". Tech. Rep. GVR-4-68038-210-5, 2022.

[13] U. K. Jena, P. K. Das, and M. R. Kabat, "Hybridization of meta-heuristic algorithm for load balancing in cloud computing environment," *Journal of King Saud University-Computer and Information Sciences*, vol. 34, no. 6, pp. 2332–2342, 2022.

[14] H. Jiang, Y. Yao, X. Hong, Y. Zhao, and Y. Li, "Research on reconstruction algorithm of time series harmonic analysis based on sg filtering and denoising," *Journal of Jilin Normal University(Natural Science Edition)*, vol. 03, pp. 133–140, 2021.

[15] J. Kumar and A. K. Singh, "Cloud datacenter workload estimation using error preventive time series forecasting models," *Cluster Computing*, vol. 23, no. 2, pp. 1363–1379, 2020.

[16] J. Kumar and A. K. Singh, "Performance assessment of time series forecasting models for cloud datacenter

networks' workload prediction," *Wireless Personal Communications*, vol. 116, pp. 1949–1969, 2021.

[17] Z. Li, "Neural network economic forecast method based on genetic algorithm," *IET Software*, 2023.

[18] C. W. Liu, W. F. Hsien, C. C. Yang, M. S. Hwang, "A survey of attribute-based access control with user revocation in cloud data storage", *International Journal of Network Security*, vol. 18, no. 5, pp. 900-916, 2016.

[19] C. W. Liu, W. F. Hsien, C. C. Yang, and M. S. Hwang, "A survey of public auditing for shared data storage with user revocation in cloud computing", *International Journal of Network Security*, vol. 18, no. 4, pp. 650-666, 2016.

[20] M. Madhi and N. Mohamed, "Improving gm (1, 1) model performance accuracy based on the combination of optimized initial and background values in time series forecasting," *Open Access Library Journal*, vol. 9, no. 4, pp. 1–17, 2022.

[21] J. P. B. Mapetu, Z. Chen, and L. F. Kong, "Low-time complexity and low-cost binary particle swarm optimization algorithm for task scheduling and load balancing in cloud computing," *Applied Intelligence*, vol. 49, pp. 3308–3330, 2019.

[22] P. Neelima and A. R. M. Reddy, "An efficient load balancing system using adaptive dragonfly algorithm in cloud computing," *Cluster Computing*, vol. 23, pp. 2891–2899, 2020.

[23] S. Negi, M. M. S. Rauthan, K. S. Vaisla, and N. Panwarm, "Cmodlb: an efficient load balancing approach in cloud computing environment," *The Journal of Supercomputing*, vol. 77, pp. 8787–8839, 2021.

[24] D. B. Percival and A. T. Walden, *Wavelet methods for time series analysis*, vol. 4. Cambridge university press, 2000.

[25] L. Ruan, Y. Bai, S. Li, S. He, and L. Xiao, "Workload time series prediction in storage systems: a deep learning based approach," *Cluster Computing*, pp. 1–11, 2021.

[26] W. Si, J. Taheri, and A. Zomaya, "A distributed energy saving approach for ethernet switches in data centers," in *37th Annual IEEE Conference on Local Computer Networks*, pp. 505–512. IEEE, 2012.

[27] G. L. Stavrinides and H. D. Karatza, "An energy-efficient, qos-aware and cost-effective scheduling approach for real-time workflow applications in cloud computing systems utilizing dvfs and approximate computations," *Future Generation Computer Systems*, vol. 96, pp. 216–226, 2019.

[28] M. P. Yadav, N. Pal, and D. K. Yadav, "Workload prediction over cloud server using time series data," in *2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, pp. 267–272. IEEE, 2021.

[29] X. X. Yu, F. L. Song, Y. B. Zhou, and H. F. Liang, "Investigations on the impact of new infrastructure on electricity forecast and power system planning during the 14th five-year plan period," *Electric Power*, vol. 54, no. 7, pp. 11–17, 2021.

[30] Q. Zhang, Z. H. Meng, X. W. Hong, Y. H. Zhan, J. Liu, J. B. Dong, T. Bai, J. Y. Niu, and M. J. Deen, "A survey on data center cooling systems: Technology, power consumption modeling and control strategy optimization," *Journal of Systems Architecture*, vol. 119, p. 102253, 2021.

Biography

Jianhua He is a teacher of Sichuan University of Science & Engineering, and is also a Ph.D. Candidate at Assumption University in Thailand. He Received his MS degree in control science and engineering from Sichuan University of Science & Engineering, 2021. His research interests include multi-objective optimization algorithms, data mining methods, and optimal control methods

Dongli Wu is a teacher of Sichuan University of Science & Engineering. She Received her MS degree from Chongqing Normal University, 2017. Her research interests focus on educational data mining and deep learning.

