IJNS
International Journal of Network Security

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

Vol. 26, No. 5 (September 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

Volume: 26, No: 5 (September 1, 2024)

International Journal of Network Security

A Regulatable Privacy Protection Scheme for Transactions Based on Consortium Blockchain

Pengshou Xie, Wanjun Shao, Tao Feng, Ye Lu, Yinchang Pan, and Xinyu Fan
(Corresponding author: Wanjun Shao)

School of Computer and Communications, Lanzhou University of Technology
No. 36 Peng Jia-ping road, Lanzhou, Gansu 730050, China
Email: 2443404684@qq.com

(Received June 16, 2023; Revised and Accepted Sept. 22, 2023; First Online Aug. 16, 2024)

Abstract

Existing privacy protection methods for blockchain transactions achieve strong privacy without revealing any information but are challenging to regulate. To address the problem of simultaneous privacy and regulation, a regulatable privacy protection scheme for transactions based on consortium blockchain is proposed. The scheme is deployed on the consortium blockchain, red using group signature and one-time random address to hide the identity information of both parties in the transaction, using homomorphic encryption and zero-knowledge proof to encrypt the transaction amount, and realizing the legitimacy verification of the transaction amount. In order to avoid the abuse of rights, audit nodes, and tracking nodes are introduced to play the role of regulatory nodes. Theoretical analysis and experimental results show that the scheme has strong privacy and reliability and reasonable computational overhead.

Keywords: Blockchain; Homomorphic Encryption; Privacy Protection; Regulable; Zero-knowledge Proof

1 Introduction

As a distributed ledger-sharing technology, blockchain is decentralized, tamper-proof, traceable, and jointly maintained by multiple parties [2,11], which can reduce transaction costs and establish trusted value transfer between peers. The features of blockchain make it widely used in industries such as finance, energy, the Internet of Things, and healthcare [1, 21]. Public blockchain, consortium blockchain, and private blockchain [6, 30] are the three main application scenarios of blockchain at present. Public blockchain does not restrict the joining or withdrawal of nodes, while only specific authorized nodes can join in private blockchain and consortium blockchain. The consortium blockchain is a blockchain managed by multiple institutions, and the data in the chain is only allowed to be read and recorded by member nodes, which streamlines the number of access nodes and bookkeeping nodes and improves the efficiency of transaction processing and platform operation. Therefore, the consortium blockchain is mainly an institution-to-institution specific application and is widely used in the fields of commodity traceability, supply chain management, and asset trading.

However, the ledger on the blockchain is publicly stored, and the data records and operation rules can be reviewed and traced by the nodes of the whole network with high transparency. While this storage method brings convenience, it also causes the risk of user privacy leakage to a certain extent [17]. Attackers can effortlessly obtain these transaction data exposed on the blockchain [15, 28] and maliciously mine users' transaction habits, transaction rules, and other information through big data analysis technology, which seriously threatens users' personal information. The data publicly available on the blockchain is both the user's personal privacy information and the core confidential data of the organization.To further ensure the privacy of transaction data on the blockchain, there have been some studies [3, 19] to hide the public data on the chain through cryptographic techniques such as coin mixing mechanisms and ring signatures, which increase the difficulty of data analysis. However, the strong privacy protection strategy leads to unregulated transactions on the blockchain, which invariably builds a solid protection barrier for various illegal and criminal activities, so that blockchain gradually degenerates into a tool for underground money laundering, tax evasion, extortion, and illegal transactions.Therefore, there is an urgent need for a blockchain transaction method that can balance privacy and regulation.

Aiming at the above problems, this paper proposes a regulatable privacy protection scheme for transactions based on consortium blockchain using group signature, one-time random address, Paillier homomorphic encryption and Bulletproof zero-knowledge proof. It not only realizes the privacy protection of the identity information of both parties of the transaction and the transaction amount but also enables the transaction amount to be in the legal positive value interval and circumvents the ap-

pearance of negative value. In order to avoid the abuse of rights, the regulatory nodes have audit nodes and tracking nodes together to realize the regulation of the transaction process.

The remainder of this paper is organized as follows. Section 2 describes the research work related to privacy protection of blockchain transactions, Section 3 describes the blockchain technology and related cryptographic knowledge used in this paper, Section 4 describes the scheme, Section 5 describes the specific implementation of the scheme, Section 6 provides theoretical analysis and experimental validation of the scheme, and Section 7 concludes the paper.

2 Related Works

Privacy issues on the blockchain include two main components: user identity privacy and transaction content privacy [25, 31]. User identity privacy in cryptocurrencies is achieved through pseudonyms, but pseudonyms do not lead to anonymity if a user's transactions are linkable, and UCoin [22] mixes transactions from multiple users into a single aggregated transaction to complete, breaking the linkability between transaction input and output addresses. Using pseudonyms alone does not fully ensure anonymity and unlinkability, SofitMix [27] relies on a hybrid server that not only achieves anonymity in a Bitcoin-compatible manner but is also effective against DoS attacks and collusion attacks. Complete anonymity using the ring signature technique can guarantee the privacy of user identity in blockchain [16, 32], but complete anonymity cannot be traced to malicious users. Group signatures [9, 14, 29] not only have good anonymity, but also group managers can track specific signature users, which protects the privacy of user identity while regulating the transaction behavior on the blockchain. The above methods are effective in achieving the concealment of transaction user identity, but ignore the concealment of transaction amount.

The symmetric key in QSHE [4] is generated by the transaction participants using the Diffie-Hellman key exchange protocol to generate a specific transaction key to hide the transaction amount in the cryptocurrency. Dumb Account [26] uses Paillier homomorphic encryption algorithm to achieve the hiding of the transaction amount and verifies the ciphertext amount by a commitment proof. Zero-knowledge proofs prove the correctness or incorrectness of an assertion without revealing any other information. Efficient non-interactive zero-knowledge proofs are constructed based on a homomorphic public key encryption scheme on account model blockchain [20], which not only hides the transaction amount and balance information but also ensures the validity of the transaction. Based on homomorphic encryption, Li [13] proposed a blockchain privacy-preserving algorithm that can both encrypt transaction data and support zero-knowledge proofs. Smart contracts are combined with zero-knowledge proofs [8] to verify the consistency and availability of data without revealing any data privacy. These schemes have achieved some results in achieving transaction amount hiding and verification, but privacy during transactions is more than just transaction amount information.

Zerocoin uses zk_SNARK zero-knowledge proof technology [23], which allows complete hiding of information such as the two parties and the transaction amount. BlockMaze [10], together with the double balance model, achieves strong privacy protection by hiding the account balance, the transaction amount, and the link between the sender and the receiver. Ring-secret transaction protocols [12] are widely used in cryptocurrencies to protect the privacy of user identities and transaction amounts, but the expensive computation and storage costs reduce the performance of the system. Aggregated ring-secret transactions [7] construct a compact aggregated signature for multiple accounts of the sender that can linearly reduce the signature size. The MimbelWimble protocol and aggregated signatures to achieve privacy protection for blockchain transactions [5], but the interaction between the two parties of the transaction needs to be online, which is not convenient in practice. Monero [18] provides a high degree of user and transaction anonymity, but many criminal activities may be carried out under the anonymity protection of cryptocurrency transactions. Therefore traceability is particularly important in cryptocurrencies.

In summary, most of the existing research proposals fail to take into account both the anonymity and traceability of blockchain systems.

3 Preliminaries

3.1 Consortium Blockchain

Blockchain is a decentralized infrastructure and distributed computing paradigm that incorporates multiple technologies. Blockchain systems can be divided into two categories: licensed and unlicensed blockchain, depending on whether new nodes need authorization and authentication. Unlicensed blockchain, also known as public blockchain, is completely open and each node is free to join or exit the network without the need for anyone to grant permission or authorization. The licensed blockchain means that each node in the participating system is licensed, and unauthorized nodes are not allowed to access the system. Licensed blockchain is further divided into consortium blockchain and private blockchain [30]. Where private blockchain is completely private, and only designated members have ledger read and write access; consortium blockchain is between the public and private blockchain and is a kind of polycentric or partially decentralized blockchain. Multiple organizations form a stakeholder alliance to jointly maintain the healthy operation of the blockchain. The consortium blockchain is widely welcomed in various industries because it combines the

openness and low trust of the public blockchain and the privacy protection and single high trust of private chains.

3.2 Group Signature

Group signatures extend signatures to the group environment, allowing group members to generate valid signatures on messages in the name of the group that can only be verified as having been generated by a member of the group, and cannot be specifically identified to a specific member. When a signature is disputed, the group manager can open the group signature to determine the specific identity information of the signer.

The group signature used in this paper is an improvement of the SM9-based multi-KGC group signature proposed by Yang et al [29]. Unlike it, this paper uses the group manager and the tracking nodes to negotiate to generate the key, and the identity information of the nodes is kept by the group manager, which no longer sets multiple independent KGCs, reducing the number of participants and simplifying the transaction process.

3.3 One-Time Random Address

Cryptonote uses a one-time public-private key pair to construct a one-time random address to hide from the transaction receiver. The sender generates a one-time random address for this transaction based on the permanent public key address of the receiver. Since the selection of random numbers is random, the address generated for each transaction cannot be repeated even for the same receiver, breaking the correlation between different transactions of the receiver.

The one-time random address is described as follows:

Parameter settings: G is a base point on the elliptic curve, p is its prime order; Hs is a cryptographic hash function. Assume that the permanent public key address of the transaction re ceiver is (A, B), and the permanent private key address is (a, b); where A=aG, B=bG; a is stored jointly by the tracking nodes and the transaction receiver as a tracking key, and b is stored individually by the transaction receiver as a unique key.

One-time random address generation: The transaction sender selects a random number $r \in [1, p-1]$, calculates R=rG, calculates the one-time random address $Address = Hs(rA)G + B$, broadcasts R and Address.

Authentication: The receiver uses the private key a and the public key B to compute $Address' = Hs(aR)G + B$, and if Address=Address', the authentication passes.

One-time private key calculation: The receiver uses the private key (a, b) to calculate the corresponding one-time private key $d = Hs(aR) + b$.

3.4 Paillier Homomorphic Encryption

Paillier homomorphic encryption algorithm is a probabilistic public key encryption algorithm. The advantage is that it is simple, easy to implement, and satisfies additive homomorphism and number multi-plication homomorphism, which is also known as the semi-homomorphic encryption algorithm. With the Paillier homomorphic encryption algorithm, the ciphertext information of the transaction amount and account balance can be added without decryption to verify the legitimacy of the transaction amount.

The Paillier homomorphic encryption algorithm is as follows:

Key Generation: $KeyGen \rightarrow (PK, SK)$. Randomly select two large prime numbers p, q to satisfy $gcd(pq, (p-1)(q-1)) = 1$; calculate n=pq, $\lambda = lcm(p-1, q-1)$, where lcm denotes the least common multiple; randomly selected integers $g \leftarrow Z_{n^2}^*$, define the function $L(x) = \frac{x-1}{n}$, calculate $\mu = \left(L\left(g^\lambda \bmod n^2\right)\right)^{-1} \bmod n$; This generates the public key $PK = (n, g)$, the private key $SK = (\lambda, \mu)$.

Encryption process: $Encrypt(m, PK) \rightarrow c_m$. Enter a plaintext message m and the public key PK, select a random number $r \in (0, n)$, and output ciphertext $c_m = g^m r^n \bmod n^2$.

Decryption process: $Decrypt(c_m, SK) \rightarrow m$. Enter the ciphertext c_m and the private key SK , output plaintext $m = L\left(c_m \lambda \bmod n^2\right)\mu \bmod n$.

3.5 Zero-Knowledge Proof

In a zero-knowledge proof, the prover can convince the verifier that an assertion is correct without providing any information to the verifier. A zero-knowledge proof is essentially an agreement involving two or more parties, i.e., a series of steps that two or more parties need to take to complete a task.

Bulletproofs is a range proof designed using the inner product method, which has a short proof time and does not require a confidential set. Bulletproofs zero-knowledge proof algorithm uses a "range proof" approach to prove that the amount of the transaction is within a given positive value interval, avoiding negative values, which is also essential to verify the legitimacy of the transaction. It relies on Pedersen commitments to hide secret inputs and provide computational integrity checks.

4 Description of Privacy Protection Methods for Regulated Transactions

4.1 Node Composition

The privacy protection method for regulated transactions uses a consortium blockchain based on the account model,

which is similar to the storage model of banks and is more efficient and intuitive, and prevents replay attacks through the nonce field. As shown in Figure 1, the scheme mainly consists of three types of nodes: group manager (GM), user nodes (UN), and regulatory nodes (RN).

1) Group manager: Combining consortium chains with the concept of groups in group signatures, a group manager is set up for the consortium chain, which is responsible for managing the joining and exiting of user nodes, updating the revocation list, and distributing relevant keys. In order to prevent a single group manager from committing evil, the group key pair is generated by the tracking nodes and the group manager together in the key distribution phase.

2) User nodes: User nodes are the main participating nodes in the transaction process, consisting of the transaction sender (TS) and transaction receiver (TR). When a transaction is conducted between user nodes, TS generates transaction information using techniques such as group signature, one-time random address, Paillier homomorphic encryption and Bulletproof zero-knowledge proof, and packages it for uploading to the chain by the Raft consensus algorithm.

3) Regulatory nodes: Regulatory nodes do not participate in the transaction process, but can regulate the transaction behavior, mainly responsible for verifying the legitimacy of the transaction, tracking the illegitimate transaction information, as well as packaging the legitimate transaction hash on the chain. In order to prevent the regulatory nodes from having too much power and to avoid the phenomenon of power abuse, there are audit nodes (AN) and tracking nodes (TN) together to form the regulatory nodes to regulate the transaction information.

4.2 Implementation Processes

A regulatable privacy protection scheme for transactions based on consortium blockchain adopts group signature, one-time random address, Paillier homomorphic encryption and Bulletproof zero-knowledge proof to realize the privacy protection of the identity information of both parties of the transaction and the transaction amount. In order to avoid the abuse of rights, the regulatory nodes have audit nodes and tracking nodes together to realize the regulation of the transaction process, and the implementation processes of the method is shown in Figure 2.

In the transaction between user nodes, TS first sends a request to GM to join the consortium chain, GM verifies the legitimacy of its identity, agrees to TS's request to join the chain, records its identity information, and generates a group key pair for TS in collaboration with TN; TS generates a one-time random address of TR, and at the same time, broadcasts the encrypted transaction amount to the blockchain after the group signature, which is packaged and uploaded on the blockchain by the Raft consensus algorithm; AN verifies the legitimacy of the transaction, if it passes the verification, the transaction is legitimate, hashes the legitimate transaction block on the blockchain, and updates the account balance information of both parties to the transaction; otherwise, it is an illegitimate transaction, and sends a request for tracking the transaction to TN, which receives the request, and traces the identities of both parties to the transaction through the tracking key and other information.

5 Design of Privacy Protection Methods for Regulated Transactions

5.1 System Initialization

5.1.1 Initialization parameters

Input safety factor λ, output the initialization parameters of the system $pp = \{\lambda, g_1, g_2, g, h, G_1, G_2, G_T, H_1, H_2\}$, G_1 and G_2 are additive cyclic groups, G_T is a multiplicative cyclic group, the orders of G_1, G_2 and G_T are all large prime numbers N, g_1 and g_2 are the generators of groups G_1 and G_2 respectively, $g, h \in G_T$, H_1 and H_2 are derived cryptographic functions of the hash function.

Initialize an empty undo list RL, used to store the undo tag of the undo user.

5.1.2 User registration application

The transaction sender TS submits a registration application to the group manager GM, who verifies the identity of TS and agrees to its application to join the group and records the identity information ID_{TS} of TS. The group manager GM and the tracking nodes TN agree on the master key $ks \in [1, N-1]$, then the master public key is $P_s = ks \cdot g_2$, namely the master key pair (ks, P_s).

Group logo $ID_G = ID_{GM} \parallel ID_{TN}$,where ID_{GM} and ID_{TN} are the identity information of the group manager and the tracking nodes, respectively.

Calculate $t_1 = H_1(ID_{TS} \parallel hid, N) + ks$, $d_1 = ks \cdot t_1^{-1}$, and get TS's private key $ds_{TS} = d_1 \cdot g_1$;

Calculate $t_2 = H_1(ID_G \parallel hid, N) + ks$, $d_2 = ks \cdot t_2^{-1}$, and get the group private key $ds_G = d_2 \cdot ds_{TS}$.

The group key pair of TS is $(ds_{TS}, ds_G, ID_{TS}, ID_G)$.

5.2 User Nodes Generate Transactions

5.2.1 Zero-knowledge proof generation

TS generates zero-knowledge proofs that enable the audit nodes to verify $m \in [0, 2^n - 1]$ without revealing the transaction amount m.

Construct a_L, a_R so that they satisfy $\langle a_L, 2^n \rangle = m$, $a_R = a_L - 1^n$; construct a commitment $C_a = h^{\alpha} g^{a_L} h^{a_R}$ of a_L, a_R;

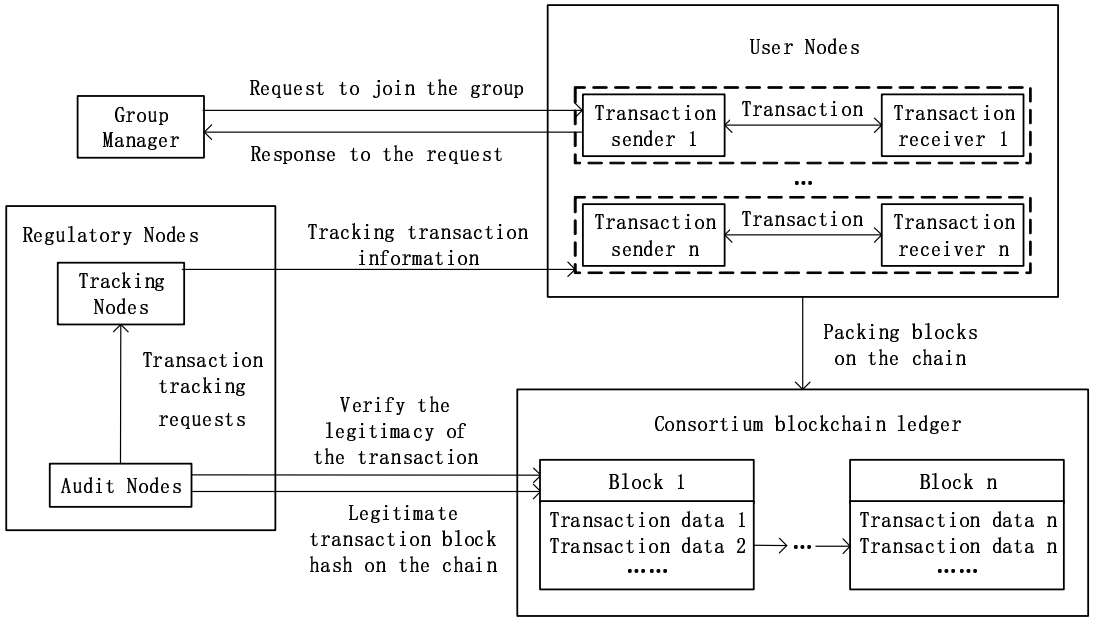


Figure 1: Node composition of privacy protection methods for regulated transactions

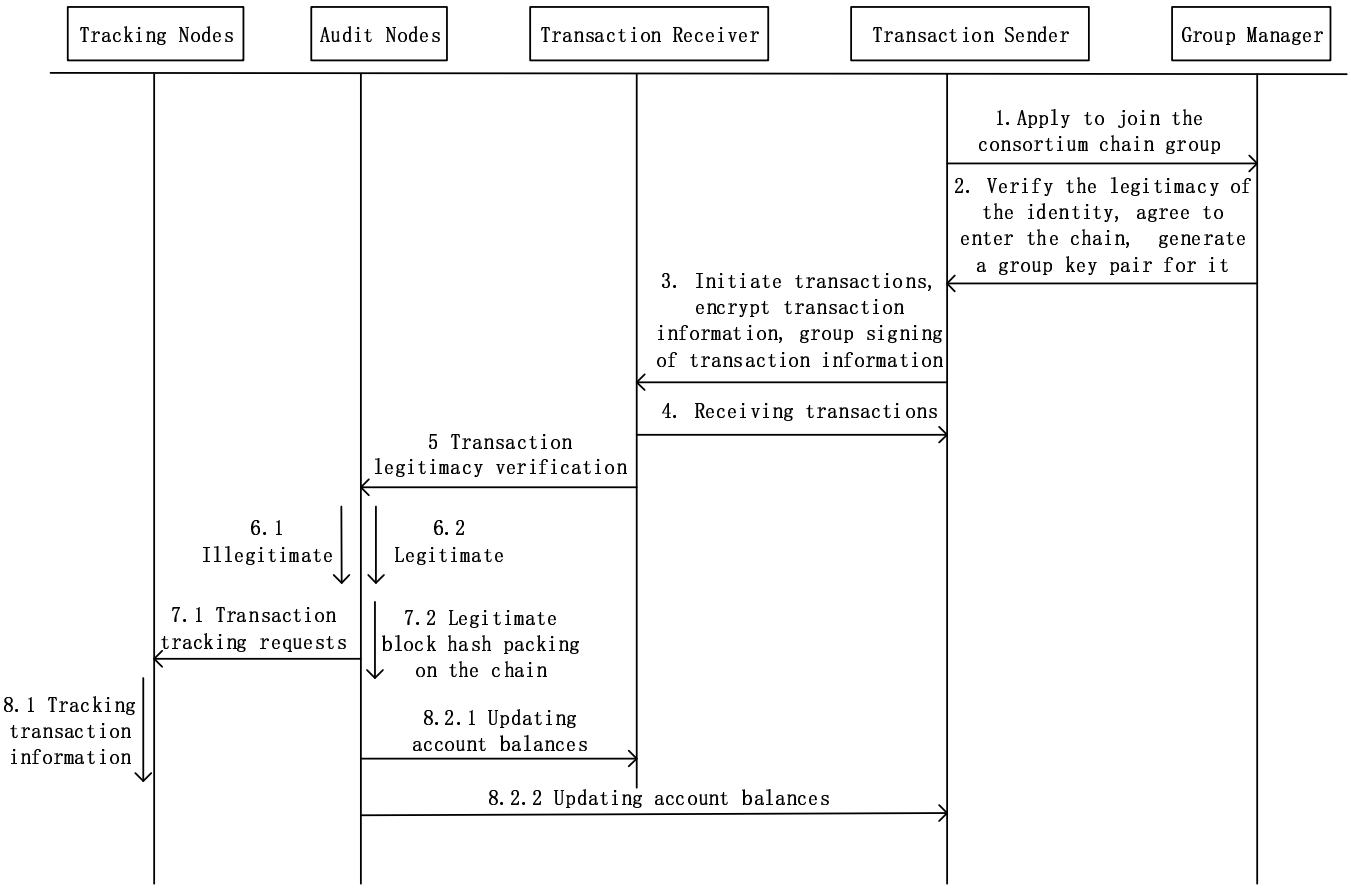


Figure 2: Implementation processes of privacy protection methods for regulated transactions

Randomly selected blind factors s_L, s_R, construct a commitment $C_s = h^{\rho} g^{s_L} h^{s_R}$ of s_L, s_R;

$$y = H_1(C_a, C_s) \tag{1}$$

$$z = H_1(C_a, C_s, y) \tag{2}$$

Randomly selected τ_1,τ_2, construct commitments $C_{\tau_i} = g^{t_i} h^{\tau_i}$, $i \in \{1,2\}$ of τ_1,τ_2;

$$x = H_2(C_{\tau_1}, C_{\tau_2}, z) \tag{3}$$

$$l(x) = (a_L - z \cdot 1^n) + s_L \cdot x \tag{4}$$

$$r(x) = (a_R + z \cdot 1^n + s_R \cdot x) \circ y^n + z^2 \cdot 2^n \tag{5}$$

$$t(x) = \langle l(x), r(x) \rangle = t_0 + t_1 \cdot x + t_2 \cdot x^2 \tag{6}$$

$$\tau(x) = \tau_2 \cdot x^2 + \tau_1 \cdot x + z^2 \cdot r \tag{7}$$

where is randomly selected and $r \in Z_p$;

Calculate $\mu = \alpha + \rho \cdot x$, and generate a commitment on m as $C_m = g^r h^m$;

From this, zero-knowledge proves that

$$\eta = \{\tau(x), \mu, t(x), l(x), r(x), C_m\} \tag{8}$$

5.2.2 Group signature generation

TS will compose a transaction message list $M_{TX} = \{Address, c_m, c_{\triangle m}, c_M, \eta\}$ with the one-time random address Address, zero-knowledge proof η, transaction amount cipher c_m, post-transaction account cipher $c_{\triangle m}$ and pre-transaction account cipher c_M, and perform group signature on the transaction message list M_{TX}.

Calculate $g_s = e(g_1, P_s)$, randomly selected $r_1, r_2 \in [1, N-1]$, $w = g_s^{r^2}$;

$$h_{TX} = H_2(M_{TX} \parallel w, N) \tag{9}$$

$$S_1 = \left(r_1^{-1}\right)(r_2 - h_{TX}) \cdot ds_{TS} \tag{10}$$

$$S_2 = \left(r_1^{-1}\right)(r_2 - h_{TX}) \cdot ds_G \tag{11}$$

$$h_1 = H_1(ID_{TS} \parallel hid, N) \tag{12}$$

$$P_{TX} = (h_1 \cdot g_2 + P_s) \cdot r_1 \tag{13}$$

Get the group signature of TS for M_{TX}:

$$Sign = (h_{TX}, P_{TX}, S_1, S_2) \tag{14}$$

The transaction information after th group signature is broadcasted to the blockchain, and the Raft consensus algorithm packages the transaction on the chain.

5.3 Regulatory Nodes Regulate Transactions

5.3.1 Transaction legitimacy verification

Audit nodes verify the legitimacy of transactions on the chain, including transaction identity legitimacy verification and transaction amount legitimacy verification.

1) Transaction identity legitimacy verification

The audit node AN verifies the transaction information on the blockchain to see if the signing user is a member of the group ID_G.

Calculate $h_2 = H_1(ID_G \parallel hid, N)$, $P_2 = h_2 \cdot g_2 + P_s$, $u_1 = e(S_2, P_2)$, $u_2 = e(S_1, P_s)$, and determine if u_1 and u_2 are equal. If they are not equal, the verification does not pass. If equal, continue the calculation $u = e(S_1, P_{TX})$, $g_s = e(g_1, P_s)$, $t = g_s^{h_{TX}^{'}}$, $w^{'} = u \cdot t$, $h_{TX} = H_2\left(M_{TX}^{'} \parallel w^{'}, N\right)$; determine if $h_{TX}^{'}$ and h_{TX} are equal, If they are equal, the message is proven to have been signed by a member of the group ID_G, and the identity of the sender is verified.

2) Transaction amount legitimacy verification

AN uses the additive homomorphism of Paillier homomorphic encryption to verify that the account balance information before the transaction sender TS transaction is equal to the sum of the amount of the transaction and the account balance information after the transaction, namely $c_M \stackrel{?}{=} c_m + c_{\Delta m}$.

AN verifies the zero-knowledge proof η and determines whether the transaction amount m is within a positive value range.

Determine $g^{t(x)} \cdot h^{\tau(x)} \stackrel{?}{=} C_m^{z^2} \cdot g^{C_s} \cdot C_{\tau_1}^{x} \cdot C_{\tau_2}^{x^2}$;

According to a_L, a_R, s_L, s_R generated commitment C_a, C_s, generated commitment $C_p = C_a \cdot C_s^x \cdot g^{(-z)} \cdot \left(h^{'}\right)^{z \cdot y^n + z^2 \cdot 2^n}$ of $l(x)$, $r(x)$, where $h^{'} = \left(h_1, h_2^{y-1}, h_3^{y-2}, \cdots, h_n^{y-n+1}\right)$;

Determine $C_p \stackrel{?}{=} h^{\mu} \cdot g^{l(x)} \cdot \left(h^{'}\right)^{r(x)}$, $t(x) \stackrel{?}{=} \langle l(x), r(x) \rangle$;

If all the above judgments are correct, the transaction amount m sent by TS is legitimate.

5.3.2 Illegitimate transaction tracking

When the AN verifies that at least one of the transaction identity and the transaction amount is illegal, it sends a transaction information tracking request to the tracking nodes TN. The TN can track the identity information of the two parties of the transaction through the information it holds such as the tracking key, and is able to find out the irregularities of the user nodes quickly.

1) Tracking of the transaction sender. When TN wants to verify the source of the group signature information, TN makes an application to the group manager GM to open the group signature, GM verifies the identity of TN, encrypts the held user identity information ID_{TS} with the master public key P_s, TN decrypts it with the master key ks, obtains the user identity information ID_{TS}, calculates the user private key ds_{TS}, and thus traces the real identity information of the signed user.

2) Tracking of the transaction receiver. TN calculates $Address \stackrel{?}{=} H_s(a \cdot R) \cdot G + B$ by the tracking key a. If the equation holds, the one-time random address generated for this transaction is based on the address of the TR, which traces the information of the receiver of this transaction; if the equation does not hold, the identity of the receiver of the transaction is uncovered.

5.4 System Maintenance

The system maintenance phase consists of two main parts: the revocation of group members and the updating of the books of both parties to the transaction.

When a user wants to quit the group voluntarily, or is kicked out of the group due to illegal actions, the group manager GM performs a revocation operation, adds the user's identity and private key information to the revocation list RL, records it as a "revoked flag", and then issues a new revocation list RL.

When the AN verifies that both the transaction identity and the transaction amount are legitimate, the legitimate transaction block hash is packaged and uploaded to the chain through the Raft consensus algorithm, and at the same time, it updates the account balance information of the transaction sender and the transaction receiver in the local database.

6 Scheme Analysis and Validation

6.1 Security Analysis

6.1.1 Anonymity

A scheme is said to be anonymous if the advantage $Adv_A(A)$ of winning in game $Game_1$ for any probabilistic polynomial time attacker A is negligible.

The anonymity of the scheme is described using the following game $Game_1$ between the challenger C and the attacker A. The attacking game consists of the following five phases.

Initial phase: C runs the system to build the algorithm, enters the safety factor λ, and sends the output system parameters pp to A.

Query phase: A executes the following queries with polynomial boundedness.

1) Key pair query: A selects a user identity u_1 and requests a public-private key pair from C. C runs the setup public-private key pair generation algorithm to generate the corresponding public-private key pair (PK_{u_1}, SK_{u_1}), and sends it to A.

2) Public key query: A selects a user identity u_2 and requests a public key from C. C runs the setup public key generation algorithm to generate the corresponding public key PK_{u_2} and sends it to A.

Challenge phase: A stops the query phase, selects a public key query to get the public key $PK_u^{'}$, and sends it to C. C chooses any $\beta \in \{0,1\}$, and calculates $\sigma = \left(PK_u^{'}, R_\beta, Address_\beta\right)$.

Re-query phase: A may continue to execute polynomially bounded queries, but may not query the public key $PK_u^{'}$ against the corresponding private key $SK_u^{'}$.

Guessing phase: A outputs a $\beta^{'}$, if $\beta^{'} = \beta$, then A wins the game $Game_1$. where the advantage of A winning the game $Game_1$ is $Adv_A(A) = \left|Pr\left[\beta = \beta^{'}\right] - \frac{1}{2}\right|$.

6.1.2 Unforgeability

Under adaptive message selection and marking attacks, a scheme is said to be unforgeable if the probability of an attacker A winning in game $Game_2$ at any probabilistic polynomial time is negligible.

The unforgeability of the scheme is described using the following game $Game_2$ between the challenger C and the attacker A. The attacking game consists of the following four phases.

Initial phase: C runs the system to build the algorithm, enters the safety factor λ, and sends the output system parameters pp to A.

Attack phase: A performs a polynomially bounded adaptive query similar to the one in the game $Game_2$.

Forgery phase: A forges the group key pair $\left(ds_u^{'}, ds_G^{'}, ID_u^{'}, ID_G\right)$ of user u and uses this key pair to perform group signature $Sign = \left(h_m, P_m, S_1^{'}, S_2^{'}\right)$ on the message m.

Verification phase: C runs the verification algorithm to verify that $u_1 = e\left(S_2^{'}, P_2\right)$ and $u_2 = e\left(S_1^{'}, P_s\right)$ are equal, If $u_1 = u_2$, then A wins in the game $Game_2$; otherwise, A fails in the game $Game_2$.

Since A has no access to the system master key ks, the probability that u_1 and u_2 are equal is extremely small and can be ignored, so this scheme can resist the unforgeable attack.

6.2 Traceability Analysis

6.2.1 Tracking of the transaction sender

Tracking Nodes TN applies to the group manager GM to open the group signature, and the GM verifies the identity of TN and sends the encrypted group membership information C_{ID} to TN using the master public key P_s, namely $Encrypt\left(\left(ID_1 \cdots ID_{TS} \cdots ID_n\right), Ps\right) \rightarrow C_{ID}$.

TN uses the master key ks agreed with the GM to decrypt and obtain the plaintext identity information of the group members, namely $Decrypt\left(C_{ID}, ks\right) \rightarrow \left(ID_1 \cdots ID_{TS} \cdots ID_n\right)$.