Research on Blockchain Secret Key Sharing and Its Digital Asset Applications

Chia-Chun Wu[1], Chun-Tse Chang[2], Iuon-Chang Lin[2,4], and Min-Shiang Hwang[3,4]
(Corresponding author: Min-Shiang Hwang)

Department of Industrial Engineering and Management, National Quemoy University, Taiwan (ROC)[1]
Department of Management Information Systems, National Chung Hsing University (ROC)[2]
Department of Computer Science & Information Engineering, Asia University[3]
Fintech and Blockchain Research Center, Asia University, Taiwan[4]
500, Lioufeng Rd., Wufeng, Taichung 41354, Taichung, Taiwan, R.O.C.
Email: mshwang@asia.edu.tw

(Received Oct. 1, 2023; Revised and Accepted Dec. 21, 2023; First Online Dec. 29, 2023)

Abstract

Public Key Infrastructure (PKI) is used in blockchain technology to authenticate entities and ensure the integrity of the blockchain. The overall security of a wallet depends on how anyone secures their private key, and the most critical data stored in a wallet is the user's private key, which is also the only identifier of ownership of encrypted digital assets. We often see cases of lost or stolen wallet private keys or even hacked cases in the domestic and international news, which makes people pay more attention to the issue of wallet private key management. Since there are some problems with traditional key management, such as cold and hot wallets and exchanges, it is no longer possible to fully guarantee the security of key management. All key management protocols have some weaknesses; if these weaknesses are not considered, the designed encryption tools are still insecure. For example, DH cannot protect against MITM (man-in-the-middle) and DoS (denial-of-service) attacks, and cryptosystems cannot be developed without proper protection against these attacks. To solve these challenges, we will propose research topics to improve the mechanism through smart contracts, which can achieve a more efficient approach. These research topics will result in a blockchain key management system with privacy, cross-trainability, and efficiency.

Keywords: Blockchain; Cryptocurrency; Key Management; Secret Sharing

1 Introduction

In recent years, with the rise of digital assets and virtual currencies, wallet security has received more and more attention. The private key is the only and most important line of defense for managing wallets. The private key may face the risk of being lost or even stolen. Once the wallet's private key is stolen or lost, we will likely not be able to get our digital assets back. Therefore, it is necessary to find a better solution to protect the security of private keys. Current tools for protecting private keys include Cold Wallet and Hot Wallet [11]. Although cold wallets can provide good protection, they are more restrictive in use and may also encounter problems with broken hardware, while hot wallets may be hacked—the risk of guest intrusion. In addition, some users are willing to host their assets in centralized institutions, such as cryptocurrency exchanges. However, this approach violates the decentralization and anonymity uniqueness of the blockchain [4, 8–10, 25, 29]. Among them, abnormal events in which assets are locked or lost prove that the security of exchange custody keys is questionable.

This research is based on the "Secret Sharing (SS)" private key partitioning proposed by Adi Shamir and George Blakley [16]. Assume that when users need to use private keys today, they can use the SS scheme to split the private keys. If it is just a general SS scheme, there will be a central person, which may lead to the problem of private key fragments being leaked or allowing the central person to crack it privately. If the SS scheme is improved through the smart contract of the blockchain, when the private key is needed, the requester can call this smart contract to generate the key and then distribute it. Or the user provides the master secret key, and then the system generates the subkeys and divides them. The subkeys can be restored according to the (t, n) setting to improve the above shortcomings. The following are the issues this research will explore and solve: secure storage after splitting private keys, seizure of illegal income, and inheritance and transfer of digital assets.

1) The problem of safe storage after private key division: The wallet's private key is essential. Once stolen or

lost, it will cause irreparable losses. Therefore, the security of the user's wallet private key must be ensured. This research topic must propose a practical solution by dividing private keys through smart contracts on the blockchain.

2) The problem of seizure of illegal proceeds from crypto assets: After supervisors seize the wallet keys of criminals, they may collaborate with other criminals to remove the keys and assets, or the supervisors may steal them themselves. This research will use the private key fragment decentralized encryption and decryption key hosting mechanism developed in the previous topic. Use blockchain smart contracts as a platform for private key custody to extend the scope of related applications. Therefore, this research topic will be applied to legalizing illegal gains from digital assets. After the inspectors seize the criminal's wallet key, it must be kept in custody. A more secure supervision mechanism can be achieved through smart contracts by dispersing keys among multiple people. This prevents criminal associates from removing keys or assets quickly and prevents supervisory personnel from guarding against theft.

3) The problem of inheritance and transfer of digital assets: If important information such as private keys is not provided to relatives when the holder passes away, they cannot transfer their encrypted assets to his family members. Therefore, it is hoped that through the smart contract mechanism, the key distribution can be set first and then triggered through the smart contract to achieve an effective mechanism for the custody of the key inheritance. The topic of this research is the application of digital asset inheritance, solving the problem of digital asset inheritance through smart contracts.

2 Main Contributions

This study has the following main contributions:

1) Solve the wallet private key theft problem that has occurred in recent years.

   In recent years, private key management has received much attention [3, 17, 21, 26]. It is a paramount personal privacy and confidential information. In addition, more and more people have begun to invest in cryptocurrency, NFT (Non-Fungible Token), etc., which has given rise to the problem of information security. For example, the wallet key is lost or stolen, mainly if used in a personal digital asset wallet. Therefore, a safer and more appropriate key management mechanism is needed. When storing management keys, it can provide higher security to ensure the security of private key management.

2) Improve the "Secret Sharing" private key split proposed by Shamir.

   Shamir proposed the "Secret Sharing" key sharing method (t, n). Private key splitting refers to dividing a private key (master key) into several fragments (as shown in Figure 1), and at least one part of them is required to restore the master key. This method can be used to prevent unauthorized persons from obtaining the entire private key to use it to sign or decrypt messages digitally [5–7, 15, 22, 36, 37]. Similar to the concept of jointly holding safe, multiple people need to hold and restore the private key before there is a way to unlock it and solve the problem of private key management. Since the administrator or a third party generates this private key division, this centralized operation will lead to the problem of being defrauded of the key. Therefore, smart contracts will be developed to improve the "Secret Sharing" private key division method proposed by Shamir. This research and development mechanism will be put on the chain to solve this problem.

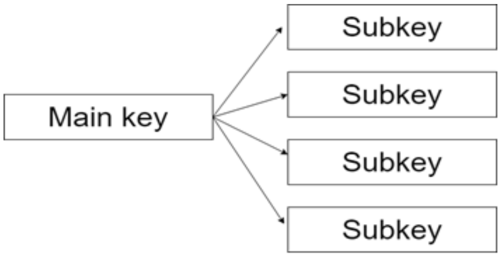


Figure 1: Properties of private key splitting

3) Solve the problem of private key custody through blockchain smart contracts.

   Most traditional key escrow has some flaws. As long as the device that stores the private key or the auxiliary tool for retrieving the private key is lost, the user can no longer retrieve the wallet. In addition, it is vulnerable to intrusion, allowing intruders to discover the key through any mechanism such as brute force cracking and weak encryption [27] or even the vicious collapse of the custodian or exchange [24], resulting in the inability to obtain the private key properly, of safekeeping. Therefore, this research is expected to use blockchain technology to prevent tampering and key leakage, protect the private key of the entire wallet, develop a better private key custody, achieve decentralization, and reduce the risk of being hacked.

4) Solve application scenarios related to the financial aspect of wallet assets.

   More and more criminal cases involving criminal crypto assets occur, which require the seizure of supervisory personnel. Due to the decentralization of cryptocurrencies, they cannot be directly seized like ordinary bank accounts. This type of legal seizure scenario of illegal gains (The criminal's private key has not been leaked out in a short time) uses public power to directly detain the criminal and transfer the assets to a new supervisory wallet. How to avoid supervisors colluding with criminals to transfer money or Transfers secretly are all urgent problems that digital asset finance needs to solve in the future. In addition, most of us in the Internet age have invested in or purchased cryptocurrency. Once the holder of these assets passes away, the assets cannot be transferred without leaving relevant information about the private key to relatives. This is a digital asset. Inheritance issues. This research will use the wallet control master key mechanism generated in the first year to develop automatic triggering smart contract technology to complete the transfer of assets.

3 Research Topics

Public key infrastructure (PKI) is used in blockchain technology to authenticate entities and ensure the integrity of the blockchain [27]. The overall security of a wallet depends on the security mechanisms protecting its private keys. The most important information stored in the wallet is the user's private key, also the unique identifier of the ownership of encrypted digital assets [23]. Cases of wallet private keys being lost, stolen, or even hacked often occur, making people pay more attention to the wallet private key management issue. Due to some problems in traditional key management, such as cold wallets, hot wallets, and exchanges, the security of key management cannot be fully guaranteed. All key management protocols have some weaknesses, and if these weaknesses are not considered, the designed encryption tool will still be insecure. For example, DH cannot protect against MITM (man-in-the-middle) and DoS (denial of service) attacks. Successful development of secure systems will not be possible without cryptographic systems that adequately protect against these attacks [1]. To solve these challenges, this study proposes a holistic solution: improve this mechanism through smart contracts to achieve a more efficient approach.

This study proposes three research themes:

1) First, use a blockchain smart contract to design the division and divergence of a key, and finally, implement the merger to ensure the security of key custody;

2) Under the framework of theme one, Extend related application scenarios in the blockchain part. The legal seizure of wallets illegally obtained from cryptocurrency allows the seizure of criminals' wallets to be appropriately supervised;

3) It is also a related application scenario, continuing the structure of theme one and improving it, triggering operations through smart contracts to solve the legacy of digital assets Referral question.

Integrating this research theme is a private, scalable, efficient blockchain key escrow system.