Calculates $h_1 = H_1\left(ID_{TS} \parallel hid, N\right)$, $P_1 = h_1 \cdot g_2 + P_s$; $h_2 = H_1\left(ID_G \parallel hid, N\right)$, $P_2 = h_2 \cdot g_2 + P_s$;

$$\begin{aligned} u_3 &= e\left(S_1, P_1\right) \\ &= e\left(\left(r_1^{-1}\right)\left(r_2 - h_{TX}\right) \cdot ds_{TS}, h_1 \cdot g_2 + P_s\right) \\ &= e\left(\left(r_1^{-1}\right)\left(r_2 - h_{TX}\right) \cdot d_1 \cdot g_1, \left(h_1 + ks\right) \cdot g_2\right) \\ &= e\left(g_1, g_2\right)^{\left(r_1^{-1}\right)\left(r_2 - h_{TX}\right) \cdot ks \cdot \left(h_1 + ks\right)^{-1}\left(h_1 + ks\right)} \\ &= e\left(g_1, g_2\right)^{\left(r_1^{-1}\right)\left(r_2 - h_{TX}\right) \cdot ks} \end{aligned}$$

$$\begin{aligned} u_1 &= e\left(S_2, P_2\right) \\ &= e\left(\left(r_1^{-1}\right)\left(r_2 - h_{TX}\right) \cdot ds_G, h_2 \cdot g_2 + P_s\right) \\ &= e\left(\left(r_1^{-1}\right)\left(r_2 - h_{TX}\right) \cdot d_2 \cdot d_1 \cdot g_1, \left(h_2 + ks\right) \cdot g_2\right) \\ &= e\left(g_1, g_2\right)^{\left(r_1^{-1}\right)\left(r_2 - h_{TX}\right) \cdot ks \cdot \left(h_2 + ks\right)^{-1} \cdot d_1 \cdot \left(h_2 + ks\right)} \\ &= e\left(g_1, g_2\right)^{\left(r_1^{-1}\right)\left(r_2 - h_{TX}\right) \cdot ks \cdot d_1} \end{aligned}$$

namely, $u_1 = u_3^{d_1} = u_3^{ks \cdot (h_1 + ks)^{-1}} = u_3^{ks \cdot \left(H_1\left(ID_{TS} \parallel hid, N\right) + ks\right)^{-1}}$.

Therefore, the identity of the sender of the transaction can be traced.

6.2.2 Tracking of the transaction receiver

TN according to the tracking key a and public key B of TR, calculates $Address^{'} = Hs\left(aR\right) G + B$,

The one-time random address during the transaction is

$$\begin{aligned} \text{Address} &= H_s(rA) \cdot G + B \\ &= H_s(r \cdot aG) \cdot G + B \\ &= H_s(aR) \cdot G + B \end{aligned}$$

namely, $Address = Address^{'}$.

Therefore, the identity of the receiver of the transaction can be traced.

6.3 Performance Analysis

6.3.1 Performance comparison

The privacy protection methods for regulated transactions proposed in this paper is compared with the existing blockchain transaction methods as shown in Table 1.

Literature [14] achieves anonymity and regulability of the identities of both parties to a transaction by improving on group signatures, but it does not provide privacy protection for the transaction amount. Literature [26] uses Paillier homomorphic encryption to hide the transaction amount, but it creates an additional account for receiving the transaction amount and the encryption and verification phases require i cycles, which is inefficient and fails to satisfy the need for anonymity of the transaction identities. Literature [24] improves CryptoNote, the underlying technology of Monero, which can realize the supervision of the transaction process while protecting the privacy of the user's identity and the privacy of the transaction amount, but the scheme fails to eliminate the impact of the number of output addresses with consistent transaction amounts on the efficiency of the transaction. In this paper, group signature, one-time random address, Paillier homomorphic encryption and Bulletproof zero-knowledge proof and other technologies are used, not only to realize the privacy protection of the identity information of the two parties of the transaction and the transaction amount, but also in order to avoid the abuse of the right, the regulatory node has the audit node and the tracking node together, to realize the regulation of the transaction process, and the scheme's computational overhead is relatively small.

6.3.2 Calculation Overhead

Let T_E denotes the computation time overhead of exponential operation, T_B denotes the computation time overhead of bilinear pair, and T_H denotes the computation time overhead of hash operation.In order to reduce the error in the testing process, 100 experiments are conducted on the system to take the average value as the basic cryptography operation time, and the operation time is shown in Table 2, it can be seen $T_H < T_E < T_B$, and the bilinear pair operation time is much higher than other cryptographic operation times.

Table 3 shows the computational overhead of the scheme in this paper, and it can be seen that the computational overhead of this paper is mainly concentrated in the transaction generation and transaction verification stages, in which the computational overhead of the transaction verification stage is related to the range of transaction amount. Although the computational overheads of the two stages of transaction generation and transaction verification are large, the computational overheads are mostly exponential operations, and the bilinear pair operations are relatively small, so the computational overheads of the scheme in this paper are relatively reasonable. The above results are based on theoretical analysis and not actual implementation results, which will be tested by experiments in the next section.

6.4 Experimental Verification

The experiments were done on a virtual machine VMware with 4GB RAM, Intel(R) Core(TM) i5-8250U CPU @ 1.60GHz, and operating system Ubuntu 18.04 64-bit. The scheme is implemented in go language, version 1.19.4, us-

Table 1: Performance comparison between this paper and other schemes

Schemes	Transaction identity privacy	Transaction amount privacy	Identity legitimacy verification	Amount legitimacy verification	Regulability
Literature [14]	✓	×	×	×	✓
Literature [26]	×	✓	×	✓	×
Literature [24]	✓	✓	×	✓	✓
This paper	✓	✓	✓	✓	✓

Table 2: Basic cryptographic operation time

Operation	T_E	T_B	T_H
Time (ms)	1.98 ms	10.08 ms	0.06ms

Table 3: The computational overhead of the scheme in this paper

Procedure	Computational overhead
User registration	$2T_H$
One-time random address generation	T_H
Transaction generation	$15T_E + T_B + 5T_H$
Transaction verification	$(11 + 2n)\,T_E + 4T_B + 2T_H$

ing GoLand as the integrated development tool, in which the bilinear pair mapping uses the pbc library, the Hash algorithm uses the cryptographic hash algorithm SM3, secp256k1 is selected as the elliptic curve algorithm, and the NTL library is called to implement the Paillier algorithm. The key steps of the system are experimentally implemented in simulation, as well as tested and analyzed for the generation verification time of the Paillier encryption and decryption algorithm and Bulletproofs range proofs.

In the performance analysis phase, a theoretical analysis reveals a connection between the verification time of Bulletproofs range proof and the range of the amount proved. The secp256k1 elliptic curve algorithm is chosen, and the range of transaction amounts is represented by an exponent with a base of 2. The exponents are taken to be 4, 8, 16, 32, 64, and 128 for experiments to test the effect of the size of the range of transaction amounts on the total time overhead of performing a range proof. With an exponent of 128, the range of transaction amounts has reached a large value, which is fully sufficient for practical applications.The appropriate range of transaction amounts can be selected based on specific application requirements, which can result in relatively small transaction validation times. As shown in Figure 3, it can be seen that as the range of transaction amounts increases, the time consumed to execute a range proof increases, which confirms the conclusion obtained in the performance analysis phase.

The stability of the Paillier encryption algorithm is

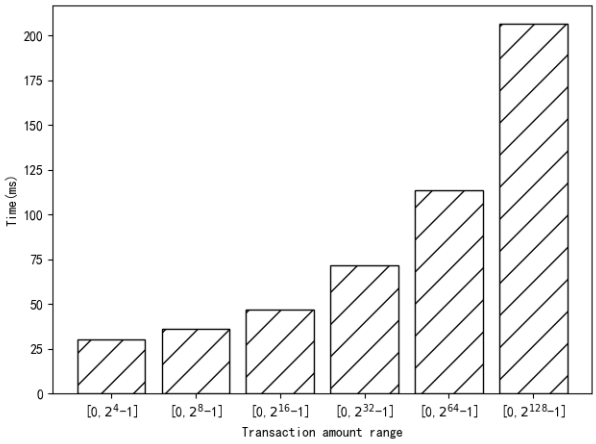


Figure 3: Time overhead for performing a range proof for different transaction amount ranges

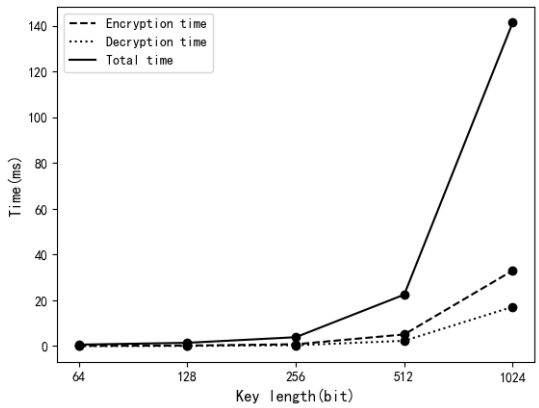


Figure 4: Paillier encryption and decryption time and total time overhead for different key lengths

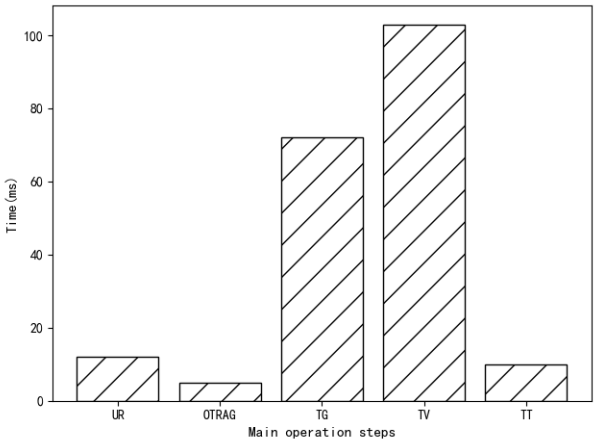


Figure 5: Time overhead of the main operation steps of the scheme in the same transaction amount range

based on the compound residual class difficulty problem, in the case that the key length is large enough, the algorithm can be secured from easy attack cracking according to cryptanalysis theory. Since the time overhead of the Paillier encryption algorithm is mainly in the encryption phase and decryption phase, the settings of the system key lengths are 64-bit, 128-bit, 256-bit, 512-bit, and 1024-bit to test the time overhead of Paillier encryption and decryption and the total time overhead of performing one Paillier encryption and decryption, respectively. To better reflect the overhead of Paillier encryption and decryption time and total time with different bit key lengths, 100 experiments are conducted on the algorithm, and the data obtained from the statistics are averaged, and the experimental results are shown in Figure 4. It can be seen that the time overhead of encryption and decryption increases with the increase of key length, when the key length reaches 1024-bit, the total time overhead of executing the Paillier encryption and decryption algorithm reaches 141.47ms, and the key length of 1024-bit has reached a high order of magnitude in the application, but in this paper the choice of 512-bit is sufficient for the privacy and regulatory aspects of the scheme.

Simulated implementation in a virtual machine of the five main operational steps of the scheme: user registration(UR), one-time random address generation(OTRAG), transaction generation(TG), transaction verification(TV), and transaction tracking(TT). Where the key length is 1024-bit with high security factor and the range of transaction amount is $\left[0, 2^{32}-1\right]$, which can meet the security and practical application requirements of the scheme. From Figure 5, it can be seen that the time overhead of the scheme in the user registration, one-time random address generation, and transaction tracking phases is relatively small, but the time overhead in the transaction generation and transaction verification phases is relatively large, and how to improve the efficiency of the transaction generation and transaction verification phases with a certain range of transaction amounts is the focus of our next research work.

7 Conclusion

In this paper, we propose a regulatable privacy protection scheme for transactions based on consortium blockchain by combining technologies such as group signature, one-time random address, Paillier homomorphic encryption, and Bulletproofs range proofs to achieve both privacy and regulability of blockchain transaction data. The scheme not only protects the anonymity of the identities of both parties of the transaction, but also can verify the legitimacy of the transaction amount while protecting the privacy of the transaction amount. By introducing audit nodes and tracking nodes to play the role of regulatory nodes, not only can user nodes effectively avoid illegal transactions using random or negative amounts, but also can expose the evil user nodes. Theoretical analysis shows that the scheme in this paper realizes a high degree of anonymity, but not absolute anonymity, and due to the intervention of the regulatory nodes, the information of the two parties conducting illegal transactions can be traced. Experimental tests show that the computational overhead of the scheme has some advantages, but the time overhead of the scheme in the transaction generation and transaction validation phases is large, and the time overhead of the validation phase increases with the increase of the transaction amount range, how to improve the efficiency of the scheme in the transaction generation and validation phases, as well as to reduce the impact of the transaction amount range on the efficiency of validation is an important research topic for us in the next step.

Acknowledgments

This study was supported by the National Natural Science Foundation of China under grants 61862040 and 62162039. The authors thank the anonymous reviewers for their helpful comments and suggestions.

References

[1] T. Alamr, "A survey on the use of blockchain for the internet of things," *International Journal of Electronics and Information Engineering*, vol. 13, no. 3, pp. 119–130, 2021.

[2] O. Ali, A. Jaradat, A. Kulakli, and A. Abuhalimeh, "A comparative study: Blockchain technology utilization benefits, challenges and functionalities," *IEEE Access*, vol. PP, no. 99, pp. 1–1, 2021.

[3] A. K. K. Ansah, F. L. Zhang, and D. Adu-Gyamfi, "Ringcoin: an accountable mix for achieving bitcoin anonymity," *International Journal of Network Security*, vol. 23, no. 3, pp. 505–515, 2020.

[4] S. J. Bai, G. Yang, C. M. Rong, G. X. Liu, and H. Dai, "Qhse: An efficient privacy-preserving scheme for blockchain-based transactions," vol. 112, pp. 930–944, 2020.

[5] G. Betarte, M. Cristiá, C. Luna, A. Silveira, and D. Zanarini, "Towards a formally verified implementation of the mimblewimble cryptocurrency protocol," in *Applied Cryptography and Network Security Workshops: ACNS 2020 Satellite Workshops, AIBlock, AIHWS, AIoTS, Cloud S&P, SCI, SecMT, and SiMLA, Rome, Italy, October 19–22, 2020, Proceedings 18*, pp. 3–23. Springer, 2020.

[6] Y. R. Chen, H. Chen, Y. Zhang, M. Han, M. Siddula, and Z.P. Cai, "A survey on blockchain systems: Attacks, defenses, and privacy preservation," *High-Confidence Computing*, vol. 2, no. 2, p. 100048, 2022.

[7] J. K. Duan, L. Z. Gu, and S. H. Zheng, "Arct: An efficient aggregating ring confidential transaction protocol in blockchain," *IEEE Access*, vol. 8, pp. 198118–198130, 2020.

[8] T. Feng, P. Yang, C. Y. Liu, J. L. Fang, and R. Ma, "Blockchain data privacy protection and sharing scheme based on zero-knowledge proof," *Wireless Communications and Mobile Computing*, vol. 2022, pp. 1–11, 2022.

[9] B. Gong, C. Cui, M. S. Hu, C. Guo, X. C. Li, and Y. H. Ren, "Anonymous traceability protocol based on group signature for blockchain," *Future Generation Computer Systems*, vol. 127, pp. 160–167, 2022.

[10] Z. S. Guan, Z. G. Wan, Y. Yang., Y. Zhou, and B. T. Huang, "Blockmaze: An efficient privacy-preserving account-model blockchain based on zk-snarks," *IEEE Transactions on Dependable and Secure Computing*, vol. 19, no. 3, pp. 1446–1463, 2020.

[11] L. Y. Guo, H. Xie, and Y. Li, "Data encryption based blockchain and privacy preserving mechanisms towards big data," *Journal of Visual Communication and Image Representation*, vol. 70, p. 102741, 2020.

[12] Y. X. Jia, S. F. Sun, Y. C. Zhang, Q. Z. Zhang, N. Ding, Z. Q. Liu, J. K. Liu, and D. W. Gu, "Pbt: A new privacy-preserving payment protocol for blockchain transactions," *IEEE Transactions on Dependable and Secure Computing*, vol. 19, no. 1, pp. 647–662, 2020.

[13] G. L. Li, D. B. He, B. Guo, and S. Lu, "Blockchain privacy protection algorithms based on zero knowledge proof," *J. Huazhong Univ. Sci. Technol.(Natural Science Edition)*, vol. 48, no. 7, 2020.

[14] P. L. Li and H. X. Xu, "Blockchain user anonymity and traceability technology," *Journal of Electronics & Information Technology*, vol. 42, no. 5, pp. 1061–1067, 2020.

[15] T. Li, Z. J. Wang, Y. L. Chen, C. M. Li, Y. L. Jia, and Y. X. Yang, "Is semi-selfish mining available without being detected?," *International Journal of Intelligent Systems*, vol. 37, no. 12, pp. 10576–10597, 2022.

[16] X. F. Li, Y. R. Mei, J. Gong, F. Xiang, and Z. X. Sun, "A blockchain privacy protection scheme based on ring signature," *IEEE Access*, vol. 8, pp. 76765–76772, 2020.

[17] X. Q. Li, P. Jiang, T. Chen, X. P. Luo, and Q. Y. Wen, "A survey on the security of blockchain systems," *Future generation computer systems*, vol. 107, pp. 841–853, 2020.

[18] Y. N. Li, G. M. Yang, W. Susilo, Y. Yu, M. H. Au, and D. X. Liu, "Traceable monero: Anonymous cryptocurrency with enhanced accountability," *IEEE Transactions on Dependable and Secure Computing*, vol. PP, no. 99, pp. 1–1, 2020.

[19] Y. Liu, M. X. He, and F. Y. Pu, "Anonymous transaction of digital currency based on blockchain.," *International Journal of Network Security*, vol. 22, no. 3, pp. 442–448, 2020.

[20] S. L. Ma, Y. Deng, D. B. He, J. Zhang, and X. Xie, "An efficient nizk scheme for privacy-preserving transactions over account-model blockchain," *IEEE Transactions on Dependable and Secure Computing*, vol. 18, no. 2, pp. 641–651, 2020.

[21] D. D. F. Maesa and P. Mori, "Blockchain 3.0 applications survey," *Journal of Parallel and Distributed Computing*, vol. 138, pp. 99–114, 2020.

[22] M. R. Nosouhi, Y. S., K.Sood, M. Grobler, R. Jurdak, A. Dorri, and S. G. Shen, "Ucoin: An efficient privacy preserving scheme for cryptocurrencies," *IEEE Transactions on Dependable and Secure Computing*, 2021.

[23] A. Rahimi and M. A. Maddah-Ali, "Multi-party proof generation in qap-based zk-snarks," *IEEE Journal on Selected Areas in Information Theory*, vol. 2, no. 3, pp. 931–941, 2021.

[24] H. B. Tian, H. Z. Lin, P. R. Luo, and Y. X. Su, "Scheme for being able to regulate a digital currency with user privacy protection," *Journal of Xidian University*, vol. 47, no. 5, p. 8, 2020.

[25] D. Wang, J. D. Zhao, and Y. J. Wang, "A survey on privacy protection of blockchain: the technology and application," *IEEE Access*, vol. 8, pp. 108766–108781, 2020.

[26] Q. Wang, B. Qin, J. K. Hu, and F. Xiao, "Preserving transaction privacy in bitcoin," *Future Generation Computer Systems*, vol. 107, pp. 793–804, 2020.

[27] H. M. Xie, S. F. Fei, Z. Yan, and Y. Xiao, "Sofitmix: A secure offchain-supported bitcoin-compatible mixing protocol," *IEEE Transactions on Dependable and Secure Computing*, 2022.

[28] G. Y. Yang, Y. L. Wang, Z. J. Wang, Y. L. Tian, X. M. Yu, and S. Z. Li, "Ipbsm: an optimal bribery selfish mining in the presence of intelligent and pure attackers," *International Journal of Intelligent Systems*, vol. 35, no. 11, pp. 1735–1748, 2020.

[29] Y. T. Yang, J. L. Cai, X. W. Zhang, and Z. Yuan, "Privacy preserving scheme in block chain with provably secure based on sm9 algorithm," *Journal of Software*, vol. 30, no. 6, pp. 1692–1704, 2019.

[30] S. Q. Zeng, R. Huo, T. Huang, J. Liu, S. Wang, and W. Feng, "Survey of blockchain: principle, progress and application," *Journal on Communications*, vol. 41, no. 1, pp. 134–151, 2020.

[31] A. Zhang and X. Y. Bai, "Survey of research and practices on blockchain privacy protection," *Journal of Software*, vol. 31, no. 5, pp. 1406–1434, 2020.

[32] J. H. Zhang, W. L. Bai, and Z. T. Jiang, "On the security of a practical constant-size ring signature scheme.," *International Journal of Network Security*, vol. 22, no. 3, pp. 392–396, 2020.

Biography

Pengshou Xie was born in Jan. 1972. He is a professor and a supervisor of master student at Lanzhou University of Technology. His major research field is Privacy Protection, Security on Internet of Vehicles, Security on Industrial Internet. E-mail: xiepsh_lut@163.com.

Wanjun Shao was born in Jan. 1998. She is a master student at Lanzhou University of Technology. Her major research field is network and information security. E-mail: 2443404684@qq.com.

Tao Feng was born in Dec. 1970. He is a professor and a supervisor of Doctoral student at Lanzhou University of Technology. His major research field is modern cryptography theory, network and information security technology. E-mail: fengt@lut.cn.

Ye Lu was born in May. 1986. He is currently a lecturer in Lanzhou University of Technology. His research interests include cyber security and blockchain. E-mail: luye@lut.edu.cn.

Yinchang Pan was born in Mar. 1993. He is a master student at Lanzhou University of Technology. His major research field is network and information security. E-mail:1713974116@qq.com.

Xinyu Fan was born in June. 2000. She is a master student at Lanzhou University of Technology. Her major research field is network and information security. E-mail: 438573214@qq.com.

MalYolo5: A Malicious Code Visualized Detection Model Based on Yolov5 with CA Attention Mechanism

Zipeng He, Yuntao Zhao, and Yongxin Feng

(Corresponding author: Yuntao Zhao)

School of Information Science and Engineering, Shenyang Ligong University
Shenyang 110159, China
Email: zhaoyuntao2014@163.com

(Received June 11, 2023; Revised and Accepted Jan. 5, 2024; First Online Aug. 16, 2024)

Abstract

The growing popularity of the internet and the expanding cyberspace have led to increased attention towards security by researchers. To address the issue of detecting and classifying malicious code, we propose visualizing the code as color images and converting their features into image visualization features. This approach facilitates the extraction of malicious code features. An improved lightweight YOLOv5s model was proposed to optimize the amount of calculation. The novel model includes both the SE and CA attention mechanisms. The experimental comparison shows that the MAP of the YOLOv5s model with CA attention is better than that of the traditional YOLOv5s model with SE attention added. The MAP of the latter's experimental results is also better than the YOLOv5s model without attention added. Finally, we have determined that the YOLOv5s model with CA attention is the improved YOLOv5 model for malicious code classification and detection.

Keywords: Attention Mechanism; Malicious Code Visualization; YOLOv5

1 Introduction

In today's era of data explosion, the emergence of the Internet has significantly impacted people's lifestyles and daily routines, resulting in a steady increase in the number of Internet users each year. According to a report published by the China Internet Network Information Center (CNNIC) in August 2022, the total number of Internet users in China as of June 2022 was 1.051 billion. However, the widespread use of the Internet has also led to numerous security concerns across various domains. Network hackers often use network attacks to steal people's information for extortion, damage equipment, and disrupt production. One of the most common methods of network attacks is through malicious code. The proliferation of malicious codes has caused significant harm to various countries and societies around the world. The national network security monitoring department's statistics show that in 2018, over 100 million samples of computer malicious programs were captured. These samples belonged to more than 500,000 families of malicious programs, and on average, they spread 5 million times per day on computers across the country. It is important to note that this information is objective and not subjective. Therefore, the study of quickly and accurately identifying malicious code is of great research significance in the field of network security.

This paper proposes a method for visually detecting malicious code using Yolov5 with an added coordinate attention mechanism to detect and classify colour malicious code images. Additionally, a YOLOv5-CA target detection model is established, which outperforms the Yovov5 model.

2 Domestic And International Status Quo

According to the national cooperation network analysis of the collected WOS literature data [1], in the research of malicious code detection, the United States and China are in the leading position with 130 and 105 papers respectively, accounting for about 47% of the total number of papers; from In terms of time distribution, the research of malicious code detection started earlier in the United States, Germany and Japan, but the research results of the United States are far ahead of Germany and Japan. Since 2010, more and more countries have also started to pay attention to malicious code detection research. Although China started later than other countries, recent research suggests that it has made significant progress in the field of malicious code detection.

Malicious code detection can be classified into three

types: static detection, dynamic detection, and hybrid detection, depending on whether the executable program is running. Static detection does not require the execution of malicious code; instead, it focuses on effectively and accurately extracting the static features of malicious code. The objective of obtaining the static features of malicious code can be achieved through decompiling, disassembling, and analyzing the file structure. This involves extracting instructions, byte sequences, and file header information from the malicious code file. The static detection method analyses the operation code sequence information, function call control flow graph, and Windows API call sequence of malicious codes through their assembly files. This analysis extracts effective features that are used for detecting and classifying malicious codes. This method offers advantages such as low energy consumption, fast speed, low risk, high coverage of malicious samples, and quick capture of syntax and semantic information for comprehensive analysis. Malicious code interference after obfuscation technology processing. Dynamic detection technology involves checking and analyzing malware after running it. To ensure system and network safety, a virtual environment must be built before running the malware due to its malicious nature. To fully demonstrate the behaviour and hidden functions of the malicious code, it is necessary to grant sufficient permissions to the code before running it. Additionally, it is important to ensure that the code can run normally with different attack content. When running malicious code in a virtual machine, it is important to use monitoring tools to track its process, listen to traffic packets, check the system registry, log files, and more. This allows for the analysis and extraction of effective features after the malicious code has run. Dynamic detection has the advantage of identifying new types of malicious code and effectively addressing false positives and negatives that may occur in static detection.

Hybrid analysis is often used to analyse packed malicious codes, combining static and dynamic methods. The packed malicious code is first unpacked through dynamic analysis, and then static analysis is used to extract the features of the unpacked code. This allows for the identification of the attack behaviour of the malicious code. Malicious codes can generate numerous variants, leading to the failure of signature-based detection methods. However, visualizing malicious codes as images does not significantly alter the image texture and structural features. This method can effectively combat malicious code obfuscation. Nataraj et al. visualised the binary data of the malicious code .text block as a grayscale image [2]. They used the GIST algorithm to extract image features and the K-Nearest Neighbor (KNN) algorithm to classify the data. This research marks the beginning of the development of a malicious code detection method based on visualisation. At the 2015 Black Hat Conference, Davis et al. proposed a method for vectorizing disassembled malicious code by encoding the source of the disassembled hexadecimal data seat. They multiplied each binary bit by 255, corresponding to the pixel gray value of a grayscale image (0 or 255). Jiang Yongkang and others further explored the selection of parameters, such as encoding length and amount. Wang Runzheng et al. disassembled the malicious code and converted the data of each block into an RGB color image. Convolutional neural networks are effective in extracting features and classifying images. Therefore, using them for malicious code image analysis has great research potential. Currently, the ever-increasing production of variants, modifications, and updates of malicious codes requires up-to-date professional knowledge to combat these emerging threats. The original method of feature extraction may no longer be suitable for new malware families, resulting in the need for feature engineering. This process can be time-consuming and inefficient. The research aims to reduce the burden of manual feature engineering and extract valuable information directly from raw data. The goal is to enhance the accuracy and efficiency of malware detection by enabling the model to autonomously learn relevant features. Techniques are being developed to facilitate self-learning capabilities within the model, enabling it to effectively discern and classify malicious software.

3 Malicious Code

3.1 Types of Malicious Code

3.1.1 Computer Virus

A computer virus, also known as computer malware, is a set of instructions or program code injected into a computer program with the intention of disrupting computer operations or damaging data. It interferes with the regular functioning of a computer and has the ability to self-replicate. Computer viruses typically possess certain characteristics, including infectivity, concealment, latency, and payload. Malware has the capability to spread from one computer to another, often disguising its presence and remaining dormant until triggered to execute malicious actions. The life cycle of malware is divided into seven stages: development, infection, latency, attack, discovery, digestion, and demise.

3.1.2 Trojan Horse

A computer Trojan horse virus is a type of malicious code that is embedded in normal programs. It acts as a secret backdoor program, equipped with specific functions such as file destruction and deletion, password theft, keystroke logging and denial of service (DoS) attacks. Trojan horses are designed to appear harmless or even attractive to unsuspecting users. Trojan horses are often hidden within gaming or graphic software, disguising their true nature and intentions. Once these seemingly safe programs are run, they can perform illegal actions such as deleting files or formatting the hard disk. A complete Trojan horse program typically consists of two parts: the server side

and the controller side. The term 'hit by a Trojan horse' refers to a server-side program that has been installed with a Trojan horse. If a server-side program is installed on your computer, someone with the corresponding controller can control your computer through the network.

3.1.3 Worm

A worm [3] is a self-replicating piece of code that spreads through a network, usually without human intervention. Email attachment propagation is the primary delivery method for network worms. Worm authors send emails to users. When users click on email attachments, networnataraj2011malwarek worms infect the computer. Once the worm has invaded and taken control of a computer, it will use that computer as a host and search for other computers to infect. If the invasion is successful, it will continue to use the computer as a host and then invade other computers. The number of worm infections therefore increases in a similar recursive manner.

3.1.4 Backdoor Virus

The characteristic of the backdoor virus is that it spreads through the network, opening the back door to the system and bringing security risks to the user's computer. In early 2004, IRC backdoor viruses began to appear on a large scale on the global network. On one hand, there is a risk of leaking local information. On the other hand, viruses can infiltrate the LAN, block the network, disrupt normal work, and cause losses. As the virus source code is public, anyone with access to it can compile and generate a new virus with minimal modifications. Furthermore, various shells have led to the emergence of numerous IRC backdoor virus variants.

3.2 Malicious Code Visualization

3.2.1 Malicious Code is Visualized As a Gray Image

Nataraj et al. presented a visualization of the binary data from the malicious code's .text block as a grayscale image. Nataraj utilises malicious code binary as the encoding source. The value interval corresponding to each 8-bit binary conversion is [0,255], which corresponds to the pixel interval and can be regarded as the value of a pixel. The number and length of the malicious code contained in the file vary, including different types of attacks, resulting in different lengths of converted binary sequences and, therefore, different sizes of converted visual images. By fixing the width of the image, the malicious code is transformed into a long grey image. Please refer to the Figure 1 below for specific steps. Nataraj vectorization shares similarities with the B2M algorithm. This grayscale visualization method has been extensively used in the detection of malicious code [4–7]. Han et al. [8] improved upon Nataraj vectorization by adding an entropy map. They used the features of the entropy map to better

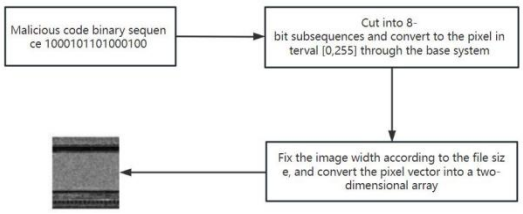


Figure 1: Nataraj vectorization specific steps

assess the similarity of malicious code images. Additionally, they improved the image texture feature extraction method and similarity measurement strategy.

In addition to the Nataraj vectorization method, David [9] proposed vectorizing disassembled malicious code by using the hexadecimal encoding source. Each hexadecimal is converted into 4-bit binary resulting in 64-bit binary for 4 hexadecimals. Each bit of binary is then multiplied by 255 resulting in a pixel gray value of either 0 or 255. The method presented in this study represents malicious code as a grayscale image with pixel values of either 0 or 255. Each row of vectors in the image corresponds to a machine code. Jiang Yongkang et al. [10]expanded on this approach by exploring the selection of parameters, such as coding length and amount, and proposing specific deep learning models.

3.2.2 Malicious Code Visualized As a Color Image

As grayscale images have only one channel, they can contain less information. This means that malicious code attack information may not be fully reflected in the image, resulting in inconspicuous features of the visualized grayscale image. Therefore, the characteristics of the malicious code may not be well reflected. Compared to grayscale images of malicious code, visualizing them as colour images retains the main features of grayscale while also emphasising repeated data fragments in binary files [11]. This means that colour images of the same malicious family have similar textures, colours, and structural features.

Wang Bo et al [12] divided the binary sequence of malicious code into three RGB channels. Each 24-bit group forms the RGB value of a pixel. However, not all malicious code binary digits are multiples of 24. Therefore, it is necessary to determine if the number of digits in the binary sequence is an integer multiple of 24 bits. If not, the digits are supplemented with 1 to visualize the malicious code as a color image. Refer to Figure 2 for the specific steps.

Figure 3 displays the image data used in this experiment. The colour images produced by the malicious code exhibit distinct texture features. The target detection model classifies the images based on these features. This classification enables the identification of malicious code.