3.1 Research Topic 1: Research and Development of Blockchain Key Escrow

Most of the currently existing methods and traditional SS solutions require a central person to do the tasks of distributing and recovering keys. The SS scheme has a weakness in that dishonest participants cheat the share, and honest participants are forced to reconstruct fake secrets [33]. Therefore, it is necessary to split the private key fragments, store them with decentralized encryption and decryption, and perform key hosting. Pal *et al.* proposed a secure and efficient blockchain technology Group Key Management (GKM) framework [27]. In this GKM framework, a multi-layer architecture is assumed. The upper-level nodes have more privileges and rights than the lower-level nodes. At each level, there are multiple groups, each containing multiple nodes. Nodes belonging to the same group have the same permissions.

Private key segmentation is also used in many other areas, such as authentication and key management in the Internet of Things [28]. The number of devices connected to the Internet of Things is growing, and the data generated by these devices requires secure and effective access control mechanisms to ensure the privacy of users and data. Most traditional key management mechanisms rely on a trusted third party, such as a registry or key generation center, to generate and manage keys. Trusting a third party leads to a centralized architecture. Panda *et al.* solved these problems by designing a blockchain-based distributed IoT architecture that uses hash chains for secure key management. It is also used in intelligent transportation systems (ITS) and is an essential application of the Internet of Things. Encryption is an effective way to protect the confidentiality of data transmitted in ITS. Zhou *et al.* [38] proposed a method to implement threshold key management in blockchain-based ITS. The proposed threshold key management scheme is efficient and secure for multiple users in blockchain-based ITS, especially for material-sharing scenarios. Zhou *et al.* studied three secret sharing schemes: Shamir, Blakley, and CRT secret sharing schemes. It is concluded that the CRT-based scheme is more effective than the scheme based on Shamir secret sharing. In the system model, shared data is owned by the vehicle. Zhou *et al.* stored the data in the cloud for convenience of use and sharing. However, storing plaintext data may bring security and privacy issues, so these vehicles use private key segmentation to generate keys to encrypt the data, ensuring data storage security. A secret sharing scheme in the system generates keys, which are dispersed into parts and distributed to vehicles in a secure channel. In this way, only some vehicles can recover the key, each vehicle can control the data, and the data can be protected.

The main problem with systems based on blockchain technology is that users will not be able to access their keys due to device damage or loss [20]. Although blockchain technology is decentralized, key management has become the user's sole responsibility. This differs from the centralized approach, where the general account password has some administrator or IT person. The latter allows you to reset your password or create a new password if it is lost or forgotten. The current lack of this functionality in the system will result in the loss of assets. According to Chainanalysis, approximately 20% of Bitcoins (worth approximately $210 billion) have not been moved in the past five years and are therefore considered lost.

In the field of wallet and private key storage research, Boneh and Goldfeder *et al.* [2] proposed a threshold signature wallet scheme: the private key is divided into n fragments and stored by n participants. Completing each transaction requires more than the threshold t All participants jointly sign. Thota [34] proposed a software wallet that resides on the user's mobile device and uses the Hyperledger Fabric blockchain network. Jian [18] proposed a single-point failure protection scheme for blockchain wallets based on trustless central threshold elliptic curve digital signatures. This scheme can complete each transaction through the collaboration of multiple people, increasing the account itself to a certain extent: reliability and security. Dikshit [12] proposed a scheme based on participant identity, which provides different weights for participants.

Rezaeighaleh [32] proposed a new digital scheme for secure backup of hardware wallets that relies on display-enabled channel human visual verification on the hardware wallet. Gutoski [13] proposed a hierarchical deterministic (HD) wallet that effectively manages multiple private keys. Still, the hierarchical characteristics make the private keys have a fixed relationship and can have a certain degree of security and privacy. He *et al.* [14] proposed a new cryptocurrency wallet management scheme based on Decentralized Multi-Constrained Derangement (DMCD) to safely and stably store keys in a decentralized network. Wei *et al.* [35] proposed a new digital signature algorithm and used it to design an online wallet that can help users derive signatures without obtaining the user's private key.

With the rise of digital assets in recent years, wallet security has received much attention. As long as the private key is accidentally stolen or lost, severe losses will be caused. Therefore, custody can be achieved by dividing the private key, and effectively protecting the security and integrity of the private key is a research topic worth exploring. This research will solve the problems caused by general private key segmentation and propose feasible solutions to protect private keys. Designing a smart contract through the blockchain can avoid problems such as stealing a subkey or cheating to recover the key. This research topic needs to address the following questions:

1) Improve the private key segmentation of the SS scheme

   As far as the SS solution is concerned, today's private key hosting will have a central role. If executed through blockchain smart contracts, such problems can be avoided. This research can design a secure and private private key storage and hosting solution to solve this problem. The concept of the SS solution will be combined with the decentralized storage method of the blockchain to design a private key subkey storage solution that is both secure, private, and scalable.

2) Key management protocols have some weaknesses.