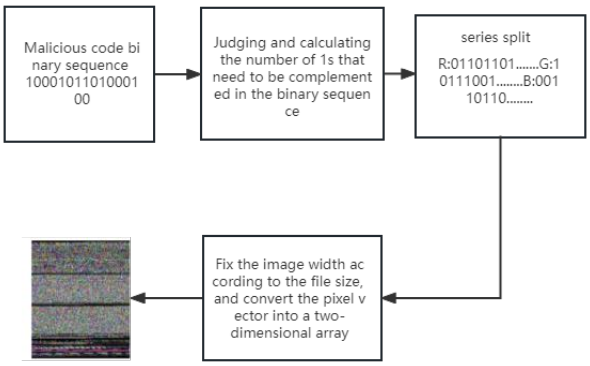


Figure 2: Nataraj vectorization specific steps

Figure 3: .Malicious code color image dataset example

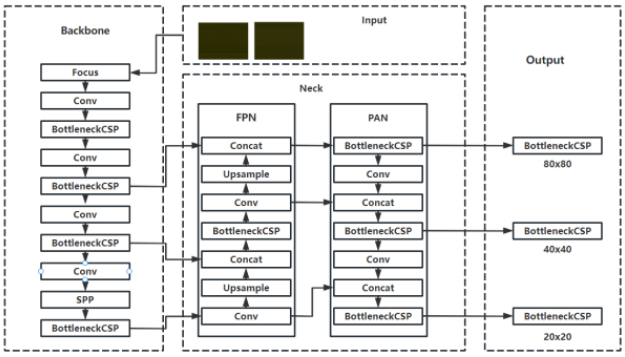


Figure 4: YOLOv5s network model structure diagram

4 Yolo5s Model and Attention Mechanism

4.1 Yolo5s Model

In 2016, Redmon proposed YOLO, which stands for 'You Only Look Once'. The latest model series is YOLOV5, which includes four models: The latest model series is YOLOV5, which includes four models: The latest model series is YOLOV5, which includes four models: The YOLOv5 models, including YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x, are arranged from shallow to deep [13]. As the depth of the network model increases, the number of model parameters progressively grows, which poses challenges for deploying experiments that cater to diverse requirements. Therefore, this paper chooses the lightweight YOLOv5s model. YOLOv5s consists of four parts: the input terminal (Input), the backbone network (Backbone), the neck (Neck), and the output terminal (Output), as shown in Figure 4.

The Yolov5s backbone, which extracts features, is composed of Focus, Conv, BottleneckCSP modules, and an SPP structure. Prior to entering the backbone, the picture is sliced and a specific operation is performed to extract values for alternate pixels within an image, resembling adjacent downsampling. This process results in four complementary pictures where the information remains largely intact. Consequently, the image's width (W) and height (H) information is consolidated within the channel space, effectively expanding the input channel by a factor of four. This means that the stitched picture has 12 channels compared to the original RGB three-channel mode. Finally, the new image undergoes a convolution operation, resulting in a double downsampling feature map without any loss of information. Its role is to crop, stack, and downsample the image. The BottleneckCSP module performs convolution operations to extract image feature information. To address the issue of non-uniform input image size, the backbone network includes a spatial pyramid pooling layer (SPP) [14].

The neck of Yolov5s consists of a bottom-up feature pyramid (Feature Pyramid Networks, FPN) and a top-down path aggregation network structure (Path Aggregation Network, PAN). Through further feature extraction using FPN and PAN, the image is fused with multi-scale features. This ensures that the feature map contains both the semantic and feature information of the target image, allowing for accurate recognition of images of different sizes.

4.2 Attention Mechanism

The attention mechanism was initially used for machine translation tasks and has since been widely adopted in various fields of deep learning, including image segmentation, speech processing, computer vision, and natural language processing. Its effectiveness is evident in all of these areas. In the field of cognitive science, humans exhibit a selective focus on specific information while disregarding other information due to the limitations of information processing capacity. This phenomenon can be understood as a bottleneck in cognitive processing. Similarly, neural networks, when faced with a large amount of information, quickly direct their attention towards crucial information for efficient processing. This is known as the attention mechanism.

The attention mechanism lacks a precise mathematical definition. It includes various techniques, such as traditional local image feature extraction and sliding window methods. In the context of neural networks, the attention mechanism often involves an auxiliary neural network that can effectively highlight specific portions of the input or assign distinct weights to different elements of the input.

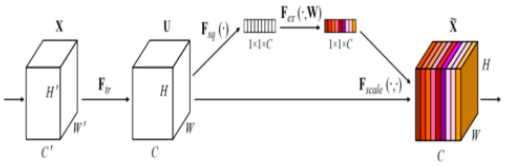


Figure 5: SENet attention structure diagram

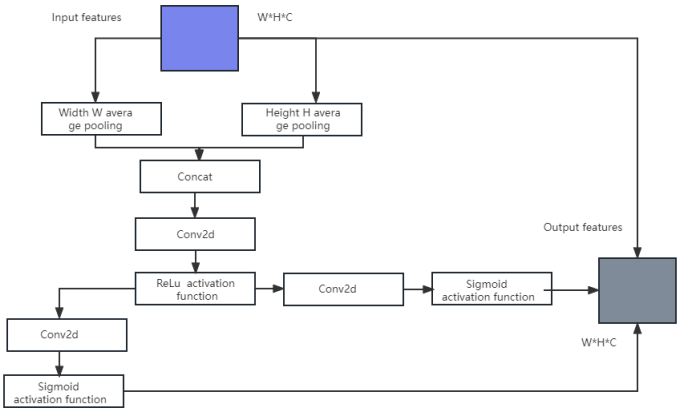


Figure 6: Coordinate attention mechanism structure diagram

Currently, attention can be classified into four types: channel attention [15], spatial attention [16], hybrid attention, and coordinate attention. These types are applied to the channel domain, spatial domain, mixed domain (a combination of channel and spatial), and coordinate domain, respectively. They enable the filtering of useful information from a large amount of data.

4.2.1 SENet Attention Mechanism

SENet, which was introduced in 2017, is a well-known implementation of the channel attention mechanism. It achieved the top position in the previous ImageNet competition. The diagram below illustrates its implementation. SENet places emphasis on the weights assigned to each channel of the input feature layer. By incorporating SENet, the network can selectively prioritize channels that require more attention, allowing for enhanced focus on essential features. The neural network model can acquire the significance of each feature channel through training by employing the channel domain attention mechanism. This allows the model to allocate greater attention to channels with higher weights while suppressing channels with lower weights. Consequently, the model becomes more adept at focusing on the most informative and relevant feature channels, leading to improved performance in various tasks.

The implementation method involves performing global average pooling on the input feature layer, followed by applying two fully connected layers. The first layer has a smaller number of neurons, while the second layer matches the size of the input feature layer. After completing these two layers, the Sigmoid activation function is applied to confine the resulting values between 0 and 1, yielding the weights assigned to each channel of the input feature layer. Using the obtained weights, we can multiply them with the original input feature layer to adjust the importance of each channel.The SENet schematic is shown in Figure 5.

4.2.2 CA Attention Mechanism

SE attention primarily focuses on encoding inter-channel information while neglecting the significance of location information. However, location information plays a critical role in capturing the structural aspects of objects in vision-related tasks. Therefore, we present a novel attention mechanism that integrates location information into channel attention, enabling mobile networks to focus on large areas while avoiding a large computational overhead. To elaborate, our approach uses two 1D global pooling operations to aggregate input features vertically and horizontally. This results in two directional feature maps that encapsulate direction-specific information. These feature maps are then separately encoded into attention maps, capturing long-range correlations along their respective spatial directions. This ensures that the attention maps preserve location information. Subsequently, the input feature map is multiplied element-wise with the two attention maps to emphasize relevant representations. This attention mechanism is referred to as 'coordinate attention' because it can differentiate spatial directions (coordinates) and generate corresponding attention maps.

The advantages of coordinate attention, as described, are incorporated. Firstly, the model captures both inter-channel information and orientation-aware, position-sensitive information. This enables more accurate localization and identification of objects of interest, enhancing its object detection capabilities. Additionally, our approach offers flexibility and efficiency, making it easily integrable into fundamental components of mobile networks. It can be easily integrated into established building blocks, such as the inverted residual block from MobileNetV2 or the hourglass block from MobileNeX, to enhance feature extraction by amplifying meaningful representations. Furthermore, when used as a pre-trained model, our coordinated attention significantly improves performance in downstream tasks involving mobile networks, especially those that require dense predictions, such as semantic segmentation.The Coordinate attention mechanism structure is shown in Figure 6.

5 Experimental Results and Analysis

5.1 Dataset Introduction

The paper's dataset converts binary malicious code into hexadecimal and then groups every two hexadecimals into an eight-bit binary number, representing a value range of 0-255. The dataset then generates gray based on the value of each pixel. The dataset includes 1547 pictures, divided into a training set and a verification set in a 9:1 ratio. The training set contains 1392 images, and the verification set contains 155 images. After dividing the training and verification sets, use the Annotation data labeling tool to label the data. The format of the annotation should be saved in XML format, and the label file name should match the picture name.

5.2 Improvement

The paper's innovations primarily focus on improving the network structure module of YOLOv5s and incorporating various attention mechanisms. To achieve this, the authors simplified the BottleneckCSP of the Backbone network and Neck part, replacing conv with CBL (Conv+BN+SiLu) while retaining the residual part. Retaining the residual structure can increase the gradient of backpropagation between layers, avoiding the disappearance of the gradient caused by deepening. This allows for finer-grained features to be extracted without worrying about network degradation. Our streamlined module, called the C3 module, simplifies the network structure of the model and reduces model parameters. This has the effect of reducing the amount of model calculation and model inference time. Figure 7shows the BottleneckCSP of the Backbone part, while Figure 8 displays the simplified C3 module. Figure 9 illustrates the BottleneckCSP of the Neck part, and Figure 10 shows the simplified module.

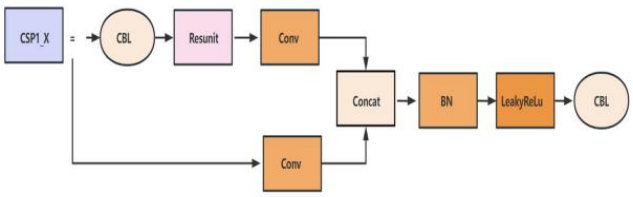


Figure 7: Network Structure of Bottleneck CSP in Backbone Part

The C3 module of the Backboone backbone network was modified by adding SEnet attention, resulting in the C3SE module. The YOLOV5S network structure was then updated to include the C3SE module with SEnet attention, resulting in the YOLOv5-Se network structure shown in Figure 11. To compare the effects of different attentions, a comparative experiment was conducted by

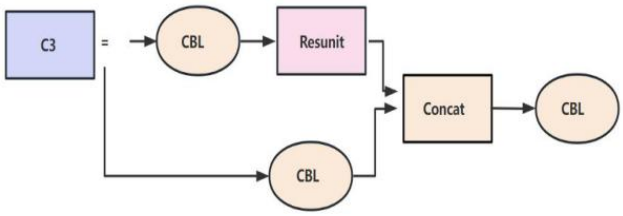


Figure 8: Backbone part C3 module network structure

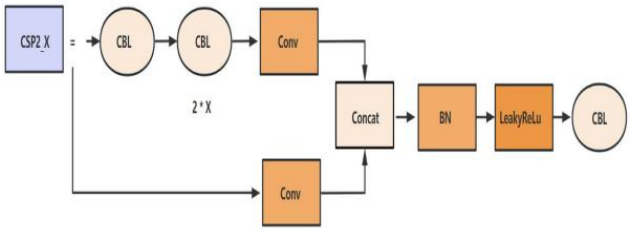


Figure 9: Neck Part BottleneckCSP Network Structure

adding coordinate attention to the C3 module. The resulting module is called C3CA. Additionally, the network structure of YOLOV5S was modified by adding CA attention, resulting in YOLOv5-CA, as shown in Figure 12.

5.3 Analysis of Results

5.3.1 Experiment Settings

The experiment's hardware settings and development environment were implemented on a cloud server, with detailed configuration provided in Table 1. Experimental parameters include input size, initial round, freezing round, unfreezing round, training freeze status, batch size, initial and minimum learning rates, optimizer, and momentum parameters. The input size indicates the size of the input image. The training process is divided into two stages: the freezing stage and the unfreezing stage. In the freezing stage, the model's backbone is frozen, and the feature extraction network remains unchanged. This stage requires less video memory. For Fine-tuning the network, the model's backbone is not frozen during the unfreezing stage. This means that the feature extraction network will change and occupy a large amount of video memory, and all network parameters will change. The initial round refers to the number of rounds at the beginning of training. If the initial round is 60, then the number of freezing rounds is 50, and the number of unfreezing rounds is 100. The training process skips the freezing stage for unfreezing training. The batch size determines how many times the training set image is input at once, and the number of rounds represents a complete iterative training. The learning rate is a parameter that ensures the network's convergence. The experimental parameter settings are shown in Table 2.

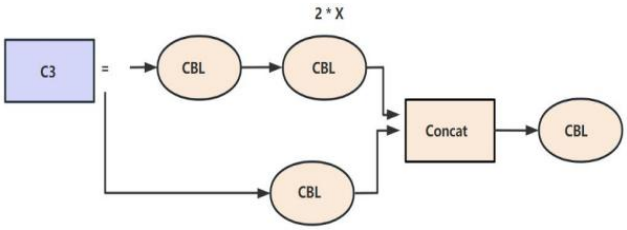


Figure 10: Neck part C3 module network structure

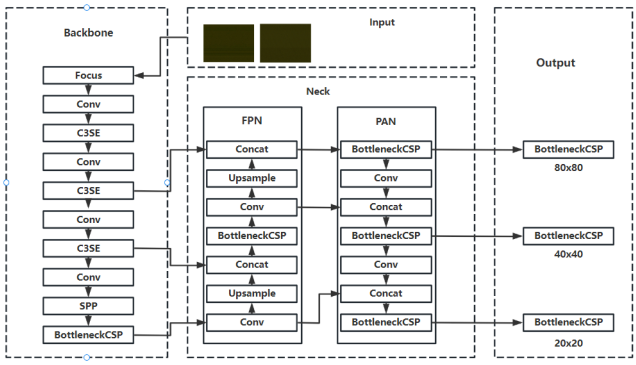


Figure 11: YOLOv5s network structure diagram with SE attention added

Table 1: Experimental environment

Configuration Items	Configuration
Operating System Version	Ubuntu 18.04.6 LTS
CPU Model	Intel(R) Xeon(R) Gold 5318Y CPU @ 2.10GHz
Graphics Card Type	NVIDIA A16
CUDA Version	11.6

Table 2: Parameter setting

input size	640*640
initial epoch	0
freeze epoch	50
Whether to freeze	True
freeze batch size	16
unfreeze batch size	8
learning rate	0.0001
optimizer	SGD
momentum	0.937

5.3.2 Model Evaluation Metrics

In this experiment, Mean Average Precision (MAP) and validation set loss (Val_Loss) are selected as the evaluation indicators of the model. MAP, which represents P

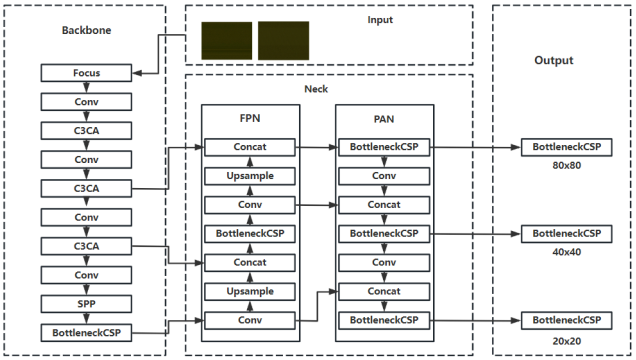


Figure 12: YOLOv5s network structure diagram with CA attention added

(Precision) accuracy. AP (Average precision) is the precision rate of the single-class label average (the average of the maximum precision rate in each recall rate), and MAP (Mean Average Precision) is the average precision rate of all class labels.

5.3.3 Experimental Dataset

The experimental dataset used in this paper was obtained from the Kaggle Malicious Code Classification Contest. The dataset consists of six collections of malicious code, comprising a total of 1547 samples, including binary and disassembled files. To expand the dataset, we employed data enhancement methods for the malicious code images. The specific classification of the malicious code is included in Table 3.

5.3.4 Experimental Results And Analysis

The aim of this study is to evaluate and compare the effects of various attention mechanisms on the model's detection performance. No changes in content have been made. The objective is to determine if and how different attention mechanisms can contribute to improving the model's ability to detect objects accurately and efficiently. The text adheres to conventional structure and formatting features, with consistent citation and footnote style. The study uses the simplified yoloV5s model of the C3 module, and the yoloV5s model with SEnet attention and CA attention. The language used is clear, concise, and objective, with a formal register and precise word choice. The text is free from grammatical errors, spelling mistakes, and punctuation errors. The models compared in the experiments are YOLOv5s, YOLOv5-CA, and YOLOv5-Se. Figure 13 shows the changes in MAP after 300 rounds of iterative training, and Figure 14 shows the loss on the validation set. Figure 13 shows a comparison chart of the map value curves for the three models on the malicious code image training and verification set. The blue line represents the YOLOv5 network model, the red line represents the YOLOv5-CA network model, and the green line represents the YOLOv5-Se net-

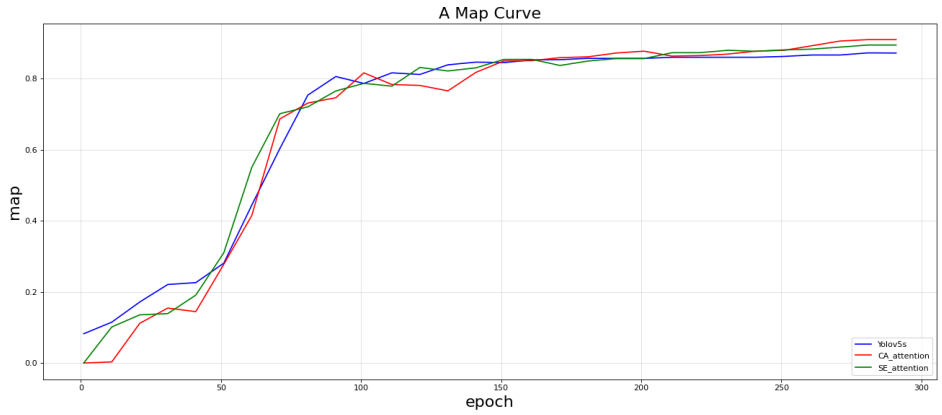


Figure 13: Model MAP curve diagram

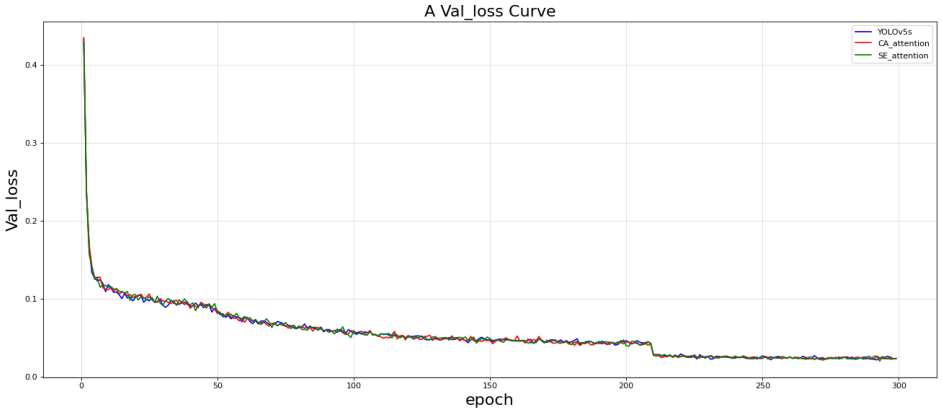


Figure 14: Model validation set loss plot

work model. The figure shows that the map values of the three network models increase as the number of training rounds increases, and they begin to stabilize around 300 rounds. The YOLOv5-Se and YOLOv5-CA network models have higher map values than the YOLOv5 network model for the same training round, as seen in the curve.

Table 4 shows that among the YOLO series algorithms, the YOLOv5 network model performs well and has a relatively lightweight size, making it suitable for application deployment. Therefore, this article selects the YOLOv5 network model as the basis for the experiment. Two new models, YOLOv5-Se and YOLOv5-CA, were created. The YOLOv5-CA model achieved the highest MAP value and the best performance.

6 Conclusions

The paper proposes using the improved YOLOv5s model to classify images of malicious code. Firstly, the Bottleneckcsp module of YOLOv5 is simplified and replaced

Table 3: Number of malicious code samples

Malicious Family Code	Number of training samples	Type
Ramnit	272	worms
Tracur	376	Trojan horse
Vundo	240	Trojan horse
Gatak	179	Back Door
Obfuscator.ACY	300	Malvertising
Lollipop	180	Malvertising

Table 4: Comparison of experimental results of malicious code classification methods

Detection algorithm	MAP	Model Size
YOLOv4	87.56%	246.19MB
YOLOv5	89.21%	27.19MB
YOLOv5-CA	94.36%	27.27MB
YOLOv5-Se	93.25%	27.25MB

with the C3 module. This simplifies the network model structure, reducing the number of network parameters and calculation amount of the model. Additionally, the SE attention mechanism and the CA attention mechanism are added. After comparing the experimental results, it was found that the model with the CA attention mechanism outperformed the model with the SE attention mechanism and the model without any attention mechanism. Therefore, the simplified yoloV5s model with added CA attention was chosen as the improved model for classifying malicious code images.

Acknowledgments

This study was supported by the 2023 Liaoning Province Applied Basic Research Program Project, 2023JH2/101300203. The authors gratefully acknowledge the anonymous reviewers for their valuable comments.

References

[1] C. Huang, Y. Ye, Y. Jin, and B. Liang, "Research progress, hotspots, and evolution of nighttime light pollution: Analysis based on wos database and remote sensing data," *Remote Sensing*, vol. 15, no. 9, p. 2305, 2023.

[2] L. Nataraj, S. Karthikeyan, G. Jacob, and B. S. Manjunath, "Malware images: visualization and automatic classification," in *Proceedings of the 8th international symposium on visualization for cyber security*, 2011, pp. 1–7.

[3] D. Xu, Y. Mi, and M. Wei, "Worm defense and detection technology based on computer network," *Scientific and Technological Innovation*, 2022.

[4] S. Li, L. Jiang, Q. Zhang, Z. Wang, Z. Tian, and M. Guizani, "A malicious mining code detection method based on multi-features fusion," *IEEE Transactions on Network Science and Engineering*, 2022.

[5] P. G. Balikcioglu, M. Sirlanci, O. A. Kucuk, B. Ulukapi, R. K. Turkmen, and C. Acarturk, "Malicious code detection in android: the role of sequence characteristics and disassembling methods," *International Journal of Information Security*, vol. 22, no. 1, pp. 107–118, 2023.

[6] J. Cheng, J. Zheng, and X. Yu, "An ensemble framework for interpretable malicious code detection," *International Journal of Intelligent Systems*, vol. 37, no. 12, pp. 10 100–10 117, 2022.

[7] A. Sejfia and M. Schäfer, "Practical automated detection of malicious npm packages," in *Proceedings of the 44th International Conference on Software Engineering*, 2022, pp. 1681–1692.

[8] K. S. Han, J. H. Lim, B. Kang, and E. G. Im, "Malware analysis using visualized images and entropy graphs," *International Journal of Information Security*, vol. 14, pp. 1–14, 2015.

[9] A. Davis and M. Wolff, "Deep learning on disassembly data," *BlackHat USA*, 2015.

[10] Y. Jiang, Y. Wu, and F. Zou, "Malicious code classification model based on image vector," *Communication Technology*, vol. 51, no. 12, pp. 2953–2959, 2018.

[11] H. Li and L. Qian, "A survey of malicious code visual detection technology research," *SOFTWARE GUIDE— SOFTWARE GUIDE*, 2022.

[12] B. WANG, H. CAI, and Y. SU, "Classification of malicious code variants based on vggnet," *Journal of Computer Applications*, vol. 40, no. 1, p. 162, 2020.

[13] Z. Wu, X. Ye, and M. Chen, "Surface defect detection method of medicinal capsules based on improved yolov5," *Packaging Engineering*, vol. 43, no. 23, pp. 297–304, 2022.

[14] H. Xiao, X. Ye, X. Luo, and P. Wang, "Pedestrian detection with multi-scale spatial pyramid pooling pcanet," *Computer Engineering*, vol. 45, no. 2, pp. 270–277, 2019.

[15] C. Xu, F. Fan, G. Ke, and J. Shen, "Action recognition method based on multi-scale channel attention mechanism," *Electronic Measurement Technology*.

[16] V. Mnih, N. Heess, A. Graves *et al.*, "Recurrent models of visual attention," *Advances in neural information processing systems*, vol. 27, 2014.

Biography

Zipeng He biography. He Zipeng is studying at Shenyang University of Technology and is currently a second-year graduate student. His research interests include mainly deep Learning, artificial intelligence, data analysis, network security.

Yuntao Zhao biography.Yuntao Zhao received the Ph.D. degree in control science and engineering from Nanjing University of Science and Technology, Nanjing, China, in 2013. He is a Post-Doctoral Researcher with the pattern recognition and artificial intelligence, Northeastern University,Shenyang, China, in 2015. He is currently a Professor with the Communication and Network Institute and also with the School of Information Science and Engineering, Shenyang Ligong University, Shenyang. He has authored over 30 papers published in related international conference proceedings and journals. He is the holder of 10 patents and software copyrights. His research interests include mainly deep Learning, AI algorithm, Cyberspace security, protocol analysis and data mining.

Yongxin Feng biography. Yongxin Feng received the Ph.D. degree in computer application technology from Northeastern University, Shenyang, China, in 2003. She is currently a Professor in the School of Information Science and Engineering, Shenyang Ligong University. She has authored over 80 papers in related international conferences and journals. She is the holder of 30 patents and software copyrights. Her research interests are in the areas of intelligent information processing, wireless sensor network, communication and information systems.

Secure Localization Method in WSN Based on Improved MCL Algorithm

Lijun Mao[1] and Yitong Zhang[2]
(Corresponding author: Lijun Mao)
School of Intelligent Science and Information Engineering, Xi'an Peihua University, Xi'an, China[1]
Email: mlj78454@yeah.net
School of Journalism and Communication, Northwest University of Political Science and Law, Xi'an, China[2]
(Received June 15, 2023; Revised and Accepted May 11, 2024; First Online Aug. 16, 2024)

Abstract

Wireless sensor networks serve as a crucial link between the physical and information world by perceiving, collecting, and transmitting data through sensor nodes. However, these networks are susceptible to various malicious attacks, highlighting the significance of ensuring accurate and secure node localization. This study enhances the Monte Carlo node localization algorithm, investigates the impact of wormhole attacks on the improved Monte Carlo localization scheme, and introduces a secure Monte Carlo localization box algorithm designed to counteract wormhole attacks. Test results demonstrate that an unknown node density of 10, an anchor node density of 1.1, and a maximum speed of 20m/s contribute to minimizing node positioning errors. Moreover, the improved secure localization algorithm effectively withstands wormhole attacks while reducing energy consumption. With 300 experiments, the energy consumption is only 40.00% and 23.53% compared to the reference algorithm; with 600 experiments, the energy consumption drops to merely 38.89% and 22.58% of the reference algorithm's consumption. The algorithm achieves high coverage for node localization, ranging from 95% to 100%. The positioning error remains stable at approximately 10m, which is 56.29% and 38.29% lower than the comparison algorithm. The secure localization algorithm successfully combats the interference caused by wormhole attacks and accomplishes the localization process securely. Overall, this study's findings contribute significantly to advancing research on secure positioning within wireless sensor networks.

Keywords: Location; MCL; Network Security; Wormhole Attacks; Wireless Sensor

1 Introduction

Wireless Sensor Networks (WSN) are distributed networks consisting of stationary or mobile micro-sensor nodes that enable data acquisition, processing, and transmission. These intelligent networks, formed through self-organization or multi-hop methods, find applications in various fields such as military, agriculture, industry, healthcare, and traffic monitoring, playing a pivotal role in shaping the world's future development [10]. Node localization in WSN refers to determining the precise position of each node within the network. Accurate node localization is vital for tasks like target tracking, area monitoring, and event detection. Node density, which refers to the distribution density of nodes in the network, is an important consideration in WSN design as it directly impacts coverage, data quality, and energy consumption. Hence, when applying wireless sensor networks, one must carefully consider both the accuracy of node localization and the selection of node density. However, WSN often operates in complex environments where large positioning deviations or abnormal attacks can lead to node failures or the loss of detection data. Consequently, ensuring the security and accuracy of node positioning is of utmost importance [8].

Various techniques exist for node localization, including distance-based methods like arrival time algorithms, received signal strength index, and angle of arrival algorithms, as well as non-distance-based approaches such as centroid localization, convex programming, DV-HOP, and periodic execution of static network localization. However, these techniques have drawbacks, ranging from high energy consumption to low localization accuracy [15, 18]. Moreover, malicious attacks on the network can further impair node localization, yet current research rarely addresses the security and performance evaluation of localization algorithms. Therefore, further research is needed to mitigate the impact of such attacks on node localization. This study presents a novel algorithm, the Monte Carlo Location algorithm (MCL), which addresses the challenges of mobility-based node localization. It analyzes the effects of wormhole attacks on the improved MCL and proposes a secure MCL capable of resisting such attacks. The study consists of four parts: an overview of the current research status on node localization in WSN, the proposal of a secure MCL using Monte Carlo algo-

rithms, validation of the secure MCL's performance, and a summary of the research results. This secure node localization algorithm aims to achieve real-time precision and security for mobile node localization while reducing energy consumption and achieving satisfactory localization outcomes.

2 Related Works

Wireless sensor networks have gained significant prominence in the field of information technology, with node localization emerging as a critical technology within WSN. Researchers have made efforts to enhance node localization algorithms and improve the security of localization in WSN. Liu *et al.* devised a WSN node algorithm that integrates particle swarm optimization and monkey algorithms to address large localization errors and low precision in wireless sensor nodes. By initializing the population using Laplace distribution and employing the particle swarm optimization algorithm, the proposed method effectively avoids falling into local optima. Evaluation results demonstrated superior performance in terms of reference node rate, node density, and communication indicators, indicating its strong localization capabilities [8]. When it comes to the self-localization of random sensors within a designated area, beacon nodes are commonly utilized. To mitigate the negative impact of the trajectory of moving beacon nodes on localization accuracy and efficiency, Sabale and Mini introduced a cosine rule-based method to locate the static trajectory of these nodes. By applying the cosine rule to received positions and deriving distances, the algorithm achieves higher localization accuracy, thereby improving overall performance [14]. In order to enhance the localization accuracy of WSN nodes, Fu *et al.* proposed a multiple center localization algorithm based on an improved receive signal strength indicator ranging technique. Through iterative filtering to refine the receive signal strength indicator and incorporating a multiple center algorithm to optimize trilateration, the proposed approach significantly enhances the accuracy of node distance calculation in wireless sensor networks. Simulation results validate the improved localization accuracy and reduced localization deviation of nodes, all achieved within existing computing resources [2]. To optimize coverage and monitoring in distributed sensor networks, Yao *et al.* introduced an algorithm for effectively deploying coordinated sensors. However, the traditional use of the receive signal strength indicator in the virtual force algorithm resulted in oscillation when approaching the optimized position. To further optimize the algorithm, the study incorporated a whale swarm algorithm to refine the cooperative positioning of adjacent nodes. Simulation results demonstrated that the coordination algorithm achieved a 95% coverage rate for distributed sensor networks, highlighting its effectiveness and feasibility [20].

To overcome the low accuracy of the traditional least squares method, Ouyang *et al.* introduced an improved adaptive genetic algorithm and evaluation function to improve accuracy and reduce distance measurement errors. Experimental results showed that the node positioning accuracy of the improved adaptive genetic DV-Hop algorithm was higher [11]. The underwater wireless channel is weak, and the topology structure of sensors may also change underwater, so the underwater working environment has higher requirements for sensor networks. On the one hand, Kaimal and Binu designed a prediction method based on energy-efficient derivatives to predict data and node terrain movement. At the same time, to resist wormhole attacks, a security-aware local constraint algorithm was designed, which can effectively manage dynamic networks and protect the confidentiality of data [4]. Faced with the unavailability of traditional positioning technology in underwater network sensor positioning, Toky *et al.* designed an underwater network sensor positioning method based on arrival angle technology, which was divided into angle estimation stage, projection stage and positioning stage. The experimental results showed that this positioning method had a high positioning rate and positioning coverage rate, and low energy consumption. Traditional receive signal strength and arrival angle measurement techniques cannot estimate the position of multiple targets [17]. Kang *et al.* designed a clustering method based on k-means and expectation maximization for multi-target localization in wireless sensor networks, which achieved the estimation of multiple target positions through multiple single-objective estimates. Simulation experiments verified its effectiveness [5]. To minimize positioning errors, Tang and Han proposed a wireless sensor network mixed received signal strength index method by weighted centroids and adaptive threshold selection. The improved algorithm reduced the positioning errors caused by complex environments and signal oscillations, and experimental results showed that the method had better positioning accuracy and higher stability than conventional signal strength index methods [16].

In summary, although many traditional positioning algorithms have been optimized and improved to improve the positioning accuracy of wireless sensor network nodes, existing research focuses mainly on reducing node positioning errors or protecting data confidentiality in wireless sensor networks [4, 17]. Research on mobile node localization that considers both localization accuracy and security is still quite rare, which provides the possibility of ensuring the security of data transmission during the localization process and improving the accuracy of node detection.

3 Node Security Localization Algorithm Based on Improved Monte Carlo

WSN consists of nodes, sensor networks, and users. Due to the wireless transmission of information in WSN, the

transmission process is easily susceptible to external intrusion and malicious attacks, leading to information leakage and damage in WSN, and greatly reducing its security performance. The secure localization of nodes requires ensuring the integrity of node communication, the privacy of node location coordinates, the availability of localization algorithms, and the authenticity of transmitted data. Based on this, this paper studies and improves MCL for mobile WSN, further analyzes the impact of wormhole attacks on the improved MCL box algorithm (MCB), and proposes a secure MCB to resist wormhole attacks.

3.1 Node Location Algorithm Based on MCL

The node localization of WSN is to obtain the absolute or relative position of the target node based on a small amount of known location information. For static wireless sensor networks, location technology can be divided into distance-based and distance-independent algorithms. For mobile sensor networks, there is a situation where anchor nodes or ordinary nodes move. If static node positioning algorithms are directly used, it will result in high computational complexity, high communication costs, and low positioning accuracy. Therefore, for mobile sensor networks, MCL is used for research. It is essentially a combination of sampling mobile nodes and particle filtering. By constructing the probability distribution of node positions through some weighted random samples, the essence of MCL is to estimate the position of nodes. MCL algorithm has a simple structure, strong stability, and adaptability, and is suitable for dealing with non-Gaussian, nonlinear, and flexible problems, such as node positioning in mobile wireless sensor networks [6, 21].

The MCL algorithm belongs to distributed algorithms, and its core essence is continuous iterative Bayesian filters. Bayesian filters can estimate node positions through node observations. When facing mobile nodes, they only need to iterate continuously and combine historical position information with current observations to achieve node position prediction. The recursive formula for Bayesian filtering is shown in Equation (1), where x_t represents the state of the system at time t, z_t represents the observed value at time t, η is a constant, $p(z_t|x_t)$ represents the observed model, $p(x_t|l_{t-1})$ represents the model of the system moving with action, and u_t represents the predicted system state.

$$Bel(x_t) = \eta p(z_t|x_t) \int p(x_t|u_t, x_{t-1}) Bel(x_{t-1}) dx_{t-1} \quad (1)$$

MCL is the most basic Monte Carlo positioning method. The algorithm is based on two important assumptions: first, time is composed of discrete time units, and second, the movement of nodes is based on the Markov chain. MCL's prediction of mobile nodes is achieved through a state transition equation $p(l_t|l_{t-1})$, where l_t represents the node position distribution at time t, and the l_t position distribution composed of N sample points is defined as $L_t = \{l_t^1, l_t^2, \cdots, l_t^N\}$. The observation equation is defined as $p(o_t|l_t)$, which represents the probability that a node conforms to the observation results, and o_t represents the observation results of the undetermined node during the time change process. For moving nodes, the prediction of node positions is determined by both o_t and L_{t-1}, and each sample point is assigned different weights [13].

The posterior probability distribution of nodes is expressed in Equation (2), and w_t^i represents the normalized weight value of sample points.

$$p(l_t|o_{0:t}) \approx \sum_{i=1}^{N} w_t^i \delta(l_t - l_t^i) \quad (2)$$

Samples are taken from the posterior probability distribution through the importance function of the known distribution. The normalized importance equation is shown in Equation (3).

$$\pi(l_t|o_{0:t}) = p(l_0) \prod p(l_k|l_{k-1}) \quad (3)$$

Finally, the update and normalization operation of sample weights is shown in Equation (4), where w_t^i represents the updated weight and w_t^i represents the normalized weight.

$$\begin{cases} w_t^i = w_{t-1}^i p(o_t|l_t^i) \\ w_t^i = \frac{w_t^i}{\sum_{k=1}^{N} w_t^k} \end{cases} \quad (4)$$

The MCL algorithm first initializes the algorithm, and the initialized sample set needs to be randomly sampled from the node deployment area. After sampling, a new sample set is generated based on the historical sample set and node movement. If $v_{\max}$ represents the possible maximum speed of node movement in the node position prediction model, the possible position at the current time will appear within the circle range where the sample is the center and $v_{\max}$ is the radius at the previous time.

The state transition equation of the node is shown in Equation (5). $d(l_t, l_{t-1}$ represents the Euclidean distance between sample points at two-time points. By obtaining a new sample set and changing the prediction model of nodes, more node motion states can be obtained, resulting in more accurate prediction results.

$$p(l_t|l_{t-1}) = \begin{cases} \frac{1}{\pi v_{\max}^2}, & d(l_t, l_{t-1}) \le v_{\max} \\ 0, & d(l_t, l_{t-1}) > v_{\max} \end{cases} \quad (5)$$

However, the node positioning determined based on this method may have some negative impacts, including some points that do not have a positive effect on prediction or are useless, as shown in Figure 1.

For nodes that may have negative impacts, the MCL algorithm introduces sample filtering operations. The filtering equation is shown in Equation (6). In Equation (6), r is the communication radius, s is the anchor node, S is the set of one hop anchor nodes, and T is the set of two hop anchor nodes. When the filtering conditions are met, $p(o_t|l_t) = 0$; Otherwise, $p(o_t|l_t) = 1$.

$$filter(l) = \forall s \in S, d(l, s) \le r \wedge \forall s \in T, d(l, s) \le 2r \quad (6)$$

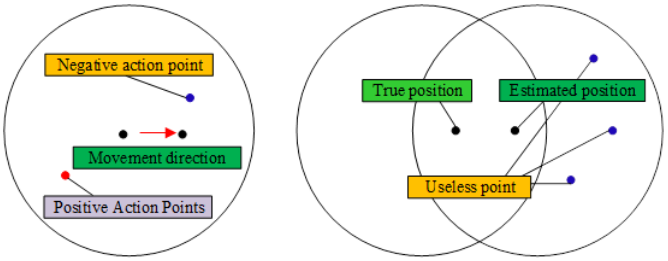


Figure 1: Schematic diagram of the negative impact of node positioning

When the sampling point is less than N, it is necessary to repeat sampling and filtering. During the process of repeated sampling, sampling is reduced with low-importance weights. In addition, a resampling sampling scale has been set to prevent algorithm degradation, as shown in Equation (7). When the sampling size is greater than the sampling threshold, it can prevent algorithm degradation. Finally, after obtaining the samples, the average of the weighted sample points is calculated to estimate the node position.

$$N_{eff} \approx \frac{1}{\sum_{i=1}^{N}(w_t^i)^2} \tag{7}$$

The overall process of the MCL algorithm is shown in Figure 2. Overall, the MCL algorithm has good comprehensive application performance, but there are still some shortcomings and shortcomings, such as insufficient positioning accuracy, consumption of node energy, and low sampling efficiency. The assumptions of equal weight values for sampling points and a fixed communication radius do not align with the real-world scenario [7].

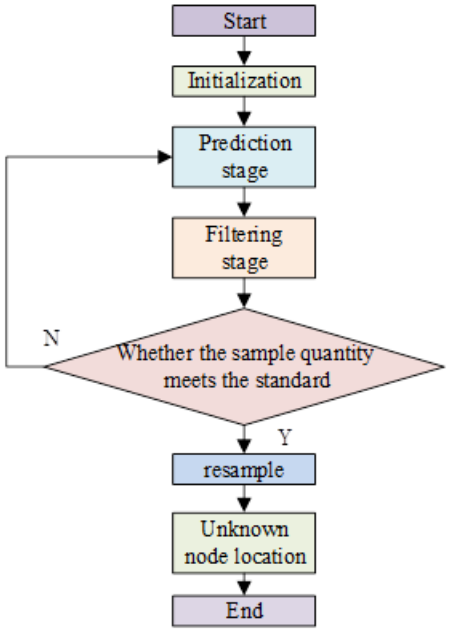


Figure 2: MCL algorithm flowchart

3.2 Localization Algorithm for Resisting Wormhole Attacks Based on Improved MCL

In response to MCL's shortcomings, an improved MCB was adopted in the study. MCB reduced the sampling area using anchor boxes on the basis of MCL, and the schematic diagram of the anchor box is shown in Figure 3 [22].

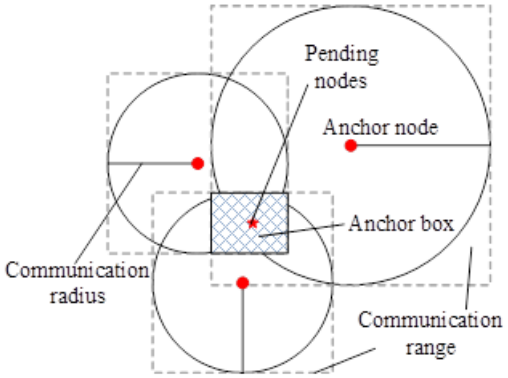


Figure 3: Schematic diagram of anchor box

The boundary definition of the anchor box of a node is shown in Equation (8), where (x_i, y_i) represents the coordinates of the anchor node and n represents the number. When $x_{\min} > x_{\max}$, the anchor box does not exist. To reconstruct the anchor box, only one hop anchor node needs to be considered, or the anchor box should be reconstructed outside the node deployment area, and the deployment boundary value should replace the boundary

value of the anchor box.

$$\begin{cases} x_{\min} = \max_{i=1}^{n}(x_i - r) \\ x_{\max} = \min_{i=1}^{n}(x_i + r) \\ y_{\min} = \max_{i=1}^{n}(y_i - r) \\ y_{\max} = \min_{i=1}^{n}(y_i + r) \end{cases} \tag{8}$$

The reconstructed sampling area is constructed based on the movement of the sample points at the previous time, and the boundary definition of the sampling box is shown in Equation (9). In Equation (9), $v_{\max}$ represents the maximum speed of the node.

$$\begin{cases} x_{\min}^{i} = \max(x_{\min}, x_{t-1}^{i} - v_{\max}) \\ x_{\max}^{i} = \min(x_{\max}, x_{t-1}^{i} + v_{\max}) \\ y_{\min}^{i} = \max(y_{\min}, y_{t-1}^{i} - v_{\max}) \\ y_{\max}^{i} = \min(y_{\max}, y_{t-1}^{i} + v_{\max}) \end{cases} \tag{9}$$

Afterward, the same sample initialization, node prediction, filtering, and resampling operations are performed as traditional MCL. The improved MCL significantly reduces computational complexity and improves positioning accuracy.

In some special application scenarios, mobile wireless sensor networks may sometimes be subjected to malicious attacks from outside the network, interfering with the positioning work of normal nodes [19]. Common malicious attack models include deception attacks that introduce malicious data, witch attacks that cause node misalignment, and wormhole attacks that form incorrect network topology. Research mainly focuses on secure localization in the context of wormhole attacks. A wormhole attack is an attack initiated by multiple attack nodes in collaboration. The attack exchanges its positioning information through dedicated wormhole links, forming incorrect network topology and causing serious negative impacts on the positioning process. The attack schematic of wormhole attack is shown in Figure 4 [1, 3]. Wormhole attacks, also known as tunnel attacks, can establish dedicated wormhole links between two attack nodes that are far apart. The transmission efficiency of stolen information in private tunnels is higher than that of normal multihop path data transmission. Nodes maliciously promote this false and efficient path, deceiving legitimate nodes to choose private tunnels. The target of wormhole attacks is to interfere with network communication and disrupt the integrity and security of the network. Wormhole attacks may cause problems such as information leakage, data tampering, or denial of service in the network. The MCL algorithm utilizes first-order and second-order anchor node information for localization. The main attack scenario of wormhole attacks on blind nodes is the presence of real and false anchor nodes around the blind node, which can lead to the MCL algorithm not obtaining legitimate samples and reducing the node localization rate.

The secure localization of WSN is that, while facing various malicious attacks, certain security techniques can

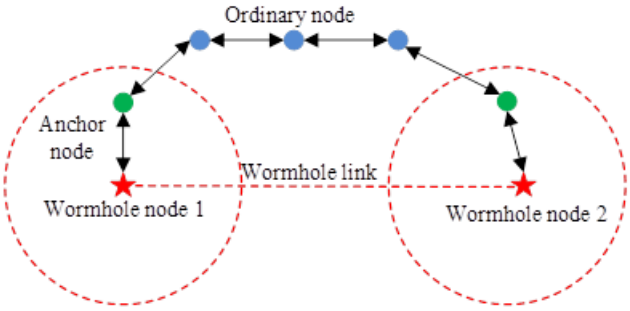


Figure 4: Schematic model of wormhole attack

still effectively and accurately obtain the estimated location information of nodes. The existing security localization ideas for responding to attacks include establishing a key mechanism, setting up a distance constraint mechanism, and conducting attack detection and location verification. The SecMCL algorithm first considers the MCL algorithm's ability to resist spoofing attacks and is also a basic secure localization method. Each node in SecMCL is equipped with a private key, which is responsible for decrypting the received node broadcast messages. Under the influence of deception attacks, if the effective sample size decreases or effective samples cannot be obtained, SecMCL adopts relaxed filtering conditions to achieve the purpose of expanding sampling. The new filtering conditions are shown in Equation (10). In Equation (10), i represents the neighbor anchor node, and Q is the neighbor anchor node set [23].

$$\begin{aligned} filter(l,i) \quad = \quad & \exists Q(Q \subseteq S \cup T)(sizeQ = i) \\ & \cap(\forall s(s \in S \cap s \in Q), d(l,s) \leq r) \\ & \cap(\forall s \in T \cap s \in Q), 2r \geq d(l,s) > r) \end{aligned} \tag{10}$$

SecMCL algorithm does not take into account the positioning error of regular nodes when responding to malicious attacks, which indicates a need for further improvement in the localization rate while under attack [9, 12]. In addressing wormhole attacks, the study seeks to enhance the MCL and proposes the MCB security localization algorithm. Wormhole attacks can potentially impact the creation of MCB anchor boxes, and thus, the proposed approach adopts a trust-based mechanism to improve MCB by pre-selecting trustworthy anchor nodes during the construction of anchor boxes. The resulting secure localization algorithm, namely DewormMCB, is illustrated in Figure 5.

Firstly, the DewormMCB algorithm is initialized, and the initial position positioning is shown in Equation (11). In Equation (11), X_{range} and Y_{range} represent known values, when the positioning node sends a positioning request.

$$\begin{cases} x_0 = X_{range} \times random(0,1) \\ y_0 = Y_{range} \times random(0,1). \end{cases} \tag{11}$$

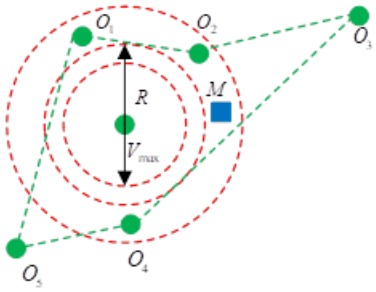


Figure 5: Schematic diagram of DewormMCB algorithm

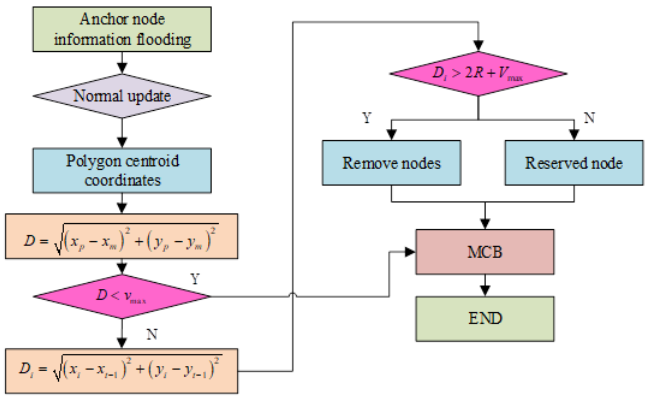


Figure 6: Schematic diagram of DewormMCB algorithm flow

Next, the polygon centroid composed of all elements in the set monitored by the positioning node at a certain time is calculated, as shown in Equation (12). In Equation (12), x_m and y_m are the horizontal and vertical coordinates of the centroid, while x_i and y_i are those of the anchor node element.

$$\begin{cases} x_m = \sum_{i=1}^{n} x_i/n \\ y_m = \sum_{i=1}^{n} y_i/n \end{cases} \tag{12}$$

The criteria for determining pseudo anchor nodes are shown in Equation (13), where x_p and y_p represent the horizontal and vertical coordinates of the positioning node. If $D > v_{\max}$, there are pseudo anchor nodes, and if there are no pseudo anchor nodes, the MCB algorithm can be directly used for secure positioning.

$$D = \sqrt{(x_p - x_m)^2 + (y_p - y_m)^2} \tag{13}$$

For situations where there are pseudo anchor nodes, it is necessary to remove them. O_t represents the set of anchor nodes, while the newly added anchor nodes for O_t and O_{t-1} are O_1 and O_2. The distances between O_1, O_2, and O_{t-1} are compared. The calculation formula is shown in Equation (14). If $D_i > 2R + v_{\max}$, O_i is removed.

$$D_i = \sqrt{(x_i - x_{t-1})^2 + (y_i - y_{i-1})^2} \tag{14}$$

Finally, using the obtained set of anchor nodes, the anchor box is reconstructed and the nodes are located. The algorithm flowchart is shown in Figure 6. The MATLAB programming language is used to develop code that simulates algorithms, while also selecting suitable experimental parameters and scenarios for testing. Additionally, the code collaborates with relevant libraries for comprehensive data processing and analysis.

4 Performance Testing of Improved Secure MCL

To verify the effectiveness of the improved MCL algorithm for secure localization, simulation experiments were conducted in Matlab. In the simulation model, all nodes are randomly and uniformly distributed within the rectangular area of 500 × 500, and the study adopts a Random WayPoint motion model with random motion speed and direction.

4.1 Parameter Settings Impact on Algorithm Performance

In the simulation experiment, the number of wormhole links was set to 6, and the wormhole attacks were scattered in the network. The node motion radius was 50, the number of samples was 50, the communication irregularity was 0, and the node movement range was always within the boundary range. The experimental results were taken as the average of 50 runs of Matlab. Model parameter experiments were conducted on three different scenarios, namely MCB in a secure environment, attacked MCB, and attacked DewormMCB. The impact of anchor node density and unknown node density on positioning error is shown in Figure 7. From Figure 7(a), the unknown node density increased by 10, and the positioning error of MCB showed a decreasing trend in all three cases; As the density of unknown nodes continued to increase, the positioning error no longer changed significantly. The positioning error of DewormMCB ultimately stabilized at around 18.00m. If the density of unknown nodes increases, more anchor points become available for node positioning, and the resulting increase in positioning information leads to enhanced positioning accuracy. However, the density of unknown nodes increased to a certain extent and no longer contributed to the available positioning information. Therefore, the most suitable density for research was 10. An increase in anchor node density was beneficial for increasing the number of one and two-hop anchor nodes, providing more positioning information. The algorithm's positioning error curve showed a downward trend before anchor node 1.1. Similarly, the contribution of anchor node density to reducing positioning errors was limited. When the anchor node density

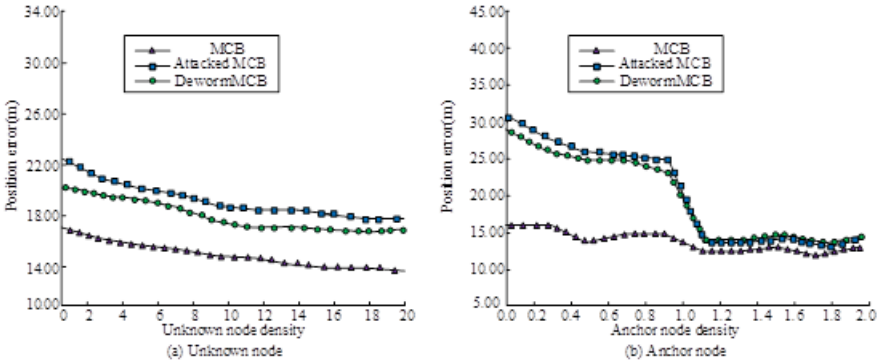


Figure 7: The influence of different parameter settings on positioning error

was too high, node positioning was affected by wormhole attacks, and the positioning error curve tended to overlap. Therefore, it was more appropriate to set the anchor node density to 1.1.

The effect of maximum speed on positioning error is shown in Figure 8. As shown in Figure 8, as the maximum velocity value continued to increase, the positioning error showed a trend of first decreasing and then increasing. When the node movement speed was low, the sampling area would also be relatively small, and the actual position may be outside the sampling area, resulting in significant positioning errors. However, excessive movement speed can lead to a large sampling area, reducing the accuracy of node positioning. Therefore, it was more appropriate to set the maximum speed to 20m/s.

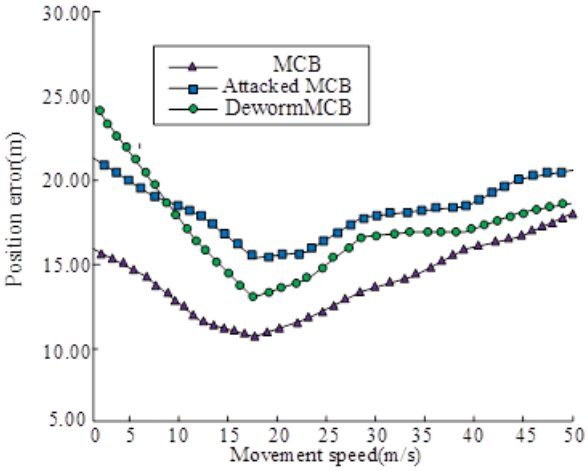


Figure 8: The influence of maximum motion speed on positioning error

4.2 Algorithm Security Localization Performance

Firstly, simulation experiments were conducted on MCB positioning errors in both secure and wormhole attack environments, with consistent parameter settings. Results are shown in Figure 9. Wormhole attacks have a significant impact on the positioning error of the MCB algorithm, with significant differences between the two curves. When there was no wormhole attack, the maximum error of the MCB was only 18.29m, with most remaining within 10.00m; When a wormhole attack occurred, the maximum positioning error of MCB exceeded 50m, and the traditional MCB algorithm failed.

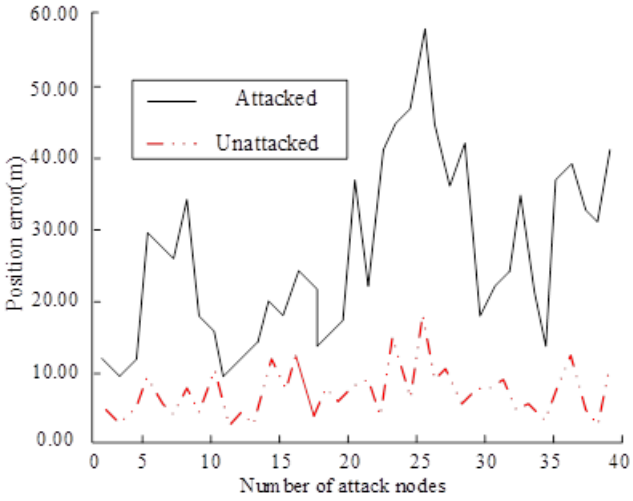


Figure 9: Comparison of wormhole attack errors

800 rounds of network experiments were conducted to compare the power consumption of DewormMCB, SecMCL, and SenLease security localization algorithms, with an initial energy of 0.5J. Results are shown in Figure 10. As shown in Figure 10, as the experiment progressed, all three algorithms continued to consume energy, with an increasing trend in energy consumption and a decreasing trend in the average remaining energy of nodes. The energy consumption gap between the three algorithms was also increasing. At 300 experiments, the energy consumption of DewormMCB was 23.53% of SecMCL and 40.00% of SenLease; At 600 cycles, the energy consumption of DewormMCB was 22.58% of SecMCL and 38.89% of SenLease. When conducting

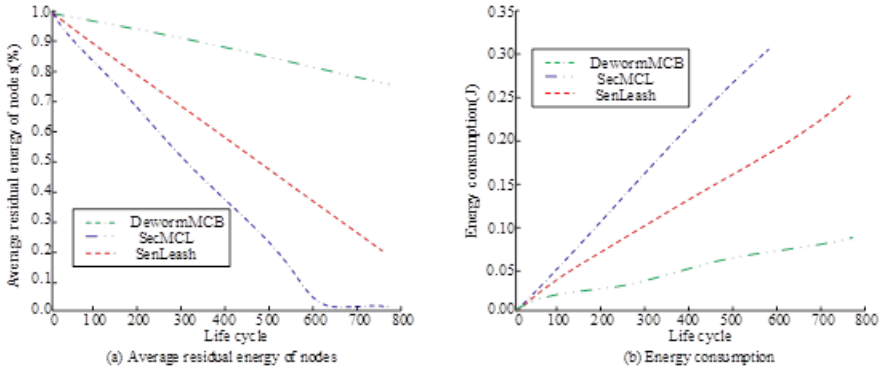


Figure 10: Comparison of power consumption and cost of different algorithms

about 600 experiments, SecMCL had almost no remaining energy and the node lifecycle was depleted. Algorithm energy refers to the algorithm under information constraints. To sum up, the DewormMCB algorithm can effectively resist wormhole attacks and save energy consumption.

The experimental results of localization coverage of different algorithms are shown in Figure 11. Location coverage represents the proportion of nodes that have been successfully located within a specific time frame to all unknown nodes. As shown in Figure 11, the DewormMCB algorithm was easier to achieve localization conditions, and the coverage of node localization was relatively high, basically within the range of 95% -100%. The SecMCL algorithm had the lowest localization coverage, with a minimum value of around 85%. The SenLeash algorithm achieved a relative improvement in localization coverage, with an increase of approximately 5-10 percentage points.

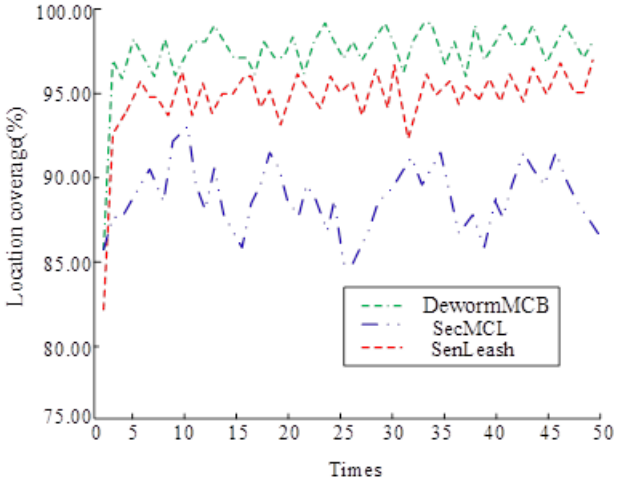


Figure 11: Location coverage of different algorithms

The experimental results of the positioning error and rate are shown in Figure 12. The DewormMCB positioning error was relatively small, while the SecMCL positioning error was relatively large. Over time, the positioning error of DewormMCB ultimately stabilized at around 10m, 56.29% lower than SecMCL and 38.29% lower than SenLease. The positioning accuracy of SenLeash and DewormMCB was relatively close, with a minimum range of over 80%; SecMCL had the lowest positioning accuracy, with a minimum accuracy floating around 70%. The DewormMCB algorithm can effectively avoid false links in wormholes, resist wormhole attacks, improve positioning errors, and improve positioning accuracy.

Finally, the research-designed secure localization algorithm is compared with other existing state-of-the-art algorithms for localization effect and performance, and the improved adaptive genetic DV-Hop algorithm (AG-DV-Hop), PSO-MA-based wireless sensor network node localization algorithm (PSO-MA-NL), and cosine rule-based localization algorithm (CRL) are chosen to use the average localization error, the localization rate, and the area AUC under the receiver operating characteristic curve of the model as evaluation metrics, and the experimental results are shown in Figure 13. As seen in Figure 13, the ROC curve of the DewormMCB algorithm is located at the top of the coordinate axis, and the ROC curve of PSO-MA-NL is located at the bottom of the coordinate axis, and the AUC values are 0.907, 0.869, 0.791, and 0.703 in the order from the largest to the smallest. The localization error curve of the DewormMCB algorithm is significantly lower than the other algorithms, and the localization rate curve is the highest, with the maximum localization rate reaching 0.842. In summary, the performance and localization effect of the DewormMCB algorithm are better than the existing advanced localization algorithms.

5 Conclusion

In response to the accuracy and security issues of node localization in WSN, this study proposes a secure MCB based on MCL to resist wormhole attacks. Test results demonstrated that as the density of unknown and anchor nodes increased, the positioning error first decreased and then stabilized, and there was an optimal value for node

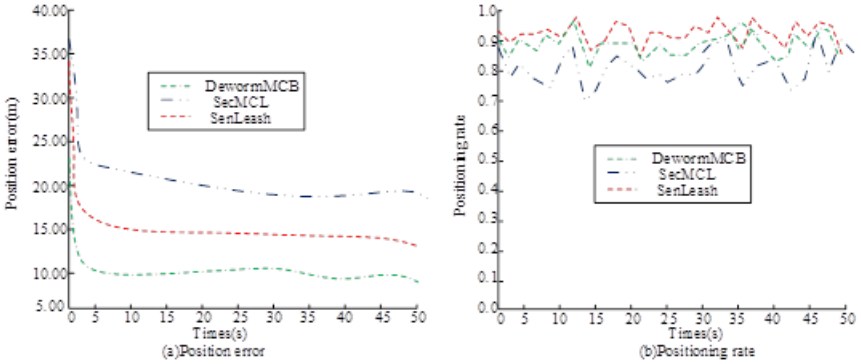


Figure 12: Positioning accuracy and error of different algorithms

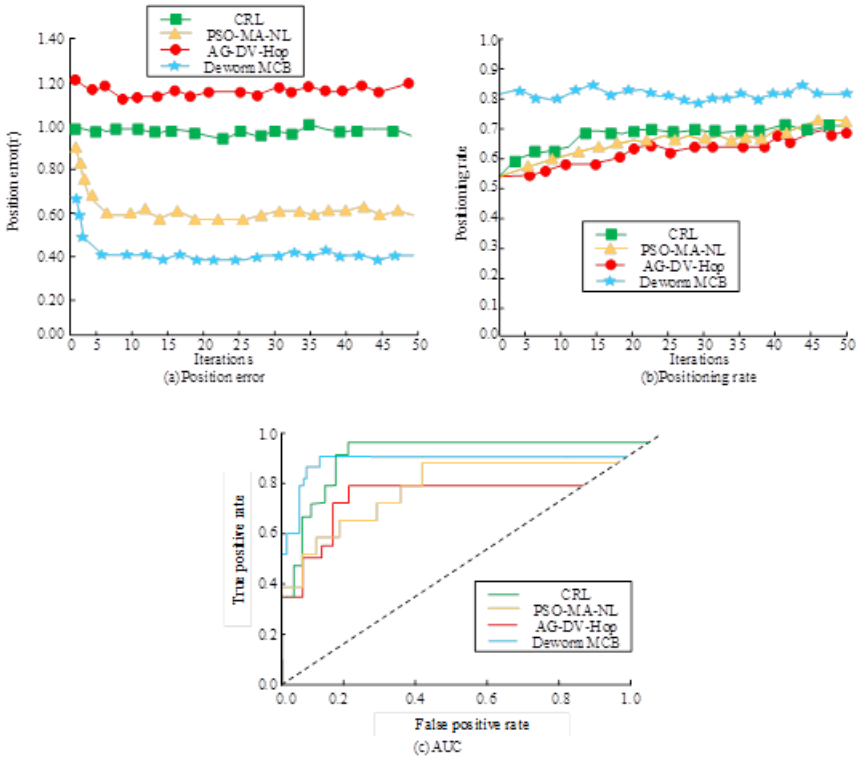


Figure 13: Comparison of performance and localization effect of different localization algorithms

density. Research suggested that an unknown node density of 10 and an anchor node density of 1.1 were more suitable. As the maximum speed value of node movement continued to increase, the positioning error showed a trend of first decreasing and then increasing. It was more appropriate to set the maximum speed of node movement to 20m/s. Wormhole attacks had a significant impact on the MCB positioning error, which indirectly confirmed DewormMCB's effectiveness. The energy consumption of DewormMCB, SecMCL, and SenLease algorithms was on the rise, while the average remaining energy of nodes was on the decline. During 300 experiments, the energy consumption of DewormMCB was 23.53% of SecMCL and 40.00% of SenLease; At 600 cycles, the energy consumption of DewormMCB was 22.58% of SecMCL and 38.89% of SenLease. When conducting about 600 experiments, SecMCL had almost no remaining energy and the node lifecycle was depleted. The DewormMCB algorithm had a high coverage rate for node localization, ranging from 95% to 100%. The SecMCL algorithm had the lowest localization coverage, with a minimum value of around 85%. Over time, the positioning error of DewormMCB ultimately stabilized at around 10m, 56.29% lower than SecMCL and 38.29% lower than SenLease. The positioning accuracy of SenLeash and DewormMCB was relatively close, with a minimum range of over 80%; SecMCL had the lowest positioning accuracy. Compared with existing state-of-the-art algorithms, DewormMCB performs better in terms of both performance and localization. The DewormMCB algorithm effectively resisted interference from wormhole attacks and completed the secure localization process. However, further research is needed on the impact of choosing more attack methods on the performance of secure localization algorithms. For future research, it is recommended to explore localization algorithms in complex scenarios involving implicit, explicit, mobile, and multi-pair wormhole attacks. Additionally, opportunities exist to apply these algorithms to real-world applications and further improve their efficacy in the context of practical constraints and real-world conditions.