   The (t, n) threshold (Threshold) secret sharing scheme in the SS scheme can be used in different cryptographic applications. However, there will be some vulnerabilities, such as shared deception vulnerabilities and misbehavior by central actors. The main reason for this type of weakness is the lack of verifiability of the shares generated by the center. In addition, other key management protocols have some weaknesses, and the designed encryption tools are still insecure. For example, Diffie–Helman (DH) does not protect against DoS (Denial of Service) and MITM (Man-in-the-Middle) attacks. If these attacks are not adequately addressed, cryptographic systems will not be developed. In addition, the SS scheme has the weakness of dishonest participants cheating on their shares, and honest participants are forced to reconstruct fake secrets, etc. [23].

To summarize the above, this research project will solve two main problems:

1) Use blockchain smart contracts to design a framework for securely storing private key fragments so that both the user's private key and the private key fragments can be safely stored in Platform;

2) It can resist various attacks and improve existing vulnerabilities to ensure the security and integrity of users' private keys.

3.2 Research Topic 2: Research and Development to Prevent Illegal Income from Escaping Seizure of Virtual Currency

Using virtual currency as a criminal tool has become increasingly common. As an emerging encrypted virtual currency, the high risks of money laundering and crime hidden in virtual cryptocurrency have already attracted the attention of governments and judicial authorities in various countries. In the first half of 2022, tens of thousands of fraud cases have been reported, with financial losses exceeding 3 billion yuan. Fraudulent groups use cryptocurrency and other methods as money laundering tools, making it more difficult to recover lost amounts. In addition to new issues related to criminal law and criminal prosecution, asset recovery and seizure issues arise due to the asset value of cryptocurrencies. The main problem here is that cryptocurrencies are neither items nor rights. Confiscation need not be considered if the virtual currency units generated due to criminal conduct will likely lose value. Ensuring the security of Bitcoin during the investigation is also a matter of substantive significance.

Lee proposed using secret sharing [19]: split the gold key and then use the user's login password to generate the gold and trading platform keys after FIDO (Fast IDentity Online) authentication of the user's device. These keys are used as secret shared encryption keys after being unpacked, and at least two can be pushed back to the original keys to protect the risk of password loss or theft. After collecting relevant information on cryptocurrency trading platforms, this study found that the platform service provided by LoclEtherm is a more secure channel that allows users to control their keys, making it impossible for platform managers to obtain user keys. , all transactions are written on the blockchain, but it still fails to solve the problem of lost user keys. Therefore, it is proposed to use the FIDO standard authentication mechanism combined with the sharing method; it can also maintain security in addition to solving this problem.

The second research topic is based on the mechanism of the first research topic and develops the legal seizure of illegal income. Mainly because, in recent years, more and more crimes have been related to virtual currencies. Assume that the suspect's wallet is now seized, and the relevant police inspectors need to create a new wallet. The wallet will have a key, and you need to decide who to split the key to, then transfer the criminal's wallet to a new wallet for seizure. After the investigation, the money will be returned or compensated to the victim. This research topic will develop a legal seizure mechanism for illegal gains from virtual currencies to solve this problem. The narrative is as follows:

1) Solving the security issues of wallet custody

   After the wallet is seized, to prevent the criminal from conspiring with his accomplices, the private key is provided to the accomplices within a short period, and they take the opportunity to remove the illegal gains. Therefore, it is necessary to design adequate supervision to immediately transfer illegal gains to the supervisory personnel's supervision wallet and obtain reasonable control.

2) Solve the internal control problems of supervisors

   When a wallet is seized to prevent supervisors from stealing it, if only one person has the key to the wallet, the security of the seizure cannot be guaranteed. Therefore, a mechanism is designed to keep encrypted assets safe. Multiple people can hold private keys to solve the security problem of wallet seizure.

3.3 Research Topic 3: Research and Development on Secure Transfer of Digital Asset Heritage

Unlike in the real world, virtual crypto-assets do not disappear when their owner dies. The essential characteristic of encrypted assets is that only those who know the key can transfer assets. Actual ownership of a crypto asset is equivalent to a key, and conversely, everyone who knows the key is considered the legal owner of the relevant crypto asset. In general, a person cannot inherit their crypto assets. Because the keys disappear when the person dies, or the ownership of these crypto assets must be shared with a third party. However, the main problem is that legal owners can only transfer encrypted virtual assets. Conversely, nothing can be transferred as long as the holder dies, making the encrypted assets' inheritance complex [30, 31].

An essential characteristic of cryptoassets is that no one can transfer wealth except the owner of the corresponding key. To resolve inheritance issues, the easiest way is to write the private key into your will. However, this is not in the spirit of cryptocurrency because it relies on trusting a third party, which conflicts with the basic concept of cryptocurrency. Therefore, it is necessary to find ways to split private keys and strengthen and design smart contracts through the blockchain key custody mechanism developed by the research theme. Achieving the inheritance of crypto assets aligns with the original spirit of cryptocurrency.