Acknowledgments

The research is supported by this paper is one of the research results of the 14th Five-Year Plan for Education Science of Shaanxi Province, "Research and Practice of Online and Offline mixed teaching mode of 'Computer and Information Technology' based on Virtue and cultivating people" (Project No: SGH22Y1822).

References

[1] M. Abrar, Alajlan, "Multi-step detection of simplex and duplex wormhole attacks over wireless sensor networks," *Computers, Materials and Continua*, vol. 70, no. 3, pp. 4241-4259, 2022.

[2] S. Fu, W. Lou, J. Wang, T. Ji, W. Liu, "Multi-barycenter nodes localization method in wireless sensor network based on improved RSSI," *Journal of Beijing Institute of Technology*, vol. 30, no. zk, pp. 210-217, 2021.

[3] D. Hemanand, N. Sankar, Ram, D. S. Jayalakshmi, "FSSAM: A five stage security analysis model for detecting and preventing wormhole attack in mobile Ad-Hoc networks using adaptive atom search algorithm," *Wireless Personal Communications*, vol. 128, no. 1, pp. 487-506, 2022.

[4] I. S. Kaimal, G. S. Binu, "A secure energy efficient event determination algorithm for underwater wireless sensor networks," *International Journal of Networking and Virtual Organisations*, vol. 25, no. 1, pp. 48-61, 2021.

[5] S. Y. Kang, T. H. Kim, W. Z. Chung, "A novel clustering method for multi-target localization based on unidentified RSS/AOA measurements in wireless sensor networks," *The Journal of Korean Institute of Electromagnetic Engineering and Science*, vol. 32, no. 9, pp. 816-825, 2021.

[6] D. Liu, X. Di, B. Xu, "Autonomous vehicle self-localization in urban environments based on 3D curvature feature points-Monte Carlo localization," *Robotica*, vol. 40, no. 3, pp. 817-833, 2021.

[7] Q. Liu, X. Di, B. Xu, 'Autonomous vehicle self-localization in urban environments based on 3D curvature feature points-Monte Carlo localization," *Robotica*, vol. 40, no. 3, pp. 817-833, 2021.

[8] W. Liu, C. Shi, H. Zhu, H. Yu, "Wireless sensor network node localization algorithm based on PSO-MA," *Journal of Web Engineering*, vol. 20, no. 4, pp. 1137-1154, 2021.

[9] C. Lv, J. Zhu, G. Chen, "A localization scheme based on Improving dynamic population monte carlo localization method for large-scale mobile wireless aquaculture sensor networks," *IET Wireless Sensor Systems*, vol. 13, no. 2, pp. 58-74, 2023.

[10] M. Majid, S. Habib, A.R.Javed, M. Rizwan, G. Srivastava, T. R. Gadekallu, J. Chun, "Applications of wireless sensor networks and internet of things frameworks in the industry revolution 4.0: A systematic literature review," *Sensors*, vol. 22, no. 6, pp. 2087-2123, 2022.

[11] A. Ouyang, Y. Lu, Y. Liu, M. Wu, P. Xu, "An improved adaptive genetic algorithm based on DV-Hop for locating nodes in wireless sensor networks," *Neurocomputing*, vol. 458, no. Oct.11, 500-510, 2021.

[12] M. Peavy, P. Kim, H. Oyediran, K. Kim, "Integration of real-time semantic building map updating with adaptive monte carlo localization (AMCL) for robust indoor mobile robot localization," *Applied Sciences*, vol. 13, no. 2, pp. 909-930, 2023.

[13] B. Russell, B. Jakub, C. Marco, M. R. Nowicki, K. Walas, M. Fallon, "Navigating by touch: haptic Monte Carlo localization via geometric sensing and terrain classification," *Autonomous Robots*, vol. 45, no. 6, pp. 843-857, 2021.

[14] K. Sabale, S. Mini, "Localization in wireless sensor networks with mobile anchor node path planning mechanism," *Information Sciences*, vol. 579, no. 1, pp. 648-666, 2021.
[15] A. Sharma, P. K. Singh, "Localization in wireless sensor networks for accurate event detection," *International Journal of Healthcare Information Systems and Informatics*, vol. 16, no. 3, 74-88, 2021.
[16] J. Tang, J. Han, "An improved received signal strength indicator positioning algorithm based on weighted centroid and adaptive threshold selection," *AEJ-Alexandria Engineering Journal*, vol. 60, no. 4, pp. 3915-3920, 2021.
[17] A. Toky, R. P. Singh, S. Das, "A localization scheme for underwater acoustic wireless sensor networks using AoA," *Recent Advances in Computer Science and Communications*, vol. 14, no. 3, pp. 690-699, 2021.
[18] G. S. Walia, P. Singh, M. Singh, M. Abouhawwash, A. K. Pandit, "Three dimensional optimum node localization in dynamic wireless sensor networks," *Computers, Materials and Continua*, vol. 70, no. 1, pp. 305-321, 2021.
[19] X. Wang, M. Cheng, J. Eaton, C. Hsieh, S. F. Wu, "Fake node attacks on graph convolutional networks," *Journal of Computational and Cognitive Engineering*, vol. 1, no. 4, pp. 165-173, 2022.
[20] X. Yao, M. Zhang, S. Hao, X. Yao, M. Zhang, "Coordinated sensing coverage optimisation in sensor networks using RSSI," *International Journal of Ad Hoc and Ubiquitous Computing*, vol. 36, no. 4, pp. 222-229, 2021.
[21] W. Ying, S. Sun, "An improved Monte Carlo localization using optimized iterative closest point for mobile robots," *Cognitive Computation and Systems*, vol. 4, no. 1, pp. 20-30, 2022.
[22] Y. Yu, Y. Li, Y. Liu, H. Yu, "Inertial optimization MCL deep mine localization algorithm based on grey prediction and artificial bee colony," *Wireless Networks*, vol. 27, no. 4, pp. 3053-3072, 2021.
[23] Y. Yu, Y. Yuan, "Hybrid adaptive extended state observer for event-triggered networked nonlinear systems subject to deception attacks and external disturbances," *International Journal of Robust and Nonlinear Control*, vol. 33, no. 5, pp. 3358-3375, 2022.

Biography

Lijun Mao (1978-), female, Han ethnicity, graduated from Xi'an University of Electronic Science and Technology in 2001 with a Bachelor's degree in Computer Application. In 2012, she graduated from Northwest Polytechnical University with a Master's degree in Computer Application Technology. In December 2023, she was awarded the title of Professor. Currently, serving as a computer teacher at Xi'an Peihua Univercity, She has led or participated in more than 10 provincial and ministerial level projects, 2 horizontal projects, and completed 2 scientific and technological achievement transformations. Published over 20 academic papers, including 10 core and above papers, published 5 undergraduate textbooks, authorized 2 national invention patents, 8 utility model patents, 7 software copyrights, guided 5 national and provincial-level innovation and entrepreneurship projects, and guided students to participate in more than 30 provincial-level and above competitions and won awards. Main research directions: data visualization, genetic algorithms, image processing, sensor applications, etc.

Yitong Zhang (2004-), female, Han ethnicity, graduated from Xi'an University of Electronic Science and Technology Affiliated Middle School in 2022. Currently, she is studying Broadcasting and Television Directing at the School of Journalism and Communication, Northwest University of Political Science and Law.

Application of NAWL-ILSTM Algorithm in Network Security Situation Awareness Prediction

Jun Ma

(Corresponding author: Jun Ma)

School of Information Engineering, Changsha Medical University
Changsha 410219, China
Email: cymajun0019@126.com

(Received July 11, 2023; Revised and Accepted June 14, 2024; First Online Aug. 16, 2024)

Abstract

The high-speed advancement of the Internet has increased the risk of network attacks and brought considerable challenges to network security. The existing network security measures, such as firewalls and virus-killing technologies, are insufficient to prevent network attacks effectively. Therefore, it is necessary to establish a network security situational awareness prediction model. According to the improved Adaptive Moment Estimation algorithm, this paper optimizes the online update mechanism of the Long Short-term Memory network, updates the parameters in real-time, and improves the model's accuracy. The Look-ahead algorithm is introduced to lift the network model's convergence rate, reduce the training cost, and deduct the prediction error. The established model was verified through experiments. The original Long short-term memory network, Support Vector Machines algorithm and Radial Basis Function were used as the comparison models. Comparative experiments have shown that the error between the predicted values of the designed model and the actual values is minimal. Compared with the support vector machine and Radial basis function, the average absolute percentage errors are 0.0198, 0.0523, and 0.0225, respectively; The Standard errors are 0.0126, 0.0326, and 0.0157. Network security situational awareness prediction accuracy is as high as 95.343%. Therefore, the proposed optimized model has particular perception and prediction capabilities for network attacks, and its development potential and reference value are worth exploring.

Keywords: Forward-Looking Algorithm; Long and Short-Term Memory; Network Security; Situation Prediction

1 Introduction

Network Security Situation Prediction (NSSP) is a major research hotspot in the past few years. It can predict the future situation and its trend built on existing network security data, providing instruction for relevant administrators to choose security policies. For the technology of network situational awareness prediction, deep learning technology is the primary choice. Among them, Long Short-Term Memory (LSTM) is a type of time cycle network, which could settle the issue of long-term dependence of RNN.

LSTM has a particular design structure that is befitting for processing and forecasting essential events with very long-intervals and delays in time series. It solves RNN problems by introducing memory units. From the current state of network security management, responding to network attacks with speed is essential in network security. Therefore, some studies suggest introducing Nadam optimization algorithm and Look-ahead method when training network models. Look-ahead is an optimizer algorithm.

By selecting the search direction through the "fast weight" sequence generated by another optimizer in advance, the training cost can be reduced and the Rate of convergence can be improved. To accurately perceive and predict the NSS and improve the efficiency of administrators' early warning response to network attacks, a NSS Perception Prediction (NSSPP) model based on improved Nadam and LSTM, as well as fusion of Look-ahead, has been proposed. The structure of the paper consists of four parts.

The first part elaborates on the background of network security under the development of computer internet, as well as the research purpose and significance of NSSPP.

The second part specifically explains the process and innovation of optimizing LSTM networks based on the improved Nadam algorithm and optimizing network models based on the Look ahead method; A new type of NSSPP has been constructed.

The third part fully elaborates on the experimental design based on NSSPP and the quantitative statistics and analysis of experimental results.

The fourth part describes the experimental conclusions, the shortcomings of this design, and the directions that need further in-depth research and exploration.

2 Related Works

The rapid emergence of computer networks has also greatly promoted the advancement of many deep learning technologies. Different scholars have proposed their own methods and strategies for perceptual prediction of NSS. For network security defense based on neural network (NN), Zhang R proposed a skill optimization algorithm and NSSP algorithm of "back promotion neural network" (BPNN) optimized by Simulated Annealing Skill Optimization Algorithm (SA-SOA). It uses the SOA to find the best adaptive individuals, obtain the best weights and thresholds, and assign them to the random initial thresholds and weights of the BPNN. Finally, the BPNN is trained to obtain the predicted values. According to SA's Metropolis criterion, this algorithm accepts bad solutions with a certain probability, avoiding the trap of falling into local optima and improving the algorithm's global search ability [19].

To more effectively detect network attacks, Li and Zhang use Honeypot technology to gather the newest network attack data, and combines Deep Neural Network (DNN) and LSTM model to propose a deep learning based network intrusion detection classification model [9]. Wei raised a new method with the combination of 3DCNN and Bidirectional RNN (Bi-RNN). Due to the fact that NSSP data includes multidimensional time series, combining spatial and sequence features can better predict NSS and improve prediction accuracy. Therefore, it adopts the 3DCNN to extract spatial features from distinctive network nodes, and uses the Bi-RNN with gated recursive units to extract sequence features. Ultimately, using the fused spatial sequence features, the prediction results of NSS are obtained [18].

Gupta *et al.* proposed a one-on-one technology system LIO-IDS based on LSTM classifier and optimized Network Intrusion Detection System (NIDS) to handle frequent and infrequent net-intrusions. LIO-IDS is a 2-layer anomaly-based NIDS that detects different network intrusions. Layer-1 uses LSTM classifier to identify intrusion from normal network; Layer2 uses integrated algorithms to sort out the detected intrusions into various attack categories. Gupta N also puts forward an Improved OnevsOne technology (I-OVO) for carrying out multi-class classification in layer2. Compared to original OVO, I-OVO only uses 3 classifiers to measure each sample, significantly deducting testing time. In addition, layer 2 uses Oversampling technology to enhance the LIO-IDS's detection capability [6]. Scholars such as Liu and Zeng have applied wavelet fuzzy neural networks (FNN) and chaotic particle swarm optimization algorithms (CPSO) to monitor the security status of IoT networks. Firstly, they analyzed the basic theory of intelligent city IoT NSS, constructed corresponding mathematical models, and designed an IoT security situational awareness framework. Secondly, they studied the basic theory of FNN, designed the FNN structure, and constructed a wavelet based NSS data processing method. Finally, the training process of FNN based on CPSO was established [12].

In the NSS evaluation method, Zhu and Du proposed an NSS evaluation mode in accordance with time-varying evidence theory to address the lack of consideration for the time-varying support rate of threat occurrence in existing multi-source information fusion technologies. By introducing time parameters into the basic possibility assignment, the threat information reflects the temporal variation pattern. Therefore, they improved the existing hierarchical threat situation quantitative assessment technology and raised a layered multi-source network security threat situation assessment model [21]. Liao *et al.* designed a NSS evaluation system according to the extended Hidden Markov model (eHMM). Firstly, he extended the standard HMM from 5 to 7 tuples, adding 2 parameters: network defense efficiency and risk loss vector. Then, he defined the initial algorithm of the State-transition matrix, and used the observation vector to create and modify the network matrix. The solving process of hidden status possibility distribution sequence grounded on eHMM is derived. Finally, he designed a method to calculate the risk loss vector based on international definitions, and calculated the current network risk value through implicit probability distribution; Then an assessment was conducted on the global security situation [10]. Lin and his team applied deep learning to analyze and discuss network details, classify and produce counter networks for sample-amplification. They take sparse noise reduction automatic encoder for selecting features, and next adopt LSTM to predict security situation [11].

In summary, NSS is divided into two aspects: situational awareness and situational prediction. Scholars use NN for security detection and defense around these two aspects. However, scholars rarely improve and optimize NN when using deep learning methods. In today's rapidly developing era, it is necessary to design a more in-depth and improved NN to establish a relatively complete network situational awareness prediction model.

3 LSTM Optimized Based on Nadam Built on Look-ahead

There are two ways to improve LSTM: one is to optimize the online update mechanism of LSTM based on the improved Nadam algorithm for situation prediction; Second, the adaptive momentum estimation algorithm of Nestervo acceleration gradient is improved by AWL's Look-ahead to rise the Rate of convergence. Therefore, an NSSP model based on ILSTM combined with NAWL is proposed.

3.1 Situation Prediction and Network Security Awareness Model on the Basis of ILSTM

LSTM is a multi-layer NN with output, hidden, and input layers, capable of processing variable length sequences and

generating new output sequences. It can also handle long-distance dependencies, reduce the matter of gradient explosion/disappearance, and learn the degree of long-term storage or forgetting between multiple time steps. RNN is a chain structure related to time series, and its hidden unit expansion diagram is Figure 1. Its three layers are closely related, and the output at a certain moment is related to both the input time and the previous output. The previously obtained information can be propagated backwards [5, 8, 14].

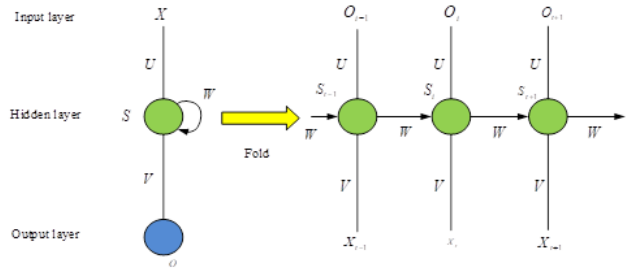


Figure 1: Hide unit expansion diagram

In Figure 1, S represents the hidden layer, O represents the output layer, U, V, W represents the weight, and represents the number of layers. LSTM could settle the above matters and lift the prediction accuracy due to its ability to introduce memory units, which can delete and memorize new information at any time. LSTM also has a chain structure, with only one Tanh layer in the hidden layer of RNN. LSTM has four interaction layers and three gates of forgetting, output, and input, each responsible for controlling the state of each unit [2]. Figure 2 shows its structure.

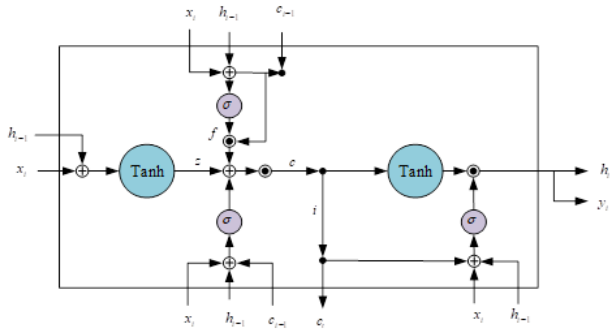


Figure 2: LSTM Cellular structure

In Figure 2, c_{t-1} represents memory cells, h_{t-1} hidden states, and state information transmitted from the previous time. σ represents the fully connected layer, $Tanh$ represents activation function, c represents candidate memory cells, x_t represents output, f represents forgetting gate, z represents output gate, and i represents input gate. In real life, the network has always been under attack. The firewall and Intrusion Detection system (IDS) can collect the information records of the attack at any time. To receive real-time network situation and optimize relevant network parameters, an Improved LSTM (ILSTM) was designed to construct an immediate and effective NSSP [15].

In reality, networks are always in a state of being attacked, so in order to optimize network parameters using real-time received network situation values, an improved LSTM network is proposed to establish an effective network situation prediction model. The improved LSTM network is an LSTM learning online update mechanism that minimizes the cost function. Based on the practical problems of network systems, a more effective ILSTM situation prediction model is proposed to update network parameters using the situation time series data transmitted online, in order to establish a more effective ILSTM situation prediction model with real-time observation of network situation data. This paper utilizes the minimization cost function to improve and optimize LSTM, making it a learning online update mechanism, thus establishing the ILSTM model. The basic process of the model specifically includes making reasonable use of existing historical records, using the observed data of the later sampling time as the true value, and adding both prediction and actual errors to the whole sample error. The model's parameters are perfected using error-minimization. With the increasing real-time updates of experimental data, the updating of model parameters becomes more proficient, and the predicted values obtained from the trained model will be closer to the actual values [13, 20]. The relevant update formula is Equation (1).

$$\begin{cases} f_t = \sigma(W_f \times [h_{t-1}, x_t] + b_f) \\ u_t = \sigma(W_t \times [h_{t-1}, x_t] + b_i) \\ \widetilde{C_t} = \tanh(W_c \times [h_{t-1}, x_t] + b_c) \\ C_t = f_t \times C_{t-1} + i_t \times \widetilde{C_t} \\ o_t = \sigma(W_o \\ times[h_{t-1}, x_t] + b_o) \\ h_t = o_t \times \tanh(C_t) \end{cases} \tag{1}$$

In Equation (1), f_t represents the forgetting gate result, i_t represents the input gate input, W_f, W_i, W_c, W_o represents the weight matrix, $\widetilde{C_t}$ represents the activation function result, C_t represents the candidate result, o_t represents the output, and h_t represents the hidden layer result. In actual net-security maintenance, the operation of it is always affected by certain external contributors, resulting in the constantly changing online situation and becoming more complex. This increases the difficulty of network situation prediction. The perception model of NSS is established accordingly. The model quantitatively analyzes the NSS situation during a certain time period, and conducts weight analysis on relevant influencing factors [17], thereby establishing and generating an NSS time series diagram. Figure 3 shows the model structure of NSS.

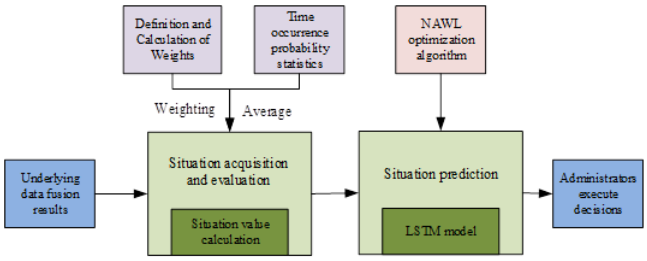


Figure 3: Structure of NSS Awareness Model

3.2 NSSP Method Based on ILSTM Combined with NAWL

Nadam is an optimization algorithm that combines Niche Genetic Algorithm (NGA) with Adam. Adam combines AdaGrad and RMSProp algorithm, and has the characteristics of random optimization of Adaptive learning rate. The weights of ILSTM are divided into four matrices $[W_f, W_i, W_c, W_o]$, with the sizes of $V \times H$, $H \times H$, $V \times H$, and $C \times H$. Firstly, connect each layer and input it into the NAWL optimization algorithm, while training and adjusting weights. The matrix weights that can be trained from this are $(2 \times V + H + C)$ in total, because $V(10k - 1M)$ has a higher order of magnitude than $H(128 - 4k)$ and $C(100 - 1k)$, and the total number of weights can be determined to be $\geq 10^8$.

To improve the Rate of convergence of Nadam and reduce the running cost, on this basis, AWL uses Look ahead to improve Adam in the Nadam algorithm, and improve the problem that the Nadam algorithm needs a lot of space when updating the first order. Introduce an equivalent formula and hyperparameter to control the strength of Look-ahead during the improvement process [4]. The relevant process formulas are Equations (2) - (6). $\theta = [W_f, W_i, W_c, W_o]$ represents the optimized parameter vector, and $f(x)$ represents the cost function.

$$\hat{g}_t \longleftarrow \frac{g_t}{1 - \prod_{i=1}^{t} \mu_u} \tag{2}$$

Equation (2) is the gradient for obtaining prospective parameters.

$$m_t \longleftarrow \mu_t \cdot m_{t-1} + (1 - \mu_t) \cdot g_t \tag{3}$$

m_t in Equation (3) is the updated first-order distance estimation. μ_t represents the Exponential decay rate of distance estimation, ranging from 0 to 1.

$$n_t \longleftarrow v \cdot n_{t-1} + (1 - v) \cdot g_t^2 \tag{4}$$

n_t in Equation (4) is the updated second-order distance estimation. v represents the same as μ_t.

$$z_t \longleftarrow \frac{\bar{m}_t}{\sqrt{\hat{n}_t} + \epsilon} \tag{5}$$

z_t in Equation (5) is the updated intermediate variable.

$$\Delta\theta \longleftarrow -\eta[(1 + \gamma)z_t - \gamma z_{t-1}] \tag{6}$$

$\Delta\theta$ in equation (6) is the vector for updating parameters. η is the learning rate. Figure 4 shows the LSTM process optimized and improved by Nadam With Look Ahead (NAWL). ϵ usually takes the value 10^8 to avoid dividing by 0 and $v = 0.999$. $\mu_0 = 0.99$ is the Exponential decay rate of the distance estimate. $\gamma = 0.9$ is a newly introduced hyperparameter [3] to control prospective intensity.

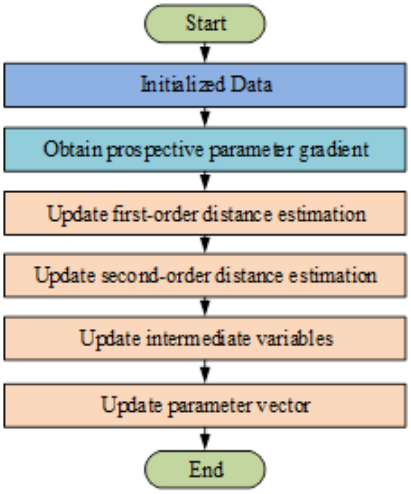


Figure 4: The flow of LSTM algorithm

The NAWL algorithm updates the weight parameters with the weight $[W_f, W_i, W_c, W_o]$ of ILSTM. It combines the strong Rate of convergence of Nadam and the advantages of AWL's Look ahead. It has a good improvement effect on Adam and Nadam's Rate of convergence, which is significantly improved compared with Adam and RMSProp [1]. Figure 5 is the graph of Look-ahead.

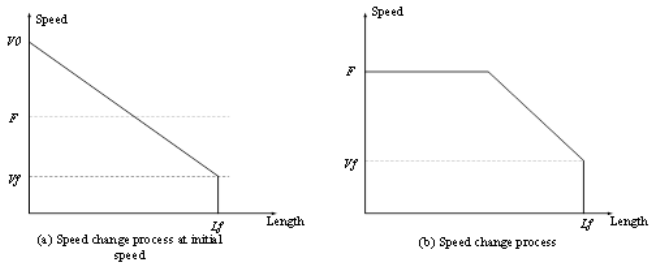


Figure 5: Implementation of Look ahead

Calculate the beginning and ending speeds of each block, in accordance with the rest length of the continuous blocks, taking the machining of a circular route as an example, as Equation (7).

$$V_0 = \sqrt{V_f^2 + 2 \cdot A \cdot L} \tag{7}$$

In Equation (7), V_0 is the feasible speed. A represents the max-allowable acceleration of the machine tool. V_f is the

next block's max-feasible entry-speed for. L represents the current block length. If the inlet feed rate V_0 is greater than the feed rate F of the command, the feasible inlet feed rate of the present block becomes F and the ending feed rate becomes V_f.

To prove the convergence property of the NAWL algorithm, the corresponding proof process was carried out. To prove that "the Nadam algorithm has asymptotic convergence, then the NAWL algorithm also has asymptotic convergence" [16]. Assuming the learning rate is stationary, the parameters for each NAWL update are Equation (8).

$$\Delta\theta = -\eta[(1+\gamma)z_t - \gamma z_{t-1}] \tag{8}$$

After updating the parameters of the NAWL algorithm, the following results are obtained, as Equation (9).

$$\lim_{t\to\infty}(\sum_{t=0}^{n} -\eta[(1+\gamma)\cdot z_t - \gamma\cdot z_{t-1})] = \lim_{t\to\infty}(\gamma\cdot z_n - \gamma\cdot z_0) + \sum_{t=0}^{n}\eta\cdot z_t \tag{9}$$

It is known that the parameter update formula of — Nadam has asymptotic convergence, that is, Nadam has asymptotic convergence, as Equation (10).

$$z_t = \frac{\bar{m_t}}{\sqrt{\hat{n_t}}+\epsilon} \tag{10}$$

Equation (10) has asymptotic convergence. From Equation (9), it can be concluded that if $\lim_{t\to\infty}(\gamma\cdot z_n - \gamma\cdot z_0)$ asymptotically converges, then NAWL is asymptotically convergent. $z_0 = 0$ and η, γ are constant constants. When training through the Nadam algorithm, z_n approaches 0, as Equation (11).

$$\lim_{t\to\infty}(\gamma\cdot z_n - \gamma\cdot z_0) = \lim_{t\to\infty}\eta\cdot\gamma\cdot z_n \approx 0 \tag{11}$$

Therefore, Equation (9) is asymptotically convergent, meaning that NAWL has asymptotic convergence. The NSSP model is established through the above mentioned methods. The online update mechanism is mainly used to improve the network parameter update of LSTM, and the look ahead method of AWL is used to improve the optimization algorithm of Nadam to improve the Rate of convergence. Figure 6 is the specific NSSP model based on ILSTM combined with NAWL.

First of all, the 1D time-series $X = (x_1, x_2, \cdots, x_n)$ is extended to a 2D matrix:

$$\begin{bmatrix} x_1 & x_2 & \cdots & x_{n-k+1} \\ x_2 & x_3 & \cdots & x_{n-k+2} \\ \vdots & \vdots & \cdots & \vdots \\ x_k & x_{k+1} & \cdots & x_n \end{bmatrix}$$

Where n is the time-series length. k represents the sample numbers. The sample is $y = (x_k, x_{k+1}, \cdots, x_n)$. Standardize the time series, as Equation (12).

$$X = \frac{x_i}{\sqrt{x_i^2 + x_{i+1}^2 + \cdots + x_{i-k+1}^2}}(i = 1, 2, \cdots, n-k+1) \tag{12}$$

Then initializing the network parameters and introduce and set hyperparameters to obtain Equation (13).

$$\begin{cases} W_f = rand(L, N) \\ b_f = rand(1, N) \\ \cdots \\ Max\_iter = M_1 \\ Error\_Cost = M_2 \end{cases} \tag{13}$$

In Equation (13), M_1 represents the max-iterations Max_iter. M_2 is the error threshold $error\_Cost$. L is the quantity of cell units in LSTM. N is the amount of neuron layers. W_f means forgetting the gate weight. b_f is offset. The cell unit status information that needs to be forgotten is Equation (14).

$$\hat{f}_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \star C_{t-1} \tag{14}$$

Calculating the output for the forgetting gate and multiplying the result with the previous unit state. Calculate the information that is able to be saved in the cell unit state at time, as Equation (15).

$$\hat{u_t} = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \star \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{15}$$

Equation (14) is divided into two parts. One is the output of input gate i_t; The second is to establish a new candidate vector C_t through the tanh function. Then multiply the candidate vector by the i_t, and calculate the C_t as Equation (16).

$$C_t = \hat{u_t} + \hat{f_t} \tag{16}$$

The net's output at is Equation (17).

$$h_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \star \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{17}$$

Calculate the output gate O_t, and then multiply the result by the current unit state to obtain the network output currently [7]. h_t means the predicted value at present. After calculating the predicted value for each sample, calculating the error between all h and the true value y, as Equation (18).

$$J_{(\theta)}(y, h; W, b) = \frac{1}{2}||y - h||^2 \tag{18}$$

Using BPTT to update net-parameters in reverse, the iterations increase till the max-error threshold or max-iteration $error > Error\_Cost$ or $iter > Max\_iter$ is reached, and then exit the loop. Enter the weight matrix $\theta_0 = [W_f, W_i, W_c, W_o]$ to be updated and train the ILSTM network model parameters using NAWL, as Equation (19).

$$\theta_t = \theta_{t-1} - \eta[(1+\gamma)z_t - \gamma z_{t-1}] \tag{19}$$

According to the method of real-time updating online observation data, new samples $X_{n+1}(x_{n-k+2}, \cdots, x_{n+1})$

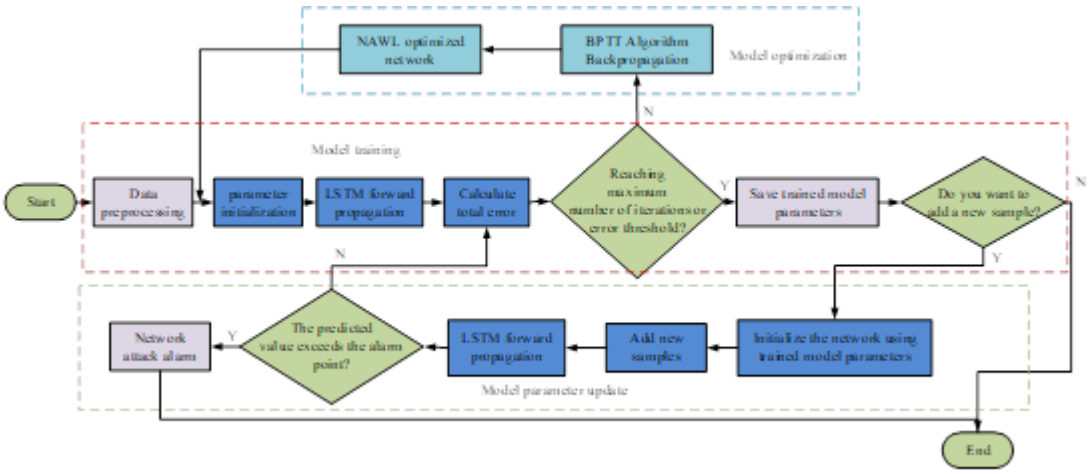


Figure 6: A NSSP Model Based on ILSTM and NAWL

and θ_0 are added to carry put forward propagation of several calculation steps from the forgotten cell unit state information to the net-output at t. For obtaining the h_{n+1} of the new sample, as Equation (20).

$$error = error + \frac{1}{2}(h_{n+1} - x_{n+2})^2 \tag{20}$$

Therefore, when the h_t at the next sampling time completes the value of the network being attacked, a warning is issued, and the net-administrator can quickly take measures.

4 NSSPP Experimental Validation

This study proposes a prediction method based on Nadam combined with Look-ahead method to improve LSTM. Fusion and innovation research uses Look-ahead to solve the problem of slow Rate of convergence of Nadam, and integrates online update mechanism into LSTM parameter update. The NSSP proposed here has been experimentally validated for its performance. By comparing the original LSTM, SVM, and Radial Basis Function NN (RBF) models, the experimental data has been analyzed to draw conclusions.

4.1 Experimental Design of Situation Awareness Prediction Model

This experiment concentrates on validate the NAWL-ILSTM prediction method's effectiveness. The experimental data used is from the historical records of network attacks collected by a certain network company's firewall and other systems for a total of 95 days in July, August, and September. Collect log records once a day, using the first three quarters (71 days) as the training dataset for the model, and the last quarter (24 days) as the testing dataset for the model. Set experimental parameters n in the experiment_ Input=28, n_ Steps=128, n_ Hidden=10, batch_ Size=128. For lifting the training speed of the constructed method, the original data is standardized. Figure 7 shows the standardized NSS time series.

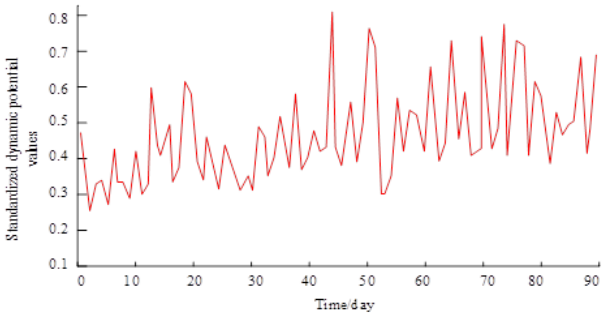


Figure 7: Standardized NSS Time Series

In the experiment, the selected comparative models were the original LSTM, SVM, and RBF models. They can more intuitively analyze the accuracy and effectiveness of designing models for predicting NSS. SVM's basic model is the Linear classifier with the biggest interval denoted in the feature space. RBF is a frequently-used 3-layer feedforward network that can be utilized for both function approximation and pattern classification. Before the experiment, it is necessary to set the parameters of several models, such as input, output, and hid-layer node numbers. Unified parameters for comparative research in experiments, see Table 1.

The data obtained in the experiment needs to be truthfully and effectively recorded, and reasonable statistics and analysis should be conducted. The analysis process uses Mean Absolute Percentage Error (MAPE), Standard error (SDE), and Mean squared error (MSE) to analyze the accuracy of the prediction model. The iterations and Rate of convergence of the model algorithm are measured

Table 1: Model's Parameter settings

Model	Input layer	Quantity of hidden layers	Amount of hidden-layer nodes	Output layer
SVM	5	1	11	1
RBF	5	1	0	1
LSTM	5	1	0	1
NAWL-ILSTM	5	1	0	1

to judge the feasibility of the model and the practicability of the algorithm.

5 Quantitative Statistics and Analysis of Experimental Results

As shown in Table 2, in order to verify the advantages and feasibility of the research algorithm, the original LSTM, RBF, and SVM models were selected, and compared with the new algorithm SOA_Compare and study the BP neural network algorithm. It can be seen that compared to the latest algorithm, the algorithm proposed in the study is generally closer to the actual value. The experiment obtained relevant experimental results and data, and calculated MAPE and SDE respectively. Table 2 shows some test results data from the test samples. Compared with RBF and SVM models, the model algorithm designed in this study has MAPE of 0.0198, 0.0523, 0.0225, and SDE of 0.0126, 0.0326, and 0.0157, respectively. The RBF model has the largest prediction error, while NAWL-ILSTM has a relatively small prediction accuracy and the highest prediction accuracy. So the prediction model designed this time has feasibility and accuracy in predicting network situational awareness.

Figure 8 shows the trend of MSE for the four algorithms over iteration. Nadam combines the properties of Look ahead and Adam, and uses the gradient descent feature of NGA algorithm in the implementation of Adam optimization algorithm to update the weight of the network, which rises the Rate of convergence. Use the Look-ahead method to improve the shortage of high cost in Nadam, so as to improve the convergence Rate of algorithm. The Rate of convergence of network training using different algorithms varies, among which the number of iterations required by NAWL is at least 60, and that of Nadam without improvement and optimization is more than 80. Adam and RMSProp have not yet reached complete convergence after more than 100 iterations.

To verify the advantages and feasibility of this algorithm, Figure 9 selected the original LSTM, RBF, and SVM models for comparison. All four methods have well-established mechanisms for processing data, but there are still some shortcomings in terms of predictive performance. The proposed prediction model can update the model parameters online, process and predict the data in real time. Compared with other methods, the error with the actual value is smaller, and the Mean absolute error and SDE are 0.0254, 0.0189. Therefore, the improved method can effectively update data in immediate online to rise the precision of prediction.

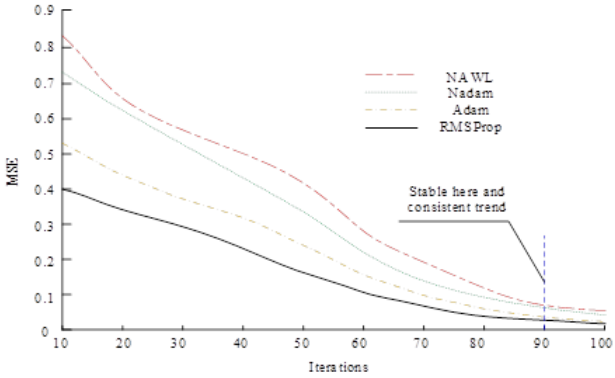


Figure 8: The change trend of Mean squared error of four algorithms with the number of iterations

From Figure 10, in September and October, there are significant differences between the two scenarios. In Figure 10(a), when the network is attacked, the situation value is in an unstable state with irregular fluctuations, and as time increases, the value continues to rise and fluctuates more and more frequently. In Figure 10(b), there is a pattern of horizontal fluctuations in the network without network attacks. This indicates that the overall risk for the 30 nodes in September was relatively low, while the curve for the 31 nodes in October showed an increasing trend. It shows that the network condition level is more secure and less risky in September, and in October, the network tends to be moderate or even very dangerous.

To more intuitively display the performance of several models on NSSP, Figure 11 extracts experimental data from 10 iterations for accuracy comparison. The accuracy of the NAWL-ILSTM model is the smallest compared to the error of 1. The error between SVM and 1 can reach over 0.4, while the error of RBF is relatively small at around 0.2. So the NSSPP designed this time has good predictive ability, almost identical to the actual value, and can provide warning for network administrators.

Table 2: Partial test result data in the test sample

Test sample	Actual value	SOA_BP	SVM	RBF	NAWL-ILSTM
1	0.5876	0.5701	0.5808	0.5538	0.5696
2	0.6501	0.6648	0.6586	0.6275	0.6746
3	0.5263	0.5499	0.5213	0.5691	0.5456
4	0.4917	0.4695	0.5246	0.5274	0.4777
5	0.5205	0.5514	0.5272	0.5543	0.5456
6	0.5546	0.5648	0.5266	0.5698	0.5296
7	0.5447	0.5604	0.5535	0.6025	0.5595
8	0.616	0.6248	0.6059	0.6368	0.6037
9	0.4962	0.4947	0.5222	0.4919	0.5124
10	0.6571	0.6741	0.6479	0.6172	0.6725

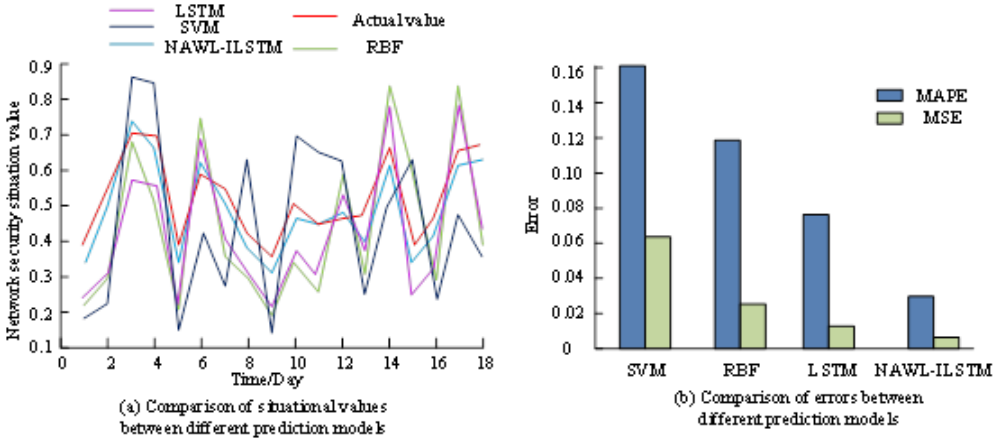


Figure 9: Comparison with the original LSTM, RBF, and SVM models

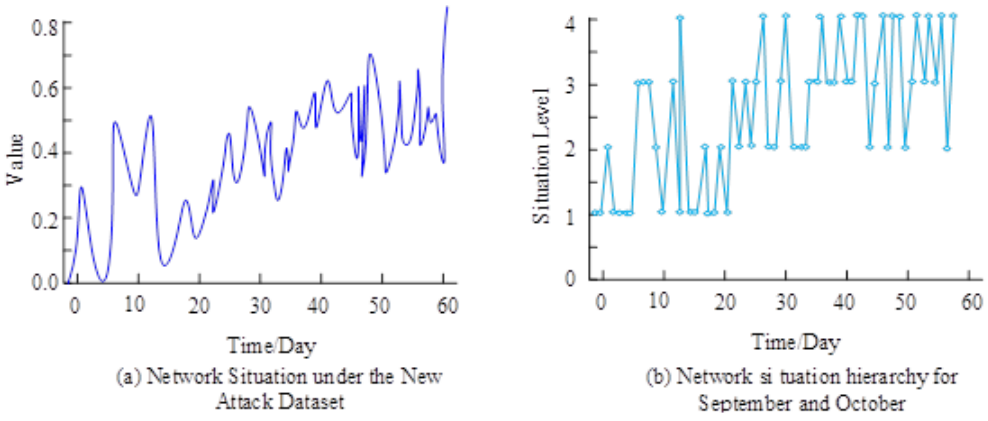


Figure 10: The situation value of the network under attack and the situation value of not receiving an attack

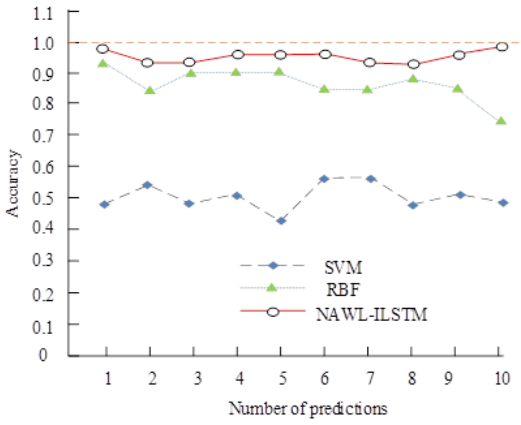


Figure 11: Several Models for Predicting Network Situation

6 Conclusion

Nowadays, NSSP gives a more integrated and viable new approach to address the shortcomings of distinctive network attack solutions, and is now popular in the network security. This manuscript come up with an optimized LSTM-NSSPP on the ground of Nadam combined with Look ahead method. It combines the advantages of LSTM time series with the online update mechanism, and combines the Look-ahead algorithm to improve the convergence Rate of the algorithm and decrease the training cost. The results of testing the model indicate that by comparing the original LSTM, RBF, and SVM, the error among the predicted values of the designed mode and the actual values is minimal; Compared with RBF and SVM models, MAPE is 0.0198, 0.0523, 0.0225, and SDE is 0.0126, 0.0326, and 0.0157, respectively, with an average accuracy of over 94%. It can respond accurately and quickly to network attacks for early warning operations. However, the model used in this study is only for predicting the perception of network situation for a certain system, and cannot comprehensively predict multiple systems. It is difficult to accurately predict situation values for nonlinear and complex time series, and more experiments are needed to adjust model parameters and further research and design.

References

[1] M. S. Al-Daweri, S. Abdullah, K. A. Z. Ariffin, "A homogeneous ensemble based dynamic artificial neural network for solving the intrusion detection problem,"

[2] H. N. Bhor, M. Kalla, "TRUST-based features for detecting the intruders in the internet of things network using deep learning," *Computational Intelligence*, vol. 38, no. 2, pp. 438-462, 2022.

[3] Z. Chen, "Research on internet security situation awareness prediction technology based on improved RBF neural network algorithm," *Journal of Computational and Cognitive Engineering*, vol. 1, no. 3, pp. 103-108, 2022.

[4] Z. Dai, "Research on network intrusion detection security based on improved extreme learning algorithms and neural network algorithms," *International Journal of Biometrics*, vol. 12, no. 1, pp. 56-66, 2020.

[5] P. Dixit, "Deep learning algorithms for cybersecurity applications: A technological and status review," *Computer Science Review*, vol. 39, no. 4, pp. 100317.1-10037.15, 2020.

[6] N. Gupta, V. Jindal, P. Bedi, "LIO-IDS: Handling class imbalance using LSTM and improved one-vs-one technique in intrusion detection system," *Computer Networks*, vol. 192, no. 4, pp. 108076.1-108076.19, 2021.

[7] R. Hou, G. Ren, W. Gao, L. Liu, "Research on cyberspace multi-objective security algorithm and decision mechanism of energy internet," *Future Generation Computer Systems*, vol. 120, no. 10, pp. 119-124, 2021.

[8] I. Jemal, M. A. Haddar, O. Cheikhrouhou, A. Mahfounhi, "Performance evaluation of convolutional neural network for web security," *Computer Communications*, vol. 175, no. 47, pp. 58-67, 2021.

[9] X. Li, S. Zhang, "Network intrusion detection methods based on deep learning," *Recent Patents on Engineering*, vol. 15, no. 4, pp. e210421180688.1-e210421180688.9, 2021.

[10] Y. Liao, G. Zhao, J. Wang, S. Li, "Network security situation assessment model based on extended hidden markov," *Mathematical Problems in Engineering*, vol. 3, pp. 1428056.1-1428056.13, 2020.

[11] Z. Lin, J. Yu, S. Liu, "The prediction of network security situation based on deep learning method," *International Journal of Information and Computer Security*, vol. 15, no. 4, pp. 386-399, 2021.

[12] Q. Liu, M. Zeng, "Network security situation detection of internet of things for smart city based on fuzzy neural network," *International Journal of Reasoning-based Intelligent Systems*, vol. 12, no. 3, pp. 222-227, 2020.

[13] J. Ren, "Network security situation assessment model based on information quality control," *International Journal of Performability Engineering*, vol. 16, no. 4, pp. 673-680, 2020.

[14] K. V. Samarthrao, V. M. Rohokale, "Enhancement of email spam detection using improved deep learning algorithms for cyber security," *Journal of Computer Security*, vol. 30, no. 2, pp. 231-264, 2022.

[15] I. Sohn, "Deep belief network based intrusion detection techniques: A survey," *Expert Systems with Applications*, vol. 167, pp. 114170.1-114170.9, 2020.

[16] I. S. Thaseen, J. S. Banu, K. Lavanya, M. R. Dhalib, K. Abhishek, "An integrated intrusion detection system using correlation-based attribute selection and

artificial neural network," *Transactions on Emerging Telecommunications Technologies*, vol. 32, no. 2, pp. 10.e4014.1-10.e4014.15, 2020.

[17] H. Wang, D. Zhao, X. Li, "Research on network security situation assessment and forecasting technology," *Journal of Web Engineering*, vol. 19, no. 7/8, pp. 1239-1265, 2020. *International Journal of Critical Infrastructure Protection*, vol. 34, pp. 100449.1-100449.23, 2021.

[18] L. M. Wei, "Network security situation prediction based on combining 3D-CNNs and Bi-GRUs," *International Journal of Performability Engineering*, vol. 16, no. 12, pp. 1875-1887, 2020.

[19] R. Zhang, M. Liu, Y. Yin, Q. Zhang, Z. Cai, "Prediction algorithm for network security situation based on BP neural network optimized by SA-SOA," *International Journal of Performability Engineering*, vol. 16, no. 8, pp. 1171-1182, 2020.

[20] Y. Zhao, G. Cheng, Y. Duan, Z. Gu, Y. Zhou, L. Tang, "Secure IoT edge: Threat situation awareness based on network traffic," *Computer Networks*, vol. 201, pp. 108525.1-108525.15, 2021.

[21] Y. Zhu, Z. Du, "Research on the key technologies of network security-oriented situation prediction," *Scientific Programming*, vol. 2021, no. Pt.3, pp. 5527746.1-5527746.10, 2021.

Biography

Jun Ma, Associate Professor, holds a Master's degree in Software Engineering from Hunan University in 2018. software designer, Dean of the School of Information Engineering at Changsha Medical University, Director of the Information Science Research Center, and a member of the Chinese Computer Society. The editor in chief and co editor of 9 sets of textbooks, holds 2 utility model patents, holds 5 software copyrights, and has published over 20 papers. Working in school for more than 20 years, mainly engaged in computer course teaching and school information construction related work.

Research on Encryption Protection of Patients' Electronic Medical Privacy Data from the Legal Perspective

Feng Wang and Fei Yang
(Corresponding author: Feng Wang)

Tibet Police College
No. 68, Duodi Road, Chengguan District, Lasa City, The Tibet Autonomous Region 850000, China
Email: wf_wfeng@hotmail.com

(Received July 14, 2023; Revised and Accepted July 28, 2024; First Online Aug. 17, 2024)

Abstract

The medical privacy of patients is protected by law. This paper briefly introduces the legal basis of medical privacy protection and the attribute-based encryption algorithm used for encrypting medical privacy data. After that, simulation experiments were conducted to test the encryption efficiency, encryption effectiveness, and security of the algorithm. It was found that the number of identity attributes contained in private data could affect the encryption efficiency of the algorithm. The more attributes, the lower the encryption efficiency. When the number of overlapping identity attributes between the enquirer and the ciphertext met the threshold value, the plaintext could be obtained smoothly; otherwise, it could not be decrypted. The ciphertext obtained by the encryption algorithm could effectively resist brute force cracking.

Keywords: Attribute; Encryption; Legal Perspective; Medical Privacy

1 Introduction

With the rapid development of information technology, various industries have gradually stepped into the field of informatization, including the medical industry [12]. In the medical industry, in the past, patients' medical data were usually recorded by paper combined with local hospital databases, and different medical institutions are independent of each other. Although the leakage caused by data circulation was avoided to a certain extent, it also increased the difficulty of data query, which was not conducive to make more appropriate diagnosis schemes [15].

With the deepening of informatization, not only patients' medical data can be stored electronically in a larger database, but also data interoperability between different medical institutions can be realized through the Internet, which greatly promotes the sharing of resources among medical institutions and provides better services for patients [4]. However, the patient's electronic medical data contains the patient's diagnosis information, treatment records, physiological indicators, and other personal privacy information, and there is a possibility of disclosure on the Internet. In order to improve its security, usually the data will be encrypted. From the legal point of view, the encryption of patients' electronic medical privacy data is not only related to the protection of patients' privacy rights and interests, but also involves the information security of the medical industry and the compliance with laws and regulations. In order to ensure the safe transmission of medical sensor information, Khan *et al.* [2] adopted two encryption methods: substitution-ceaser cipher and improved Elliptical curve cryptography and verified their effectiveness through experiments.

Naeemabadi *et al.* [10] used chaotic sequence to change the encryption key to improve encryption security. The experimental results verified that this method has the advantages of low noise sensitivity and fast encryption speed. Doss *et al.* [6] adopted an evolutionary algorithm, that is, meme algorithm, to encrypt medical information, and verified the advantages of this algorithm through experiments. This paper briefly introduces the legal basis of medical privacy protection and the attribute-based encryption algorithm used to encrypt medical privacy data. Identity attributes are involved in data encryption to achieve the effect that the same ciphertext can only be decrypted by users who meet the conditions of identity attributes. Then simulation experiments are carried out.

2 Medical Privacy Protection from the Legal Perspective

As a natural person, patients have the right to privacy, but because of their disease and need to receive treatment, the abnormal performance caused by the disease and treatment process is not in ordinary social groups,

with particularity, so there is the right to privacy of patients. The right to privacy includes the right to privacy of patients. Compared with the conventional right to privacy, the right to privacy of patients only exists in the doctor-patient relationship, and the contents of the privacy are special, which may cause spiritual and reputational losses to patients once leaked. In addition, once patients seek medical treatment, private information such as their condition and treatment process must be shared with the medical institution (at least the attending doctor) and recorded in the institution's database. Patients have little control over such information, especially in the era of information technology [14]. Therefore, it is necessary to have relevant systems, laws and professional ethics to restrain medical institutions and protect patients' privacy.

The right to privacy is protected by law. The Civil Code, the Cybersecurity Law, the Personal Information Protection Law, and the Tort Liability Law all regulate the responsibilities and obligations of medical institutions and medical personnel in confidentiality of patients' private information. The Regulations on Industry Administration strengthen the privacy protection provisions for all types of information subjects in accordance with the needs of different industries and organizations.

On the whole, there are quite a lot of legislative provisions on the protection of the right to privacy and the right to privacy of patients, which provides a reliable legal basis for the protection of patient privacy and also stipulates the responsibility and obligation for the protection of patient privacy for medical institutions and personnel.

3 Encryption of Electronic Medical Privacy Data

This paper adopts the attribute-based encryption algorithm to encrypt medical privacy data [1]. The basic principle of the encryption scheme is shown in Figure 1. The public key and private key are given by the authority after combining the medical data attribute set (used to screen the attribute set of queriable users) provided by the medical data provider [5]. The medical data provider uses the public key to encrypt and store medical data. The enquirer downloads the ciphertext from the database and then decrypts the ciphertext using the private key given by the authority [13]. The specific steps are described below.

1) The authority inputs the security parameter λ and uses the algorithm of $G(1^{\lambda}$ to generate a group with a descriptive information of $D = (p, G, G_T, e)$. G and G_T are groups with an order of p [9], and there is a mapping relationship between them:

$$e : G \times G \to G_T$$

The generating element of G is g.

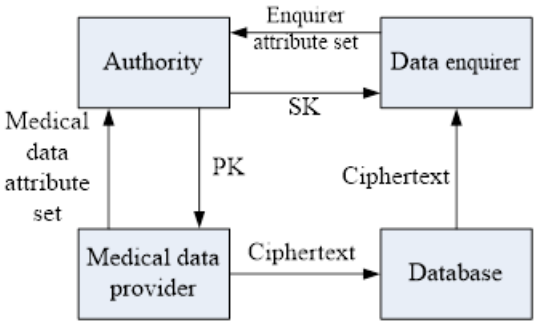


Figure 1: Principle construction of the attribute-based encryption scheme

2) The authority calculates the master key and public parameter through the generating element of G, and the equations are:

$$\begin{cases} MSK = (y, t_1, t_2, \cdots, t_n) \\ PK = (T_1 = g^{t_1}, T_2 = g^{t_2}, \cdots, \\ \qquad T_n = g^{t_n}, Y = e(g,g)^y) \end{cases} \tag{1}$$

where MSK is the master key, PK is the public parameter, y and t_i are random parameters that belongs to Z_p, Z_p is the set of integers ranging from 0 to $p-1$, and t_i is the random parameter representing the i-th attribute [8].

3) The authority calculates the private key by combining the medical data attribute set given by the medical data provider. The medical data attribute set refers to the set of identity attributes that have the permission to query the medical data. The formula for calculating the private key is:

$$\begin{cases} SK = \{D_i = g^{p(i)/t_i}\}_{i \in S} \\ p(x) = y + a_1 x + a_2 x^2 + \cdots + a_{d-1} x^{d-1} \end{cases} \tag{2}$$

where SK is the private key given to the data enquirer, S is the set of permission identity attributes that the data enquirer has, D_i is the private key of the i-th attribute (all the private keys of the attributes belonging to S are combined into SK), d is the threshold value, $p(x)$ is a $(d-1)$-degree polynomial function [7], and $a_1, a_2, \cdots, a_{d-1}$ are random numbers in the rational number field [3].

4) The medical data provider encrypts the medical data using PK issued by the authority. The encryption formula is:

$$CT = (S', M = mY^s, \{E_i = T_i^s\}_{i \in S'}) \tag{3}$$

where CT is the ciphertext, m is the plaintext, S' is an identity attribute that has permission to query medical data, and s is a random parameter belonging to Z_p. The encrypted data is uploaded to the database for storage.

5) After the data enquirer downloads the ciphertext from the database, whether the number of permission identity attributes that S coincide with S' is less than threshold value d. If it is less than, it cannot be decrypted; otherwise, d attributes are taken out from the overlapping attributes, and then $e(g,g)^{p(i)s} = e(E_i, D_i)$ is calculated, where i belongs to the d attributes that are taken out. The decryption formula is:

$$\begin{cases} m = \frac{CT}{Y^s} \\ Y^s = e(g,g)^{p(0)s} \end{cases} \tag{4}$$

where $e(g,g)^{p(0)s}$ is obtained by using the Lagrange interpolation method [11] combined with $e(g,g)^{p(i)s}$.

4 Simulation Experiment

4.1 Experimental Environment

The simulation experiment was carried out in laboratory servers. Five servers were set up in the experiment, server 1 as the authority, server 2 as the database, server 3 as the medical data provider, server 4 as the medical data enquirer, and server 5 as the third-party attacker. The configuration of the five servers was the same, all running on the Windows 11 operating system with 16 GB of memory and an i7 processor.

4.2 Experimental Setup

Since this article mainly aims to verify the effectiveness of attribute-based encryption algorithms in protecting medical privacy data, the focus is on the encryption performance for the data. Considering the difficulty in obtaining medical data and protecting patient privacy, random diagnostic data was generated and assigned with different identity attributes. Part of the data is shown in Table 1. Random generation of diagnostic data was a relatively simple way to process patients, and each processing method was endowed with 10-30 identity attributes.

1) Performance test of the attribute-based encryption algorithm:
When using the encryption algorithm proposed in this paper to encrypt the constructed medical data, in order to verify the impact of the number of identity attributes in the medical data on the performance of the encryption algorithm, diagnostic data with 10, 15, 20, 25, and 30 identity attributes were selected from the randomly generated data set, and the number of diagnostic data with different numbers of identity attributes was the same. The time taken by the encryption algorithm to generate key, encrypt data, and decrypt data under different numbers of identity attributes was tested.

2) Usage test of the attribute-based encryption algorithm:
The diagnostic data generated in server 3 was encrypted by the public key given by server 1 and then transmitted to server 2 for storage. Then, 50 enquirer IDs were generated in server 4, and each ID was assigned 15 identity attributes. After server 4 obtained the ciphertext from server 2, server 4 obtained the private key from server 1 using the enquirer ID and decrypts the ciphertext. The threshold value was set to 3.

3) Security test of the attribute-based encryption algorithm:
Server 5 acted as a third-party attacker to steal data from server 2, which stored medical data, and performed brute-force decryption on its ciphertext. The brute-force decryption lasted for 180 min. The integrity of decryption through brute force during this process was recorded.

4.3 Experimental Results

Diagnostic data with different numbers of identity attributes were constructed for encryption, and the key generation, encryption, and decryption time of the encryption algorithm are shown in Figure 2. As the number of identity attributes increased, the key generation, encryption, and decryption time of the encryption algorithm increased. By vertical comparison of the time under the same number of identity attributes, it can be seen that the encryption time was the longest, and the key generation time and decryption time were close. The plaintext and long key needed to be calculated during encryption. However, when decrypting, it only needed to use the Lagrange interpolation method to obtain a specific value and then divided the ciphertext to get the plaintext.

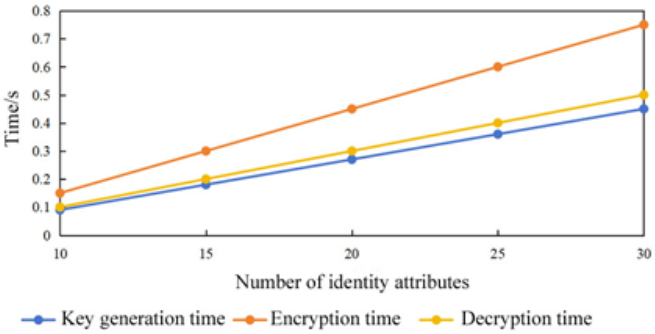


Figure 2: Time consuming of the encryption algorithm under different numbers of identity attributes

The enquirer in server 4 queried the encrypted data in server 2. Some of the query results are shown in Table 2. When the enquirer with the same ID decrypted different ciphertexts, the decryption failed because the number of identity attributes overlapping with the ciphertext did not reach the threshold (3 here), and the corresponding plaintext could not be obtained by the enquirer who reached

Table 1: Part of randomly generated diagnostic data

No.	Diagnostic data	Identity attribute
1	The patient needs to be admitted for computed tomography examination in the afternoon.	The Inpatient Department of XX Hospital; The Radiology Department of XX Hospital; The Internal Medicine Department of XX Hospital; patient's family members; ...
2	The patient has a mild cold and is prescribed two courses of cold medicine.	The Internal Medicine Department of XX Hospital; The Pharmacy of XX Hospital; Patient's family members:...
3	The patient is recovering well and is expected to be discharged in a week.	The Inpatient Department of XX Hospital; The Surgery Department of XX Hospital; Patient's family members; ...
...	...	...

the threshold.

The ciphertext stored in server 2 was brute-force cracked. The decryption integrity of ciphertext in this process is presented in Figure 3. With the increase of brute force cracking time, the decryption integrity of the ciphertext gradually increased, but the increase amplitude gradually decreased. After 100 min of brute force cracking, the decryption integrity was almost unchanged, and the decryption integrity was only 2.5%, which ensured the security of the ciphertext.

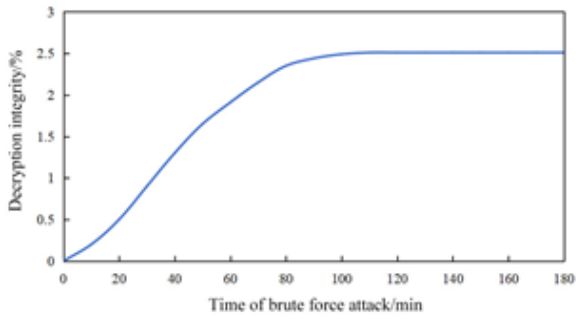


Figure 3: Decryption integrity of the ciphertext in the process of brute force cracking

5 Discussion

The development of information technology has prompted various industries, including the medical industry, to enter the field of informatization. With the help of information technology, the medical industry gradually realizes the work of less paper or even paperless, which greatly improves the efficiency. At the same time, after patients' information is stored on the Internet, information can be shared between different medical institutions and between different departments within medical institutions, which is conducive to providing more appropriate medical services for patients. However, it is the sharing of information Internet that puts patients' medical information at the risk of being easily leaked. Patients' medical information is private information, and compared with conventional private information, patients are special because they suffer from diseases and are prone to abnormal performance different from ordinary social groups in the course of treatment. Moreover, patients must interact with medical institutions during treatment, and their medical information must be shared with medical institutions. Therefore, as a special kind of private information, medical information needs to be protected by law. In addition to the relevant laws for the protection of routine privacy, corresponding laws and regulations for medical privacy information are also needed to require medical institutions to fulfill the responsibility and obligation of protecting patients' medical information.

On the other hand, the law also requires medical institutions to bear the responsibility of protecting medical information, so there is a need for encryption and protection of medical information. In the face of users with different identities, the traditional encryption algorithm will distribute the ciphertext encrypted by different keys. That is to say, in the face of the same plaintext, it needs to be independently encrypted for many times to obtain the complex ciphertext, which greatly increases the calculation amount. In this paper, an attribute-based encryption algorithm was adopted, which combines identity attributes when encrypting the plaintext, and the user must meet the conditions of identity attributes when decrypting the ciphertext. The algorithm does not need independent encryption, and only one ciphertext is needed to obtain multiple ciphertexts. Only users who meet the conditions of identity attributes can decrypt the ciphertext. Then, a simulation experiment was carried out to test the encryption efficiency, encryption effectiveness, and security of the encryption algorithm. It was found that the number of identity attributes involved in the encryption of

Table 2: Some of the query results obtained by the enquirer in server 4

The actual plaintext of the ciphertext	Number of identity attributes	Enquirer ID	Number of identity attributes overlapping with the ciphertext	Decryption result
The patient needs to be admitted for computed tomography examination in the afternoon.	15	102145	5	The patient needs to be admitted for computed tomography examination in the afternoon.
		104878	2	Unable to decrypt
		102359	6	The patient needs to be admitted for computed tomography examination in the afternoon.
		...	...	...
The patient has a mild cold and is prescribed two courses of cold medicine.	20	102145	2	Unable to decrypt
		104878	6	The patient has a mild cold and is prescribed two courses of cold medicine.
		102359	0	Unable to decrypt
		...	...	...

medical data could affect the efficiency of the algorithm. The reason is that identity attributes are required to participate in the generation process of the key. The more identity attributes, the longer the key will be, and the more time it takes to generate the key. Similarly, the calculation amount required for the encryption and decryption of the plaintext with a long key will also increase, resulting in an increase in the encryption and decryption time. The test results also verified that the algorithm could effectively block users who did not meet the conditions of identity attributes to query medical information, and moreover, it had considerable resistance to the brute force cracking of the third attacker.

6 Conclusions

This paper briefly introduces the legal basis of medical privacy protection and the attribute-based encryption algorithm used to encrypt medical privacy data. After that, simulation experiments were conducted to test the encryption efficiency, encryption effectiveness, and security of the algorithm. With the increase of the number of identity attributes of diagnostic data, the key generation, encryption, and decryption time of the encryption algorithm increased. Under the same number of identity attributes, the encryption time of the algorithm was the largest, and the key generation time and decryption time were close. When the number of identity attributes of the enquirer overlapping with those of the ciphertext reached or exceeded the threshold value, the plaintext was obtained smoothly; otherwise, it could not be decrypted. With the increase of brute force cracking time, the decryption integrity of the ciphertext gradually increased, but the increase amplitude gradually decreased. After 100 min of decryption, the decryption integrity was almost consistent, only 2.5%.