This research topic will use the research topic 1 R&D mechanism to implement relevant applications. As we live in the digital age, after we pass away in the future, we hope to transfer digital assets to our loved ones through smart contract triggering, and we can make a will to decide in advance who we want to give it to. Who can use these? The triggering conditions for key distribution are activated through smart contracts. It will only be activated when the holder dies and is distributed to their respective wallets. To withdraw money later, these people need to present it simultaneously to achieve the digital asset distribution goal: fairness and integrity. However, as the data on the chain increases, the scale of the chain will become larger and larger, which will consume considerable costs and resources for the nodes.

4 Conclusion

This article has proposed three future research topics of blockchain secret key sharing and its digital asset applications. The first topic is the research and development of blockchain key custody. First, use blockchain to design an encrypted storage architecture to ensure patient privacy and integrity. Topic two is based on the structure of Topic one and extends to the blockchain part. Expand application field research on the legal seizure of illegal proceeds from virtual currency. The third topic is to solve the transfer of digital asset inheritance. Develop private key-splitting methods to make transferring digital assets on the blockchain more efficient. Therefore, integrating these three research yopics is a platform system that can be fully applied to user private key segmentation and can have privacy, security, and integrity.

This research will be able to achieve the following three goals:

1) Ensure the security of private key custody

   This project will research this issue and is expected to propose a solution to protect users' private keys and propose a framework that combines security, efficiency, and practicality.

2) Improve the problem of private key division and centralization

   The private key custody framework designed in this project is implemented through blockchain smart contracts to solve the problem of the central role of the traditional SS solution.

3) Solve the legal issues of seizure and custody of illegal gains from cryptocurrency and inheritance of digital assets

   This study predicts that the private key custody smart contract proposed in this study will be applied to the legal seizure and custody of illegally obtained cryptocurrency and the inheritance of digital assets, breaking through the scalability limitations and improving the applicability of the blockchain.

Acknowledgments

The Ministry of Science and Technology partially supported this research, Taiwan (ROC), under contract no.: NSTC 109-2221-E-468-011-MY3, NSTC 112-2221-E-507 -005 -MY3, and NSTC 111-2622-8-468-001 -TM1.

References

[1] A. K. Biswas, M. Dasgupta, S. Ray, M. K. Khan, "A probable cheating-free (t, n) threshold secret sharing scheme with enhanced blockchain", *Computers and Electrical Engineering*, vol. 100, 2022.

[2] D. Boneh, R. Gennaro, S. Goldfeder, "Using level-1 homomorphic encryption to improve threshold dsa signatures for bitcoin wallet security", *Lecture Notes in Computer Science*, vol. 11368, pp. 352–377, 2017.

[3] C. C. Chang, M. S. Hwang, "Parallel computation of the generating keys for RSA cryptosystems", *Electronics Letters*, vo1. 32, no. 15, pp. 1365–1366, 1996.

[4] P. Y. Chang, M.S. Hwang, C.C. Yang, "A blockchain-based traceable certification system", in *Security with Intelligent Computing and Big-data Services*, pp. 363-369, 2018.

[5] T. Y. Chang, M. S. Hwang, and W. P. Yang, "A new multi-stage secret sharing scheme using one-way function", *ACM Operating Systems Review*, vol. 39, no. 1, pp. 48–55, Jan. 2005.

[6] T. Y. Chang, M. S. Hwang, W. P. Yang, "An improvement on the Lin–Wu (t, n) threshold verifiable multi-secret sharing scheme," *Applied Mathematics and Computation*, vol. 163, no. 1, pp. 169-178, 2005.

[7] T. Y. Chang, M. S. Hwang, W. P. Yang, "An improved multi-stage secret sharing scheme based on the factorization problem," *Information Technology and Control*, vol. 40, no. 3, pp. 246-251, 2011.

[8] L. Chen, Q. Fu, Y. Mu, L. Zeng, F. Rezaeibagha, M. S. Hwang, "Blockchain-based random auditor committee for integrity verification", *Future Generation Computer Systems*, vol. 131, pp. 183-193, 2022.

[9] Y. H. Chen, L. C. Huang, I. C. Lin, M. S. Hwang, "Research on the secure financial surveillance blockchain systems", *International Journal of Network Security*, vol. 22, no. 4, pp. 708-716, 2020.

[10] Y. H. Chen, L. C. Huang, I. C. Lin, M. S. Hwang, "Research on blockchain technologies in bidding systems", *International Journal of Network Security*, vol. 22, no. 6, pp. 897-904, 2020.

[11] P. Das, S. Faust, J. Loss, "A formal treatment of deterministic wallets", in *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security (CCS'19)*, pp. 651–668, 2019.

[12] P. Dikshit and K. Singh, "Efficient weighted threshold ECDSA for securing bitcoin wallet", in *ISEA Asia Security and Privacy*, pp. 1–9, 2017.

[13] G. Gutoski, D. Stebila, "Hierarchical deterministic bitcoin wallets that tolerate key leakage", in *Financial Cryptography*, pp. 497–504, 2015.