References

[1] B. Abdul, T. W. Tariq, S. Muhammad, M. Naeem, K. S. Ali, O. S. Al, "Reversible encryption and lossless data hiding for medical imaging aiding smart health care," *Cluster Computing*, vol. 26, no. 5, pp. 2977-2991, 2023.

[2] M. Ayoub, M. T. Quasim, N. S. Alghamdi, M. Y. Khan, "A Secure Framework for Authentication and Encryption Using Improved ECC for IoT-Based Medical Sensor Data," *IEEE Access*, vol. 8, pp. 52018-52027, 2020.

[3] A. M. Badr, Y. Zhang, H. G. A.Umar, "Dual Authentication-Based Encryption with a Delegation System to Protect Medical Data in Cloud Computing," *Electronics*, vol. 8, no. 2, pp. 1-15, 2019.

[4] M. D. Boomija, S. V. K. Raja, "Securing medical data by role-based user policy with partially homomorphic encryption in AWS cloud," *Soft Computing: A Fusion of Foundations, Methodologies and Applications*, vol. 27, no. 1, pp. 559-568, 2023.

[5] P. Diwan, R. Kashyap, B. Khandelwal, "Blockchain assisted encryption scheme for intellectual share estimation usingmedical research data," *Concurrency and Computation: Practice and Experience*, vol. 36, no. 2, pp. 1-19, 2024.

[6] S. Doss, J. Paranthaman, S. Gopalakrishnan, A. Duraisamy, S. Pal, B. Duraisamy, C. L. Van, D. N. Le, "Memetic Optimization with Cryptographic Encryption for Secure Medical Data Transmission in IoT-based Distributed Systems," *Computers, Materials and Continua*, vol. 66, no. 2, pp. 1578-1594, 2021.

[7] V. Jaikumar, D. K. Venkatachalapathy, "Secure Transfer of Medical data to the cloud using cyclic watermarking and Encryption Technique," *Solid State Technology*, vol. 63, no. 5, pp. 37-47, 2020.

[8] H. Li, Y. Yang, Y. Dai, S. Yu, Y. Xiang, "Achieving Secure and Efficient Dynamic Searchable Symmetric Encryption over Medical Cloud Data," *IEEE Transactions on Cloud Computing*, vol. 8. no. 2, pp. 484-494, 2020.

[9] K. S. R. Murthy, V. M. Manikandan, "A Reversible Data Hiding through Encryption Scheme for Medical Image Transmission Using AES Encryption with Key Scrambling," *Journal of Advances in Information Technology*, vol. 13, no. 5, pp. 433-440, 2022.

[10] M. Naeemabadi, A. M. Dehnavi, H. Rabbanni, K. Bahaadinbeigy, H. Khajehpour, "Statistical Analysis of One Time Pad Encryption Using Variable Chaotic Key for Medical Data Transmission & Storage," *Journal of Applied Sciences Research*, vol. 11, no. 4, pp. 10-29, 2015.

[11] T. Paka, S. Divya, "Data Storage Security and Privacy in Mobile Cloud Computing Using Hierarchical Attribute Based Encryption (HABE)," *International Journal of Computer Sciences and Engineering*, vol. 7, no. 6, pp. 750-754, 2019.

[12] A. K. Singh, B. Kumar, G. Singh, A. Mohan, "Securing Patient Data Through Multiple Watermarking and Selective Encryption," *Medical Image Watermarking*, vol.2017, pp. 195-225.

[13] M. Singh, N. Baranwal, K. N. Singh, A. K. Singh, H. Zhou, "Deep learning-based biometric image feature extraction for securing medical images through data hiding and joint encryption-compression," *Journal of Information Security and Applications*, vol. 79, no. Dec., pp. 1-11, 2023.

[14] X. Yan, S. Zhuo, Y. Wu, C. Bo, "Distributed Privacy-Preserving Fusion Estimation Using Homomorphic Encryption," *Journal of Beijing University of Technology*, vol. 31, no. 6, pp. 551-558, 2022.

[15] L. Zheng, Z. Wang, S. Tian, "Comparative study on electrocardiogram encryption using elliptic curves cryptography and data encryption standard for applications in Internet of medical things: NA," *Concurrency and Computation Practice and Experience*, vol. 34, no. 9, pp. 1-13, 2022.

Biography

Feng Wang, born in August 1983, graduated from Chinese People's Public Security University with a master's degree in June 2014. He is working at Tibet Police College as an associate professor. He is interested in procedural law.

Fei Yang, male, Han nationality, born in 1993, Bachelor degree, Bachelor of Engineering, graduated from Yunnan Police College Road traffic management engineering, main research direction is road traffic management, is now a teacher in Xizang Police College.

Network Traffic Intrusion Detection Based on FSP-VSTG-MTL Algorithm

Jianjun Wu

(Corresponding author: Jianjun Wu)

Kaifeng University, Kaifeng 475004, China

Email: henu88@126.com

(Received July 31, 2023; Revised and Accepted July 12, 2024; First Online Aug. 17, 2024)

Abstract

The study in focus creates a network threat data detection model. This model combines a traditional network detection model, classification algorithms, and multi-task grouping techniques. Test results reveal that this model outperforms other algorithms in classification performance. When the sample size was less than 150, the model surpassed others by around 4%, 9%, and 10% on average. For sample sizes above 200, it exceeded by roughly 2%, 4%, and 5%. This model also exceeds in robustness performance. At sample sizes below 150, it performed better by about 6%, 12%, 5%, and 8%. For sizes more than 200, it bettered by 4%, 5%, 6%, and 7%. Its flexibility stands higher, too. With sample sizes less than 150, it was around 5%, 10%, 6%, and 7% higher. For samples above 200, it scored 3%, 4%, 5%, and 6% higher. This demonstrates that flexible self-stepping learning and multi-task techniques incorporated in a network data stream threat detection model yield practical and feasible results. When paired with feature selection methods and optimization techniques, these techniques enhance the detection and classification of network traffic, making it a potent research object in network security.

Keywords: Invasion; Probe; Multitask; Self-step; Trait

1 Introduction

The birth of the Internet and today's Internet has brought qualitative leaps and improvements to people's lives, and the current explosive development of network science and technology has largely brought about the exponential growth of applications and vulnerabilities in network systems. Due to the depth of technological change and the continuous reorganization of the network has led to an increasing number of new and unexpected technical vulnerabilities, posing a great potential threat to the security of individual or collective users [12]. As of 2022, there are more than 20,000 new vulnerabilities and the number of new vulnerabilities is increasing every year, making network intrusion a huge and prosperous new black industry chain. Hackers even have specialized iterative update tools to conduct cyber intrusions and an ecosystem of support systems to serve the illegal behavior, as well as an infrastructure to facilitate more people to conduct cyber intrusions. Current cyber intrusions are becoming more diverse and more frequent, with new intrusion tools for cryptojacking and ransomware increasing in development rate by 75% and 42% respectively, and malware randomly appearing in more adaptive and intelligent variants as the environment changes and new intelligent technologies are applied. Hackers using these rapidly updated and innovative attack tools can not only make cyber intrusions easier, but also more sophisticated and covert [5]. Among these attacks on critical assets can cause serious financial losses to the owner, even endanger public health and safety or cause actual physical harm and ransom.Therefore, it is very necessary to monitor the network intrusion traffic. The traditional network intrusion traffic detection algorithms mainly include support vector machine, decision tree and neural network, etc., but there are problems of low recognition efficiency caused by large difference in data distribution. Under this background, a new learning algorithm model is proposed to improve the efficiency of attack data identification by creatively incorporating flexible self-stepping learning technology and multi-task learning technology.

The study explores and analyzes the technology from four aspects. The first part discusses and summarizes the current research and status related to the network data stream threat detection system. The second part analyzes the traditional network threat detection system and various common algorithms including support vector machine, decision tree and K-neighborhood algorithm, including the construction of a new network data threat detection system. In the third part, experimental validation and comparative data analysis of the network data flow threat detection model are conducted. The fourth part is a comprehensive overview of the whole paper and a reflective summary of the shortcomings.

2 Related Work

The popularity of network technology has led to the explosive growth of network intrusion tools and attacks, and the construction of detection models for monitoring malicious and abnormal traffic in network data streams has become an important area of research and exploration for some experts around the world. Almomani has proposed a hybrid intrusion detection model based on feature selection and bio-inspired elements, combined with various algorithms such as particle swarm optimization, to address the problem of network intrusion attacks that cannot be detected by firewalls, algorithms, proposed a hybrid intrusion detection model thus improving the efficiency of rare attack identification [1]. Pontes *et al.* proposed an energy-based stream classifier based on stream-based intrusion detection system combined with machine learning algorithms to improve the efficiency of real-time classification of network data streams [11]. Azizan *et al.* addressed the anomalous data in big data diagnosis problem, proposed a new network intrusion detection system based on the traditional intrusion detection system combined with various optimization-seeking algorithms, thus improving the efficiency of identifying anomalous traffic in big data [2]. Krishnaveni *et al.* proposed an integrated approach based on the traditional intrusion monitoring system combined with feature selection classification techniques for the possible network intrusion problem in the process of cloud computing, thus improving the efficiency of identifying anomalous traffic in big data, feature selection classifier to improve the efficiency of anomalous traffic identification in the process of cloud computing [8]. Zhang *et al.* proposed a multi-objective optimization algorithm based on balanced convergence and diversity based on machine learning algorithms for the security of smart cars, thus improving the security of smart cars in the process of networking [16]. Ghurab *et al.* addressed the problem of utilizing network intrusion datasets with a network intrusion detection system based on a combination of eight world famous datasets proposed an instance feature analysis method for datasets, thus effectively improving the learning training effect of network intrusion detection models [4].

At present, many multi-task learning algorithms are used in the field of network intrusion traffic detection. In anomaly detection, Wan B *et al.* proposed a multi-task deep neural network to detect with the help of global spatiotemporal context features, and the proposed method effectively improved the detection efficiency [13]. Dao and Lee proposed a probabilistic feature extraction classification network algorithm model based on stacked autoencoders based on automatic coding technology and feature extraction analysis technology to address the current problem of identifying network attack data, thus effectively improving the efficiency of identifying attack data [3]. Wu *et al.* proposed a machine learning detection model based on machine learning combined with sample. The machine learning detection model is proposed to improve the accuracy of network intrusion detection based on sample training and testing [15]. Lin *et al.* propose an IoT intrusion detection model based on cloud computing and machine learning combined with multidimensional feature extraction techniques and extreme learning algorithms to improve the identification efficiency of IoT intrusion data for the security problem of IoT [9]. Wang *et al.* propose an IoT intrusion detection model based on traditional detection systems to improve the identification efficiency of IoT intrusion data for the problem of network vulnerability to attacks. Wang *et al.* proposed an acoustic emission detection system based on traditional detection systems combined with boundary decision algorithms to improve the efficiency of identifying network attack traffic [14]. Hidayat *et al.* proposed a new threat detection model based on traditional machine learning combined with deep learning techniques and long and short-term memory to improve the efficiency of detecting network attacks for the attack data classification problem. Efficiency [7]. Guo *et al.* proposed a spam identification method based on machine learning combined with bi-directional encoder and pre-training methods for the useless email detection and classification problem in emails, thus improving the accuracy of email classification [6].

From the research of scholars in various countries, most of the current intrusion detection systems are less efficient in identifying and classifying data for multiple types of novel unknown network intrusion attacks, and the research of system identification algorithms for novel attacks is neglected. Therefore, the multi-task learning network threat detection model developed by combining multiple intelligent algorithms possesses a certain degree of innovation.

3 Internet Data Stream Threat Detection System Design and Implementation

Unlike traditional detection systems, detection models that use flexible self-stepping learning techniques as well as multi-task learning techniques based on feature selection and a series of optimization techniques are innovative, so the design and implementation of the model is particularly important to ensure that the detection accuracy of this data detection model can be continuously optimized for accuracy. Therefore, this section analyzes the implementation principle of the model, classification algorithm and system construction.

3.1 Network Threat Detection System and Algorithm Research

Network threat refers to the use of unauthorized methods through the network to arbitrarily obtain, tamper or damage the target computer system resources informa-

tion, which may lead to economic losses or privacy leaks, thus posing a threat to the safety of other people's personal property network malicious behavior. The network threat detection system can effectively monitor all the traffic data in the target computer network in real time, which helps to fix the possible vulnerability of the network before the threat arises, and the system can help administrators to find illegal network threats and other suspicious traffic faster and more accurately. Network threat detection system is called Network intrusion detection system, referred to as NIDS. Its process is shown in Figure 1.

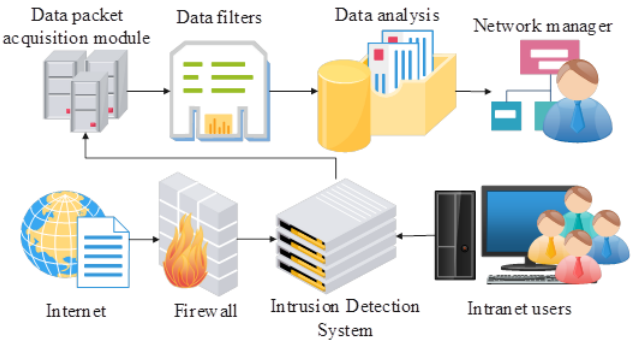


Figure 1: Flow chart of network intrusion detection system

As can be seen from Figure 1, the local network consists of three parts: the extranet, the intranet and the management and supervision network. When the network data flow enters the intranet, it will flow between each data unit, and these flows are subject to real-time capture and monitoring analysis from NIDS. If the data operation is normal, it will be recorded in the system log, if the data operation is illegal, it will promptly generate an exception report and remind the network administrator to intervene in time to restore the network security norm to ensure the safe operation of the local network. Currently, a variety of threat detection algorithms need to be analyzed for network threats, one of which is Support Vector Machine (SVM). The analysis of this algorithm is shown in Figure 2.

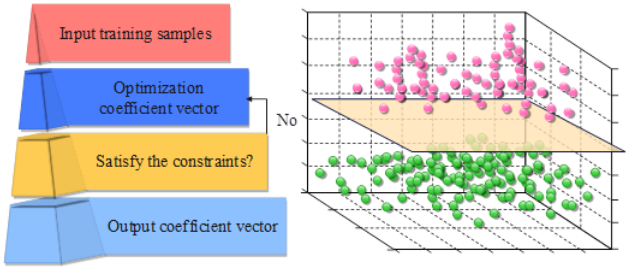


Figure 2: Parse diagram of SVM algorithm

As can be seen in Figure 2, the essence of the SVM algorithm is to find a reasonable classification hyperplane and a support vector for each data unit, i.e., a basis for classification, so that the data units are classified into two different categories according to the rules. The algorithm converts the original problem into a dyadic problem for solving to obtain the solution of the original problem, which requires the Lagrangian function as shown in Equation (1).

$$L(w,b,\alpha) = ||w||^2/2 + \sum_{i=1}^{m}(1 - y_i(w^T x_i + b)) \tag{1}$$

In Equation (1), L denotes the Lagrangian function, w denotes the normal vector, b denotes the dimensionality, denotes the support vector, α, y and x both denote the real vector, and T denotes the transpose. Most of the samples are not retained after training the SVM, making the final classification of the algorithm only related to the support vectors. The second algorithm is the Decision Tree algorithm (DT). Its flow is shown in Figure 3.

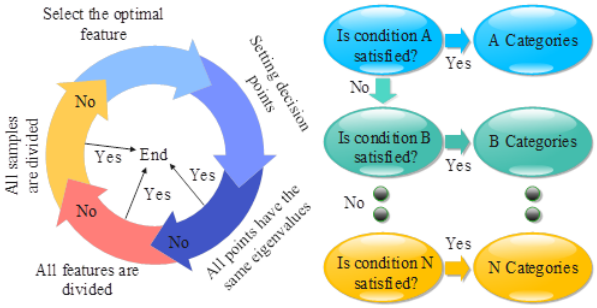


Figure 3: DT algorithm flow analysis diagram

As can be seen in Figure 3, the DT algorithm resembles the structure of a tree according to its name, and the algorithm analyzes and classifies the target data according to a set of classification knees. The DT algorithm consists of predefined variables as the target data and branching sub-nodes supported by decision rules, so it can quickly classify a large amount of complex data. The third algorithm that needs to be used is the K-nearest Neighbor (KNN) algorithm. Its mathematical expression is shown in Equation (2).

$$\begin{cases} Eucliden = \sqrt{\sum_{n=1}^{N}(x' - x'')^2} \\ Manhant\tan = \sum_{n=1}^{N}|x' - x^n| \\ Minkowski = (\sum_{n=1}^{N}|x' - x^n|^k)^{1/k} \end{cases} \tag{2}$$

In Equation (2), $Eucliden$ denotes the Euclidean function, N denotes the overall sample size of the data set, x denotes the sample, n denotes the random number, and k denotes the positive random integer. The KNN algorithm has strong noise immunity for training data with certain noise effects, but its computation requires long time memory usage, which leads to a huge increase in memory usage and wasted resources. The fourth algorithm is Adaptive Boosting (AB), which is a prediction algorithm that classifies data streams by calculating the

classification weights for each data sample. The weights are divided into weak classification weights and strong classification weights. The AB algorithm first obtains the weak classification weights of the data unit and then combines multiple weak classification weights to obtain a strong classification weight to improve the classification accuracy of the data unit. The fifth algorithm to be used is Balanced self-paced learning (BS). The mathematical expression is shown in Equation (3).

$$\min_{w,e} \sum_{n=1}^{N} e^n \eta(y^n, f(x^n, w)) + \beta R(w) - \sum_{k=1}^{n}\sum_{k=1}^{n} \lambda_k e^{nk} \quad (3)$$

In Equation (3), η denotes the loss function, k denotes the number of categories, β denotes the corresponding coefficients, l denotes the categories, R denotes the regularization function, w denotes the Gaussian noise, e denotes the natural constant, x and y both denote the random parameters, and λ denotes the control function. The algorithm can keep the complexity of different samples in line with the overall average by selecting from simple to complex samples. In addition, the algorithm requires the use of Synthetic Minority Oversampling Technique (SMOTE) to avoid overfitting problems [10]. Its mathematical expression for computing the minority samples is shown in Equation (4).

$$\begin{aligned} x^{new} &= Near(rand[1, N^+]) \\ &+rand \times (Near(rand[1, K]) - Near(rand[1, N^+])) \end{aligned} \quad (4)$$

In Equation (4), x^{new} expresses the new minority sample, $Near$ expresses the proximity space sample, $rand$ expresses the random value in a certain interval, N^+ expresses the minority training sample, and K expresses the proximity parameter. The detection system constructed using the above classification algorithms as a basis for detecting and classifying real data has a high accuracy for common common attack data.

3.2 Multi-way Learning Detection Model Design and Implementation

The detection model constructed by traditional algorithms can accurately classify most of the network data streams, but it is difficult to further learn the classification for the more hidden and unconventional intrusions. Therefore, it is necessary to make full use of the learning data of each terminal computer for effective information sharing to maximize the overall learning efficiency and recognition accuracy of the model. The model introduces a multi-task learning mechanism, whose general mathematical expression is shown in Equation (5).

$$w = \min_{[w^1,\cdots,w^T]} \sum_{t=1}^{\tau} \zeta(y^t, x^t, w^t) + \lambda\Omega(W) \quad (5)$$

In Equation (5), w denotes the coefficient vector, x denotes the input matrix, y denotes the output vector, τ denotes the total number of samples, t denotes the task number, ζ denotes the cost function, λ denotes the regularization coefficient, Ω denotes the regularization function, and W denotes the coefficient matrix. Where the mathematical expression of regularization is shown in Equation (6).

$$\begin{cases} L1 = H(p,q) + \lambda \sum_{i=1} |\beta_i| \\ L2 = H(p,q) + \lambda \sum_{i=1} \beta_i^2 \end{cases} \quad (6)$$

In Equation (6), $L1$ expresses the regularization to minimize the absolute value of the weights, $L2$ expresses the regularization to minimize the squared weights, and β expresses the weighting coefficients. $L1$ Overfitting is avoided by reducing the density of weights to simplify the data, and overfitting is avoided by decaying the weights at $L2$, so the combination of the two can be used to obtain better regularization results. Multi-task training and learning requires grouping to improve the efficiency of learning and training, which requires the introduction of Variable Selection and Task Grouping for Multi-Task Learning (VSTML) based on the low-rank principle and feature selection. The mathematical expression of the rank minimization of the objective function is shown in Equation (7).

$$\min_{w} \sum_{t=1}^{\tau} \zeta(y^t, x^t, w^t) + \lambda||W|| \quad (7)$$

In Equation (7), w denotes the coefficient vector, x denotes the input matrix, y denotes the output vector, τ denotes the total number of samples, t denotes the task number, ζ denotes the cost function, λ denotes the regularization coefficients, and W denotes the coefficient matrix. The product of two matrices with low rank structure can yield the coefficient matrix, so that the new spatial features have the feature data of the original space to promote higher shared learning efficiency among multiple tasks with low rank constraint. The expression of the objective function of the VSTML algorithm can be derived from Equation (7) as shown in Equation (8).

$$\begin{aligned} \min_{w} \sum_{t=1}^{\tau} \zeta(y^t, x^t, UV^t)/N^t + \lambda_1||U||_1 \\ +\lambda_2||U||_{1,\infty} + \mu \sum_{t=1}^{\tau} (||V^t||_k^{sp})^2 \end{aligned} \quad (8)$$

In Equation (8), U denotes the potential feature matrix, V denotes the potential task matrix, λ_1, λ_2 and μ denote the regularization parameters, t denotes the data columns, and $||\cdot||_k^{sp}$ denotes the support parametrization of the random coefficients k. VSTML has shared information through each grouping task, but there is still the problem of task inequality, so it is necessary to select new tasks that are simpler according to the difficulty of each task itself, so the objective function is improved as shown in Equation (9) is shown.

$$\min_{\{w^1,w^2,\cdots,w^\tau\}} \sum_{t=1}^{\tau} e^t \zeta(y^t, x^t, w^t) + \lambda||w^t - w^0||_2^2 + \lambda||e||_1 \quad (9)$$

In Equation (9), e denotes the complexity value, which enables the training learning of the algorithm model to

start with more basic task datasets and later gradually perform new tasks with increasing difficulty coefficients in a stepwise manner to build the information sharing base. Further introduction of Multi-task Network Threat Detection Model (MNTD) is required based on the grouping task algorithm, and the detection process of this model is shown in Figure 4.

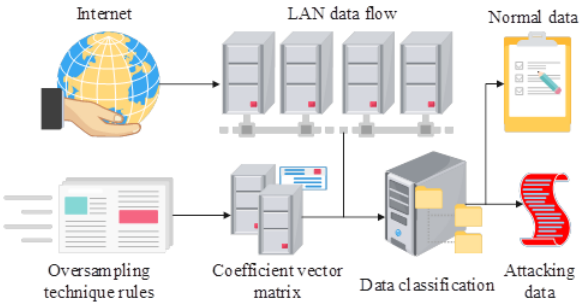


Figure 4: Flowchart of the multi-task network threat detection model

As can be seen from Figure 4, the model is mainly divided into three parts, the first part mainly contains the data terminals of the Internet and LAN; the second part mainly learns and trains the model through a certain amount of data streams to achieve the effect of model initialization; the third part mainly learns and trains the model to detect and classify the new data streams and the process of response. In order to increase the flexibility of the model it is necessary to further introduce the Flexible Self-Paced Multi-task Learning Techniques (SL), whose algorithm flow is shown in Figure 5.

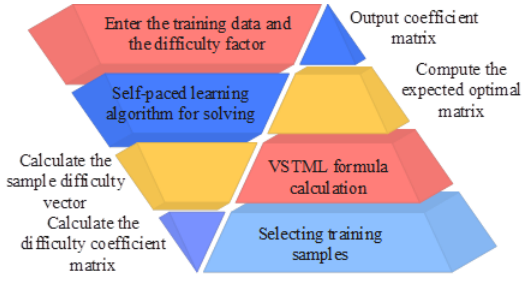


Figure 5: Flowchart of the multi-task algorithm for self-paced learning

As can be seen in Figure 5, the multi-task algorithm introducing the SL technique calculates and divides the complexity of each sample in the dataset, so that the complexity of the network data stream can be predicted in advance to improve the model efficiency. The objective function of the complexity of samples in a task is shown in Equation (10).

$$\min_{e^t}(e^t,\theta^t,x^t,y^t) = \min_{\theta^t,e^t}\sum_{n=1}^{N^t} e_1^{t,n}\zeta(y^{t,n},g(x^{t,n},\theta^t)) + |Omega(e^t,\theta^t) \quad (10)$$

In Equation (10), θ denotes the auxiliary vector, x and y both denote the input training data set, n denotes the sample number, $g(\cdot)$ denotes the binary classification function, l denotes the step size, and N denotes the total number of samples. Combining the previously mentioned VSTML algorithm and SL techniques, the study proposes the Flexible Self-Paced Variable Selection and Task Grouping for Multi-Task Learning (FSP-VSTG-MTL) detection algorithm based on the feature selection model, and its solution flow is shown in Figure 6.

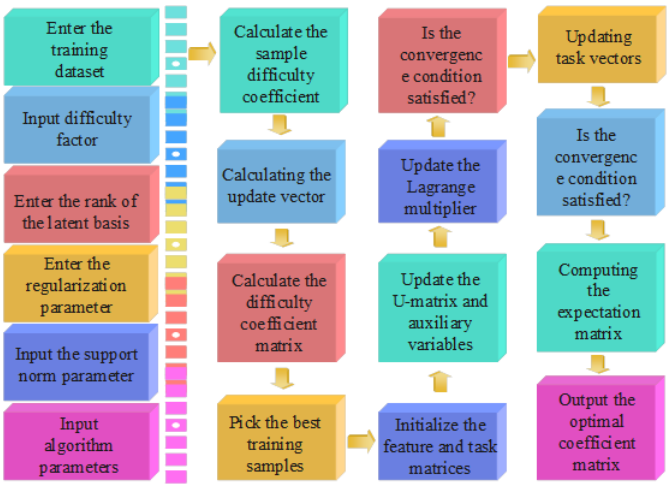


Figure 6: Flowchart of FSP-VSTG-MTL algorithm solution

As can be seen in Figure 6, the first stage of the algorithm obtains the objective function value of a task in the complexity matrix through a self-step learning technique to obtain the training samples for each assignment of the matrix during training. The second stage obtains the training samples used at different levels of complexity through a feature selection multi-task algorithm based on the obtained training samples, and iterates over them several times until they enter the convergence stage. The expression of the regular term in the algorithm is shown in Equation (11).

$$\min_{\theta^t,e^t}\sum_{n=1}^{N^t} e_1^{t,n}\zeta(y^{t,n},g(x^{t,n},\theta^t)) + \mu R(\theta_l^t) - \sum_{k=1}^{K}\sum_{nk}\lambda_k e_l^{t,nk} \quad (11)$$

In Equation (11), K denotes the training classification number of the sample, k denotes the serial number of the sample class, and CL denotes the training sample. The mathematical expression of the regularization term

is shown in Equation (12).

$$R(\theta_l^t) = \sum_{n=1}^{N^t} E_\zeta[(\theta_l^t)^T \bar{x}^{t,n}] - [(\theta_l^t)^T x^{t,n}] \tag{12}$$

In Equation (12), E_ζ denotes the expected distribution value, $\bar{x}$ denotes the noise sample, and $(\cdot)$ denotes the loss function, whose Taylor expansion is shown in Equation (13).

$$R^q(\theta_l^t) = \sum_{n=1}^{N^t} [(\theta_l^t)^T x^{t,n}] Var_\zeta[(\theta_l^t)^T \bar{x}^{t,n}]/2 \tag{13}$$

In Equation (13), q denotes the second proximity similarity value and Var_ζ denotes the variance value of the distribution expectation. The final expression of the objective function is shown in Equation (14).

$$\min_{U_l,V_l} \sum_{t=1}^{\tau} \zeta(y_l^t, x_l^t U_l V_l^t)/N^t + \lambda_1 ||U_l||_1 + \lambda_2 ||U_l||_{1,\infty} + \mu_2 \sum_{t=1}^{\tau} (||V^t||_k^{sp})^2 \tag{14}$$

In Equation (14), U denotes the potential feature matrix, V denotes the potential task matrix, and λ_1, λ_2 and μ denote the regularization parameters. The mathematical expression of the solved coefficient matrix is shown in Equation (15).

$$\begin{cases} W = U_l^* V_l^* = E[W_1, \cdots, W_L] \\ U_l^* = \arg_{U_l} \min \sum_{t=1}^{\tau} \zeta(y_l^t, x_l^t U_l V_l^t)/N^t + \lambda_1 ||U_l||_1 \\ \qquad\qquad + \lambda_2 ||U_l||_{1,\infty} \\ V_l^{t*} = \arg_{V_l} \min \sum_{t=1}^{\tau} \zeta(y_l^t, x_l^t U_l V_l^t)/N^t \\ \qquad\qquad + \mu_2 \sum_{t=1}^{\tau} (||V^t||_k^{sp})^2 \end{cases} \tag{15}$$

In Equation (15), W denotes the coefficient matrix. After convergence of the algorithm, the coefficient matrix is calculated and the final value of the coefficient matrix is obtained by calculating the expected value. The FSP-VSTG-MTL algorithm obtains the feature selection coefficient matrix of each data flow and stores it in the database, so that when a new data flow passes through the local network, the features of the data flow can be compared in real time for fast classification.

4 Model Validation and Data Analysis

To verify the practical effectiveness of the multitasking algorithm FSP-VSTG-MTL detection model, the experimental validation part of the study introduces four multitasking algorithm models, FSP-VSTG-MTL abbreviated as FV, VSTML abbreviated as VM, Self-paced Multi-task Learning Algorithm (SL) and Balanced Self-Paced Learning for Generative Adversarial Clustering Network (SG), comparing the criteria for introducing the F1 value (F-measure, F1), the number of feature dimensions (ND), and the number of tasks (F-measure, F1). dimensions (ND) and number of tasks (NT).

The laboratory uses a computer with operating system Windows, CPU 2.5 GHz and 6.0 GB of RAM. The study uses four simulation datasets with different D-dimensional features and T tasks. Dataset 1 contains 20 tasks with 25-dimensional training sample features per task; Dataset 2 contains 50 tasks with 20-dimensional training sample features per task; Dataset 3 contains 25 tasks with 40-dimensional training sample features per task; and Dataset 4 contains 50 tasks with 40-dimensional training sample features per task. Firstly, the classification performance of the algorithm is tested against the example shown in Figure 7.

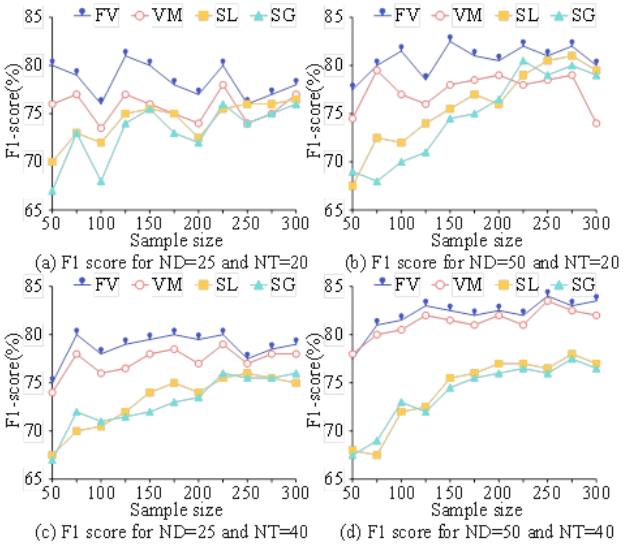


Figure 7: Comparison of F1 values under different parameters

The FV algorithm has a strong classification capability regardless of the number of feature dimensions and the number of tasks, and the VM performs better with four data parameters, with an overall performance of about 10% lower. The overall comparison shows that the multitask algorithm can better utilize the data information of different samples to promote the data sharing of the overall model to enhance the overall classification efficiency. Next, the robustness performance of the algorithms was compared and tested as shown in Figure 8.