[14] X. He, J. Lin, K. Li and X. Chen, "A novel cryptocurrency wallet management scheme based on decentralized multi-constrained derangement", *IEEE Access*, vol. 7, pp. 185250-185263, 2019.

[15] L. C. Huang, S. F. Chiou, M. S. Hwang, "New lossless random grids progressive secret sharing based on XOR for binary-image," in *National Computer Symposium (NCS'19)*, pp. 728-731, 2019.

[16] M. S. Hwang, I. C. Lin, *Introduction to Information and Network Security (6ed, in Chinese)*, Taiwan: Mc Graw Hill, 2017.

[17] M. S. Hwang and W. P. Yang, "Conference key distribution protocols for digital mobile communication systems", *IEEE Journal on Selected Areas in Communications*, vo1. 13, no. 2, pp. 416-420, Feb. 1995.

[18] Z. Jian, Q. Ran and S. Liyan, "Securing blockchain wallets efficiently based on threshold ECDSA scheme without trusted center", in *Asia-Pacific Conference on Communications Technology and Computer Science (ACCTCS'21)*, pp. 47–51, 2021.

[19] Y. S. Lee, *Key Management for Cryptocurrency Exchange Platform*, Thesis, National Chengchi University, 2019.

[20] N. Lehto, K. Halunen, O. M. Latvala, A. Karinsalo and J. Salonen, "CryptoVault - A secure hardware wallet for decentralized key management", in *IEEE International Conference on Omni-Layer Intelligent Systems (COINS'21)*, pp. 1-4, 2021.

[21] C. T. Li, M. S. Hwang, "A lightweight anonymous routing protocol without public key en/decryptions for wireless ad hoc networks",*Information Sciences*, vol. 181, no. 23, pp. 5333-5347, 2011.

[22] C. T. Li, M. S. Hwang, "An online biometrics-based secret sharing scheme for multiparty cryptosystem using smart cards," *International Journal of Innovative Computing, Information and Control*, vol. 6, no. 5, pp. 2181-2188, May 2010.

[23] G. Li, L. You, "A consortium blockchain wallet scheme based on dual-threshold key sharing", *Symmetry*, vol. 13, no. 8, p. 1444, 2021. (`https://doi.org/10.3390/sym13081444`)

[24] B. Y. Lin, *Research for Recovery and Saving Private Key*, Thesis, Tamkang University, 2021.

[25] C. Y. Lin, L. C. Huang, Y. H. Chen, M. S. Hwang, "Research on security and performance of blockchain with innovation architecture technology", *International Journal of Network Security*, vol. 23, no. 1, pp. 1-8, 2021.

[26] I. C. Lin, M. S. Hwang, C. C. Chang, "A new key assignment scheme for enforcing complicated access control policies in hierarchy", *Future Generation Computer Systems*, vol. 19, no. 4, pp. 457-462, May 2003.

[27] O. Pal, B. Alam, V. Thakur, S. Singh, "Key management for blockchain technology", *ICT Express*, vol. 7, no. 1, pp. 76–80, 2021.

[28] S. S. Panda, D. Jena, B. K. Mohanta, S. Ramasubbareddy, M. Daneshmand, A. H. Gandomi, "Authentication and key management in distributed IoT using blockchain technology", *IEEE Internet of Things Journal*, vol. 8, no. 16, pp. 12947-12954, 2021.

[29] Z. P. Peng, M. L. Chiang, I. C. Lin, C. C. Yang, M. S. Hwang, "CBP2P: Cooperative electronic bank payment systems based on blockchain technology", *Journal of Internet Technology*, vol. 23, no. 4, pp. 683-692, 2022.

[30] F. Prost, "Inheritance and blockchain: Thoughts and open questions", *arxiv*, 2212.01194, 2022.

[31] F. Prost, "On the heritage of crypto assets – Tales from the crypt protocol", *arxiv*, 2209.11194, 2022.

[32] H. Rezaeighaleh, C. C. Zou, "New secure approach to backup cryptocurrency wallets", in *IEEE Global Communications Conference*, pp. 1-6, 2019.

[33] M. Tompa, H. Woll, "How to share a secret with cheaters", *Journal of Cryptology*, vol. 1, pp. 133–138, 1989.

[34] A. R. Thota, P. Upadhyay, S. Kulkarni, P. Selvam and B. Viswanathan, "Software wallet based secure participation in hyperledger fabric networks", in *International Conference on COMmunication Systems & NETworkS (COMSNETS'20)*, pp. 1–6, 2020.

[35] Q. Wei, S. Li, W. Li, H. Li, M. Wang, "Decentralized hierarchical authorized payment with online wallet for blockchain", *Lecture Notes in Computer Science*, vol. 11604, pp. 358–369, 2019.

[36] C. C. Wu, M. S. Hwang, S. J. Kao, "A new approach to the secret image sharing with steganography and authentication," *The Imaging Science Journal*, vol. 57, no. 3, pp. 140-151, 2009.

[37] C. C. Wu, S. J. Kao, W. C. Kuo, M. S. Hwang, "Reversible secret image sharing based on shamir's scheme," in *Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pp. 1014-1017, 2009.

[38] T. Zhou, J. Shen, Y. Ren, S. Ji, "Threshold key management scheme for blockchain-based intelligent transportation systems", *Security and Communication Networks*, vol. 2021, Article ID 1864514, 2021. (`https://doi.org/10.1155/2021/1864514`)

Biography

Chia-Chun Wu received a Ph.D. degree from the Department of Computer Science and Engineering, National Chung-Hsing University, Taichung, Taiwan, in 2011. He is currently an associate professor at the Department of Industrial Engineering and Management, National Quemoy University, Kinmen County, Taiwan. His current research interests include artificial intelligence, internet of things (IoT), database security, secret image sharing, mobile applications development, and digital image techniques.

Iuon-Chang Lin received the B.S. in Computer and Information Sciences from Tung Hai University, Taichung, Taiwan, Republic of China, in 1998; the M.S. in Information Management from Chaoyang University of Technology, Taiwan, in 2000. He received his Ph.D. in Computer Science and Information Engineering in March 2004 from National Chung Cheng University, Chiayi, Taiwan. He is currently a professor of the Department of Management Information Systems, National Chung Hsing University, and Department of Photonics and Communication Engineering, Asia University, Taichung, Taiwan, ROC. His current research interests include electronic commerce, information security, cryptography, and mobile communications.

Min-Shiang Hwang received M.S. in industrial engineering from National Tsing Hua University, Taiwan in 1988, and a Ph.D. degree in computer and information science from National Chiao Tung University, Taiwan, in 1995. He was a professor and Chairman of the Department of Management Information Systems, NCHU, during 2003-2009. He was also a visiting professor at the University of California (UC), Riverside, and UC. Davis (USA) during 2009-2010. He was a distinguished professor of the Department of Management Information Systems, NCHU, during 2007-2011. He obtained the 1997, 1998, 1999, 2000, and 2001 Excellent Research Award of National Science Council (Taiwan). Dr. Hwang was a dean of the College of Computer Science, Asia University (AU), Taichung, Taiwan. He is currently a chair professor with the Department of Computer Science and Information Engineering, AU. His current research interests include cryptography, Steganography, and mobile computing. Dr. Hwang has published over 300+ articles on the above research fields in international journals.

Guide for Authors

International Journal of Network Security

IJNS will be committed to the timely publication of very high-quality, peer-reviewed, original papers that advance the state-of-the art and applications of network security. Topics will include, but not be limited to, the following: Biometric Security, Communications and Networks Security, Cryptography, Database Security, Electronic Commerce Security, Multimedia Security, System Security, etc.

1. Submission Procedure

Authors are strongly encouraged to submit their papers electronically by using online manuscript submission at http://ijns.jalaxy.com.tw/.

2. General

Articles must be written in good English. Submission of an article implies that the work described has not been published previously, that it is not under consideration for publication elsewhere. It will not be published elsewhere in the same form, in English or in any other language, without the written consent of the Publisher.

2.1 Length Limitation:

All papers should be concisely written and be no longer than 30 double-spaced pages (12-point font, approximately 26 lines/page) including figures.

2.2 Title page

The title page should contain the article title, author(s) names and affiliations, address, an abstract not exceeding 100 words, and a list of three to five keywords.

2.3 Corresponding author

Clearly indicate who is willing to handle correspondence at all stages of refereeing and publication. Ensure that telephone and fax numbers (with country and area code) are provided in addition to the e-mail address and the complete postal address.

2.4 References

References should be listed alphabetically, in the same way as follows:

For a paper in a journal: M. S. Hwang, C. C. Chang, and K. F. Hwang, ``An ElGamal-like cryptosystem for enciphering large messages,'' *IEEE Transactions on Knowledge and Data Engineering*, vol. 14, no. 2, pp. 445--446, 2002.

For a book: Dorothy E. R. Denning, *Cryptography and Data Security*. Massachusetts: Addison-Wesley, 1982.

For a paper in a proceeding: M. S. Hwang, C. C. Lee, and Y. L. Tang, ``Two simple batch verifying multiple digital signatures,'' in *The Third International Conference on Information and Communication Security (ICICS2001)*, pp. 13--16, Xian, China, 2001.

In text, references should be indicated by [number].

Subscription Information

Individual subscriptions to IJNS are available at the annual rate of US$ 200.00 or NT 7,000 (Taiwan). The rate is US$1000.00 or NT 30,000 (Taiwan) for institutional subscriptions. Price includes surface postage, packing and handling charges worldwide. Please make your payment payable to "Jalaxy Technique Co., LTD." For detailed information, please refer to http://ijns.jalaxy.com.tw or Email to ijns.publishing@gmail.com.