As can be seen in Figure 8, all four algorithms show a stepwise decrease in their classification performance as the noise factor increases. When the noise level is low and the parameters are 25 and 20, respectively, the F1 values of the FV algorithm can be maintained above 85% on average, which is about 5% higher than the other multitasking algorithms and about 7% higher than the single task. When the noise and parameters increase respectively, the FV algorithm can still maintain an average above 80%, which is about 2% higher than the other multitasking al-

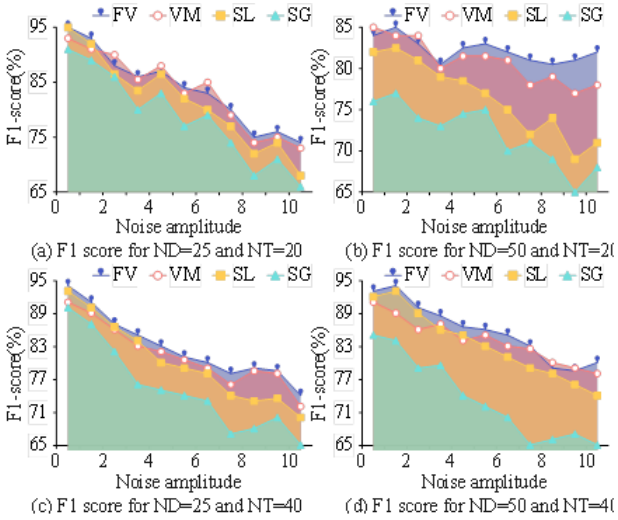


Figure 8: Comparison of robust performance test of the algorithm

gorithms and about 10% higher than the monotasking ones. Although the VM algorithm is close to the FV algorithm in various conditions, there is still an overall gap, and the experiments show that the robustness of the FV algorithm is better. The visualization of the third pair of complexity classification performance tests of the FV algorithm is shown in Figure 9.

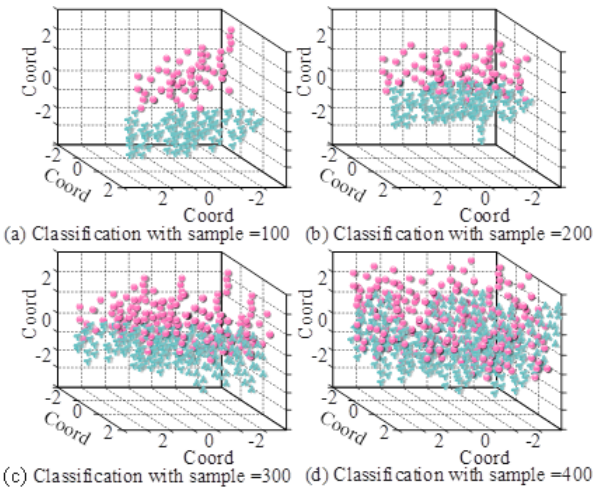


Figure 9: Visualization of the algorithm's complexity classification performance test

As can be seen in Figure 9, clear classification boundaries can be seen when the complexity of the samples is low, but as the complexity increases, the overlapping regions of their boundaries increase step by step indicating that the algorithms are not yet able to perform more effective and clear classification for overly complex data streams. Fourth, the experiments add four single-task algorithms, namely, SVM, DT, KNN, and AB, and eight algorithms are compared to identify nine types of attack data under three oversampling states, of which the three oversampling states are No oversampling technique (NOT), SMOTE, and Adaptive weighted oversampling technique (AB). semi-unsupervised weighted oversampling (ASW) are shown in Figure 10.

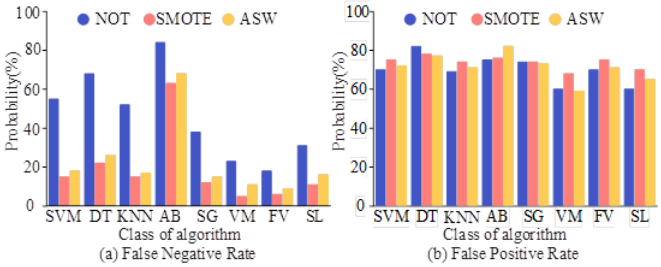


Figure 10: Comparison of false alarm rate and false alarm rate of the algorithm

As can be seen in Figure 10, the leakage rate without using the oversampling technique is about 30% and 25% higher than the other two states, respectively, and the false alarm rate is about 3% and 2% higher than the other two states. The leakage rate using SMOTE technique is about 5% lower and the false alarm rate is about 1% lower compared to the ASW technique. The eight attacks introduced in the experiment are Worm attack (W), Overflow attack (O), Scan attack (S), Detecting attack (D), Backdoor attack (B), Generic attack (G), and Vulnerability attack (G). Generic attack (G), Exploitation attack (E), Denial of Service attack (DS). The comparison criteria further introduce the False Negative Rate (FN) and the False Positive Rate (FP). A comparison of the F1 values for the experiments is shown in Table 1.

As can be seen in Table 1, the F1 values for the unused sampling technique are about 20% and 17% lower than the other two on average. The F1 values are about 3% higher using the SMOTE technique compared to the ASW technique. Finally the flexibility of the algorithm was tested as shown in Figure 11.

As can be seen in Figure 11, there is a gap between the F1 values of the FV algorithm and the SL algorithm in all four cases, with FV being on average about 5%, 8%, 6%, and 7% higher than SL. The gap is decreasing step by step as the number of samples increases, but the overall average gap is increasing step by step as the parameters increase. The experimental results indicate that the FSP-VSTG-MTL algorithm detection model has certain advantages in both classification performance, robustness performance and the final testing session for the data. The model proposed in the study is compared with the models in Literatures [7, 11, 15] and the experimental results are shown in Table 2.

From the data in the table, it can be seen that the FV algorithm outperforms other models in terms of precision, recall, accuracy and F1-score to capture the intrusion traffic accurately. In conclusion, the algorithm proposed in the study has better performance and can fulfil the need

Table 1: F1 values of the algorithm in different states

F1-score	Name	SVM	DT	KNN	AB	SG	VM	FV	SL
NOT	W	53.38		53.56		50.44	54.32	56.88	54.44
	DS	81.26	74.05	82.31	68.99	81.55	90.66	93.26	87.88
	O	56.71		55.71		54.22	60.55	60.99	60.13
	E	53.00		52.70		60.25	57.66	60.04	59.99
	B	53.33		52.88		50.01	54.11	56.99	56.44
	D	94.53	88.41	91.33	63.66	94.99	97.99	98.25	98.10
	S	87.40	81.75	86.00	80.55	91.55	97.00	98.00	95.41
	G	61.04	59.80	62.11	60.50	61.55	61.99	62.32	59.78
SMOTE	W	65.83	93.21	64.99	63.79	63.49	65.34	71.42	70.42
	DS	82.64	77.55	82.49	75.26	82.89	93.89	94.78	88.10
	O	67.00	64.89	66.99	56.74	63.49	65.46	66.49	68.12
	E	63.59	63.01	64.06	51.26	61.66	62.91	63.05	62.46
	B	63.13	62.89	62.33	58.99	61.47	64.46	65.87	65.45
	D	94.62	88.33	92.11	63.98	95.48	98.01	98.59	97.15
	S	87.34	80.16	86.79	79.99	90.89	98.02	97.89	95.11
	G	67.11	67.05	69.79	70.46	67.11	72.23	71.98	70.43
ASW	W	63.10	87.64	62.45	56.10	60.59	60.94	60.89	68.77
	DS	82.80	74.84	82.46	77.15	83.49	93.48	90.49	89.41
	O	66.89	63.15	64.15	56.45	62.15	65.15	64.16	68.94
	E	61.45	64.50	61.02		62.64	60.48	63.15	63.00
	B	60.78	61.78	60.46	54.91	60.01	61.00	60.99	62.45
	D	93.12	89.47	88.99	64.55	95.14	97.89	97.19	96.88
	S	85.20	80.04	80.15	80.59	91.22	97.48	96.49	95.13
	G	65.00	67.01	67.10	70.48	64.99	71.40	68.49	73.47

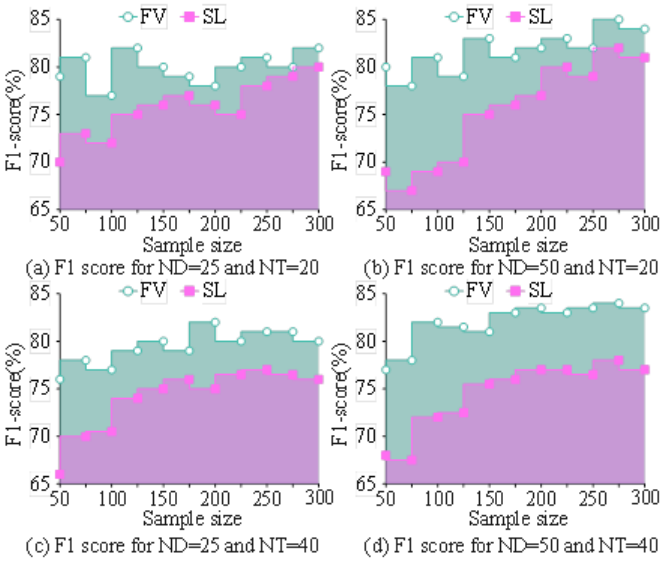


Figure 11: F1 values of the algorithm in different states

Table 2: Comparison of the performance of different models

Model	[11]	[15]	[7]	FV
F1-score	90.7	89.9	91.6	91.5
Precision	91.8	92.7	94.1	93.4
Recall	88.4	89.5	89.9	90.7
Accuracy	95.6	94.7	94.3	96.5

of detecting network intrusions.

5 Conclusion

For the network data stream threat detection classification response problem, the study combines the advantages and disadvantages of single-task algorithms as well as traditional multi-task algorithms to propose a self-stepping multi-task grouping algorithm model based on feature selection. Experiments are conducted to test and analyze the algorithm's classification type, robustness performance, as well as the identification performance and flexibility performance of the algorithm with different algorithms for attack data, and it is found that its classification performance is higher than other algorithms by

an average of about 4%, 9% and 10% when the number of samples is below 150, and about 2%, 4% and 5% when the number of samples is above 200, respectively. The robustness performance was found to be higher than other algorithms by about 6%, 12%, 5%, and 8% at sample sizes below 150, and about 4%, 5%, 6%, and 7% at sample sizes above 200 for the four state parameters. The underreporting rate is about 30% and 25% higher than the other two states, the false alarm rate is about 3% and 2% higher than the other two states, and the F1 value is about 20% and 3% higher than the other two states, respectively. The flexibility is about 5%, 10%, 6%, and 7% higher than other algorithms when the number of samples is below 150 in four states, and about 3%, 4%, 5%, and 6% higher than other algorithms when the number of samples is above 200.The research on FSP-VSTG-MTL algorithm for network traffic intrusion detection demonstrates its great potential in dealing with complex network security threats in the future. In addition to network security, the FSP-VSTG-MTL algorithm has potential applications in numerous fields. For example, in healthcare, it may help detect abnormal patterns in electronic medical records and protect patient privacy.The study also has certain shortcomings, the recognition efficiency of the model decreases step by step as the complexity of the network data stream increases, indicating that its fast and accurate recognition in large-scale complex environments needs to be further explored and optimised.

References

[1] O. Almomani, "A hybrid model using bio-inspired metaheuristic algorithms for network intrusion detection system," *Computers Materials and Continua*, vol. 68, no. 1, pp. 409-429, 2021.

[2] A. H. Azizan, S. A. Mostafa, A. Mustapha, C. F. M. Foozy, M. H. A. Wahab, M. A. Mohammed, B. A. Khalaf, "A machine learning approach for improving the performance of network intrusion detection systems," *Annals of Emerging Technologies in Computing*, vol. 5, no. 5, pp. 201-208, 2021.

[3] T. N. Dao, H. J. Lee, "Stacked autoencoder-based probabilistic feature extraction for on-device network intrusion detection," *IEEE Internet of Things Journal*, vol. 9, no. 16, pp. 14438-14451, 2021.

[4] M. Ghurab, G. Gaphari, F. Alshami, R. Alshamy, S. Othman, "A detailed analysis of benchmark datasets for network intrusion detection system," *Asian Journal of Research in Computer Science*, vol. 7, no. 6, pp. 14-33, 2021.

[5] H. Guo, J. Li, J. Liu, N. Tian, N. Kato, "A survey on space-air-ground-sea integrated network security in 6G," *IEEE Communications Surveys AND Tutorials*, vol. 24, no. 1, pp. 53-87, 2021.

[6] Y. Guo, Z. Mustafaoglu, D. Koundal, "Spam detection using bidirectional transformers and machine learning classifier algorithms," *Journal of Computational and Cognitive Engineering*, vol. 2, no. 1, pp. 5-9, 2022.

[7] I. Hidayat, M. Z. Ali, A. Arshad, "Machine learning-based intrusion detection system: An experimental comparison," *Journal of Computational and Cognitive Engineering*, vol. 2, no. 2, pp. 88-97, 2022.

[8] S. Krishnaveni, S. Sivamohan, S. S. Sridhar, S. Prabakaran, "Efficient feature selection and classification through ensemble method for network intrusion detection on cloud computing," *Cluster Computing*, vol. 24, no. 3, pp. 1761-1779, 2021.

[9] H. Lin, Q. Xue, J. Feng, D. Bai, "Internet of things intrusion detection model and algorithm based on cloud computing and multi-feature extraction extreme learning machine," *Digital Communications and Networks*, vol. 9, no. 1, pp. 111-124, 2023.

[10] M. N. Muhammad, N. Faisal, "Cybersecurity mechanism and user authentication security methods," *International Journal of Electronics and Information Engineering*, vol.14, no. 1, pp. 1-9, 2022.

[11] C. F. T. Pontes, M. M. C. De. Souza, J. J. C. Gondim, M. Bishop, M.A. Marotta, "A new method for flow-based network intrusion detection using the inverse Potts model," *IEEE Transactions on Network and Service Management*, vol. 18, no. 1, pp. 1125-1136, 2021.

[12] K. Tsiknas, D. Taketzis, K. Demertzis, C. Skianis, "Cyber threats to industrial IoT: a survey on attacks and countermeasures," *IoT*, vol. 2, no. 1, pp. 163-186, 2021.

[13] B. Wan, W. Jiang, Y. Fang, Z. Luo, G. Ding, "Anomaly detection in video sequences: a benchmark and computational model," *IET Image Processing*, vol. 12, no. 15, pp. 3454-3465, 2021.

[14] N. Wang, Y. Chen, Y. Xiao, Y. Hu, W. Lou, Y. T. Hou, Manda, "On adversarial example detection for network intrusion detection system," *IEEE Transactions on Dependable and Secure Computing*, vol. 20, no. 1, pp. 1139-1153, 2022.

[15] F. Wu, T. Li, Z. Wu, S. Wu, C. Xiao, "Research on network intrusion detection technology based on machine learning," *International Journal of Wireless Information Networks*, vol. 28, no. 3, pp. 262-275, 2021.

[16] Z. Zhang, Y. Cao, Z. Cui, W. Zhang, J. Chen, "A many-objective optimization based intelligent intrusion detection algorithm for enhancing security of vehicular networks in 6G," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 6, pp. 5234-5243, 2021.

Biography

Jianjun Wu received his master's degree in Geographic Information Systems from Henan University in 2008. His current research interests are computer applications and software development, and he is an associate professor.

Software Vulnerability Detection and Analysis Technology Integrating Static Taint Analysis and Deep Learning

Li Luo[1] and Honghua Zhu[2]
(Corresponding author: Honghua Zhu)
Information Management Center, Hainan Health Vocational College[1]
Physical Education Institute, Qiongtai Normal University[2]
Haikou 571100, China
Email: m13627585325@163.com
(Received July 7, 2023; Revised and Accepted June 6, 2024; First Online Aug. 17, 2024)

Abstract

In the era of vigorous growth of the Internet, software security issues are receiving increasing attention. In software development, due to various factors, the software has vulnerabilities and leaks the private information in the software. For this issue, based on static taint analysis technology, patch comparison technology is used to improve and extract pollution paths. A software vulnerability detection model based on bidirectional recurrent neural networks is constructed by adding multiple long- and short-term memory models. The experimental results show that the average of accuracy, recall, precision, false negative rate, false positive rate, and F1 value based on static taint analysis are 89.78%, 92.08%, 98.23%, 22.05%, 13.55%, and 86.88%, respectively. The vulnerability detection model based on static taint analysis has higher accuracy, recall, precision, and F1 value, which can better classify pollution paths and perform better. The average accuracy of the vulnerability detection model is 93.08%, and the average recall rate is 97.22%. The vulnerability detection performance in the contaminated path is excellent. The experimental results demonstrate that the vulnerability detection models based on static taint analysis and bidirectional recurrent neural networks have good detection performance. This study aims to provide a certain reference value for network security detection.

Keywords: Patch Comparison; Recurrent Neural Network; Software Security; Static Taint Analysis; Vulnerability Detection

1 Introduction

Software vulnerability refers to a security vulnerability problem caused by negligence, errors, or malicious attacks during the design, implementation, or testing of software code [6]. For developers, software vulnerabilities may only be a technical issue, while for users, the impact of software vulnerabilities may lead to property damage, personal information leakage, and other issues [1]. Due to the widespread application of deep learning (DL) technology, introducing DL into vulnerability detection has also become a new research direction. In addition, static taint technology can scan the source code without running the software, further achieving control flow analysis and data flow analysis. However, in previous studies, vulnerability mining in source code mainly focused on treating the source code as natural language text and using sequence-based recurrent neural networks for training. This approach avoids the defects of manual detection, but ignores the semantic structure information of the source code. Based on this, to better detect vulnerabilities in software, this study uses static taint analysis (STA) technology and patch comparison technology to improve and extract pollution paths. And based on a recurrent neural network, multiple bidirectional long-term and short-term memory (Bi-LSTM) models are added to construct a software vulnerability detection model based on a bidirectional recurrent neural network (Bi-RNN). This study aims to improve the accuracy of software vulnerability detection and provide some reference for the field of network security.

This study is composed of four parts. The first part is the research results of domestic and foreign scholars on software vulnerability detection and taint analysis technology. In the second part, based on the STA technology, patch comparison technology is used to improve, the pollution path is extracted, and a software vulnerability detection model is built based on Bi-RNN. The third part tests and analyzes the vulnerability detection model of STA technology and the vulnerability detection model based on Bi-RNNs. The fourth part summarizes the article and points out its shortcomings.

2 Related Works

Software vulnerabilities have always been an important and critical research issue in network security, and some experts and scholars have conducted relevant research on software vulnerability detection based on DL. Liu *et al.* found that existing machine learning-based methods have an imbalance between code representation and code. A deep balancing system has been developed to address this issue. And a deep neural network with Bi-LSTM was designed to learn code representation from tagged vulnerable and non-vulnerable codes. The experimental findings denoted that this method could significantly improve vulnerability detection performance [11]. Alenezi *et al.* believed that DL technology has been successfully applied in natural language processing. Due to the similarity between source code and natural text, an improved automatic vulnerability detection method based on character embedding technology was proposed. The proposed method was tested with a large C/C++open source code base, and the outcomes indicated that the proposed method had excellent performance [2]. Tamboli and Moparthi believed that unpredictable behavior and unknown advanced attack vulnerabilities posed significant network security challenges. Therefore, a DL model for detecting network attacks was constructed, and the effectiveness was compared with other detection models. The outcomes expressed that the proposed model had better performance than existing detection models when evaluated using the CICIDS-2017 dataset [17]. Batur Şahin and Abualigah believed that STA had a high false positive rate (FPR) in vulnerability detection. Based on this, a new vulnerability detection model based on DL was proposed by introducing the clustering theory of clonal selection algorithm. The detection ability of STA could be improved by immune based feature selection model. Comparing the model with other feature classification models, experimental results showed that this method had significantly improved classification accuracy and true positive rate [4].

The taint analyses technology is divided into STA and dynamic taint analyses, and is widely used in vulnerability information mining. Some experts and scholars have conducted relevant research based on this. Ami *et al.* found that existing STA tools had defects in their analysis, so they constructed a mutation-based robustness evaluation framework. This framework utilized mutation analysis to systematically evaluate android STA tools for identifying, documenting, and fixing defects. The experiment found that the framework found unrecorded defects in the Flow Drive framework, which stimulated the demand for system discovery and recording of unreasonable choices in sound tools, and demonstrated the opportunity to achieve this goal through mutation detection [3]. Zhang *et al.* found that existing STA techniques often only used a set of benchmarks for evaluation in mobile applications, making it difficult to generalize the results. Based on this, three of the most famous STA techniques were compared under different configurations and evaluated on a set of universal benchmark tests and real applications from Google App Store. Finally, the analysis results were compared with those in previous research reports to explore the main factors for the inaccuracy of existing tools [20]. Zhang *et al.* found that Flow Droid could not be directly used to detect DL applications because it was difficult to detect third-party application programming interfaces. To address this issue, a static information flow analysis tool based on DL has been proposed, which could effectively identify third-party application programming interfaces. The research findings denoted that 26.0% of applications had sensitive information leakage issues detected by the static information flow analysis tool based on DL which outperformed existing tools in detecting and identifying vulnerabilities [19]. Sun *et al.* found that STA techniques might report a large number of false positives due to excessive approximation. Therefore, a static method was proposed specifically for highlighting sensitive operations that were hidden. The experimental results indicated that this method successfully revealed the anti analysis code fragments that evaded detection through dynamic TA, helping security analysts verify potential sensitive information that existed before [16].

In summary, vulnerability detection based on DL has been widely applied in practice and has high stability and accuracy. STA technology can comprehensively inspect programs and has good flexibility and applicability. Existing research has conducted more independent research on vulnerability detection and STA techniques for DL, while little research has been conducted on the fusion technology of the two. Therefore, this study is based on improved STA technology to extract pollution paths. And based on a Bi-RNN, multiple Bi-LSTM models are used to improve and construct a software vulnerability detection model. This study aims to provide important reference for the field of software vulnerability detection.

3 Construction of a Vulnerability Detection Model Based on STA and DL

To achieve the purpose of vulnerability detection in software, this chapter is divided into two parts for model construction. The first part is based on STA technology, using patch comparison technology to improve and extract pollution paths. The second part builds vulnerability detection model based on Bi-RNNs.

3.1 Pollution Path Extraction Based on STA Technology

STA technology is a technique that detects and identifies security vulnerabilities and risk points according to specific rules by analyzing program source code [13]. This technology does not require program execution. By ana-

lyzing the code and data flows in the program, all possible inputs that may be encountered in the program can be checked and which positions and parameters may pose a security threat throughout the entire execution of the program [10]. However, STA technology also has some limitations in practical applications. Due to factors such as the diversity of variable data types, program processes, and logical structures, this technology sometimes has a large number of false alarms and false positives. Therefore, patch comparison technology is used for improvement, and the flowchart is shown in Figure 1.

The source code file contains a set of code containing vulnerabilities, and the patch file contains code files that fix the corresponding vulnerabilities. The source code file is compared with the corresponding patch file to generate the initial difference file. The Joern tool is utilized to extract the relevant code in the difference file, generate a control-flow graph from the extracted code, and save it [14]. In the difference file, taint selection rules are used to extract taints as initial points, and forward and backward taint propagation are performed and analyzed at the initial points. The paths formed by forward and backward taint propagation are combined to form a complete taint path. The STA structure diagram is shown in Figure 2.

The main purpose of STA is to check the source of taints in the program and provide users with suspicious sensitive data in the code caused by input or external environment [18]. The selection of taint source is one of the committed step in performing taint analyses. By determining whether the variables or parameters in the program are taint sources, it can track the flow of variables or parameters throughout the program and identify vulnerabilities and security risks. The common methods of taint source selection include manual selection, manual annotation, automatic selection and data-flow analysis. STA based on patch comparison is used to analyze a large number of different files to determine the relationship between vulnerabilities and patches. Taints are selected through differential files, following the three selection principles of calling function parameters, adding conditional variables for rows, and common variables, and selecting them in order [5]. After locating the location of the taint, the taint is analyzed using a forward and backward bidirectional taint propagation technology. The expression is shown in Equation (1).

$$X \oplus Y = \begin{cases} T & \text{if } x = T \text{ or } y = T \\ U & \text{if } x = U \text{ or } y = U \end{cases} \quad (1)$$

In Equation (1), T means the taint variable; U expresses the uncontaminated variable; $\oplus$ indicates the binary operation; x and y represent nodes. In forward taint propagation, the taint travels from the input of the program to the output of the program, which may trigger a security vulnerability in the middle. It can help detect whether input taints can spread to dangerous spots in the program, thereby discovering hidden vulnerabilities in the program. In backward taint propagation, it observes the output of the program and infers the input of the program. It reverses and analyzes whether the taints in the input will be affected by manipulation and attacks [15]. The combination of forward taint propagation and backward taint propagation can provide more comprehensive security checks and vulnerability analysis. The path obtained by using the method of automatically selecting taints will contain noisy paths. Due to differences in the location of vulnerabilities, cross data vulnerability analysis method is used to detect vulnerabilities in different functions. First, it needs to determine the relationship between functions and data flow, conduct STA on programs, determine the structure of programs, collect function call graphs and data flow graphs, and check the parameters, return values, global variable and other variables that may cause vulnerabilities. After completing relationship determination and data collection, it is necessary to sort out vulnerabilities. Common cross data types include buffer overflow, formatted string, etc. Due to the lack of judgment on whether the taint has been harmless during forward propagation, it is necessary to conduct a legitimacy check on the taint variable. When the analysis of the taint ends prematurely, it needs to check if there is a validity check for the variable in front of the taint. If there is a relevant check, it continues with the taint analyses until the true vulnerability is found.

3.2 A Vulnerability Detection Model Based on DL

Extracting vulnerabilities from code is an abstract and quantitative process that may lead to the model falling into an overfitting state. Therefore, DL models are used for vulnerability detection. Common DL models include convolutional neural network (CNN) and recurrent neural network (RNN). RNN is suitable for sequence data, which is suitable for software vulnerability detection because of its ability to process variable length sequences. The flowchart of RNN detection method is shown in Figure 3.

Before using RNN for vulnerability detection, it needs to preprocess the data, including standardizing the pollution path, converting the code into sequence, and uniformly representing the variables and functions in the code with numbers. It quantifies the standardized path for subsequent input [9]. Then the vectorized path and the corresponding label of the path are input into the Bi-RNN for training. Therefore, the problem of vulnerability detection task research has the mapping relationship shown in Equation (2).

$$f : C \to Z \to Y, C = \{c_i\}, z = \{z_i\}, Y = \{y_i\} \quad (2)$$

In Equation (2), c_i means the pollution path extracted through STA; z_i indicates the vectorization of the pollution path; y_i represents the label category corresponding to each pollution path. To solve the problem of binary classification in vulnerability detection, word vector model and Bi-RNN are used. Firstly, it standardizes the

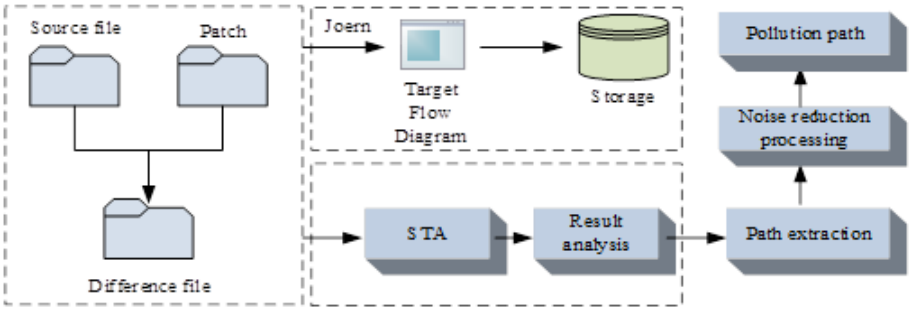


Figure 1: Flow chart of STA technology

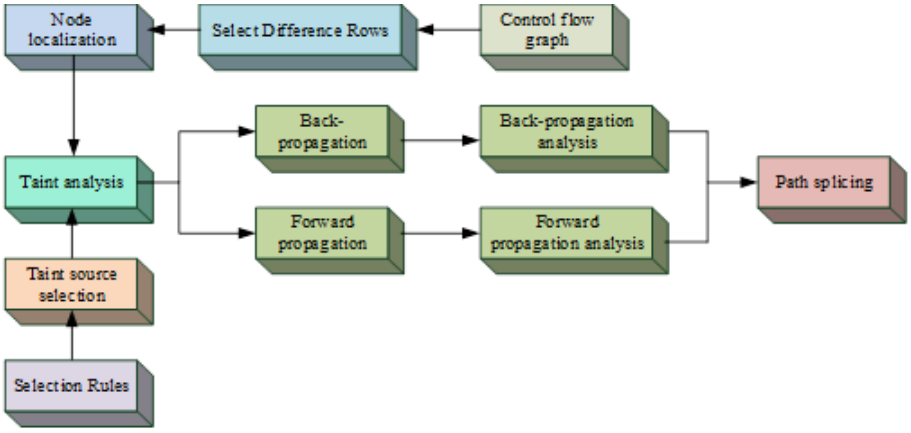


Figure 2: Structural diagram of STA

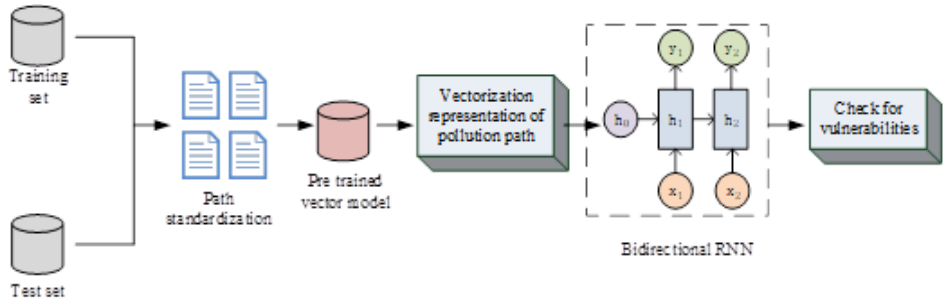


Figure 3: Flow chart of RNN detection method

input pollution path, and inputs the processed data into the word vector training model. Through the input layer, a fully connected neural network layer and a layer of continuous bag-of-words model (CBOW), it is transformed into the input word vector of the Bi-RNN [7]. The processed word vector is input into multiple BiLSTM models in the Bi-RNNs for learning, and the learning results are input into the full connected layer, and finally output through the output layer to determine whether the samples have vulnerabilities. The structure diagram of the vulnerability detection model is denoted in Figure 4.

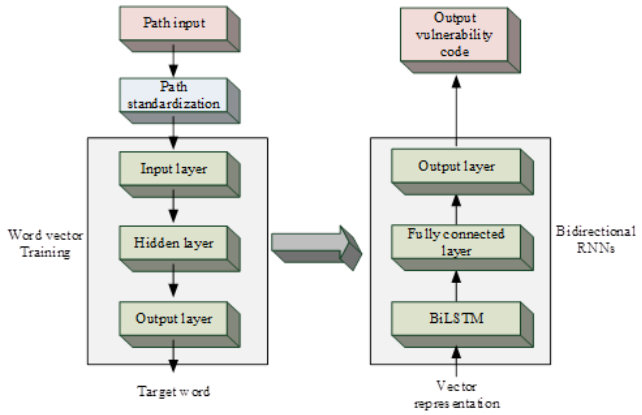


Figure 4: Structure diagram of vulnerability detection model

Due to the inability of the standardized pollution path to be directly input into the neural network model, the CBOW model is applied to vectorize the path. The expression is shown in Equation (3).

$$P(y = w_n|x) = \frac{e^{x_n}}{\sum_{k=1}^{N} e^{x_k}} \tag{3}$$

In Equation (3), x means the N-dimensional original output vector, and x_n expresses the word value of the corresponding dimension to the original output vector. In the RNN model, the expression for calculating the information contained in the states of each layer is shown in Equation (4).

$$h_t = f(Ux_t + Wh_{t-1}) \tag{4}$$

In Equation (4), h_t refers to the hidden state of the t-th layer; W stands for the weight matrix of the state transition at the moment; U means the weight matrix from the input layer to the hidden layer; x_t denotes the current input value; h_{t-1} denotes the state of the previous layer; f represents the activation function. The calculation expression for the output of the last layer is shown in Equation (5).

$$y = g(Vh_t) \tag{5}$$

In Equation (5), g refers to the Softmax Activation function. In RNN, the error of the next layer can be corrected by back propagation to the upper layer, which is conducive to the model to deal with the problem of long time series, but there is still the problem of vanishing gradient problem. LSTM is utilized to solve the vanishing gradient problem. The hidden layer of LSTM contains cell state and runs through the whole sequence [12]. LSTM includes input, output and forgetting gates. The expression for the information contained in forgetting gate is expressed in Equation (6).

$$f_t = \delta(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{6}$$

In Equation (6), δ refers to the weight; W_f denotes the matrix of the forgetting gate; b_f expresses the vector of the forgetting gate which decides whether the information is retained. The expression for the input gate is shown in Equation (7).

$$u_t = \delta(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{7}$$

In Equation (7), W_i stands for the matrix of the input gate, and b_i refers to the vector of the input gate. The information passed through the sigmoid function has a value between [0,1], with values closer to 0 indicating forgetting the information and closer to 1 indicating retaining the information. The expression for the output gate is shown in Equation (8).

$$o_t = \delta(W_0 \cdot [h_{t-1}, x_t] + b_0) \tag{8}$$

In Equation (8), W_o and b_o denote the matrix and vector of the output gate, respectively. The cell state expression is shown in Equation (9).

$$C_t = f_t * C_{t-1} + i_t * \widetilde{C_t} \tag{9}$$

In Equation (9), f_t represents the information contained in the forgetting gate; C_{t-1} indicates the cell state in the previous hidden layer; $\widetilde{C_t}$ expresses the updated cell state. The expression for cell update status is shown in Equation (10).

$$\widetilde{C_t} = \tanh(W_c \cdot [h_{t-1}, x_t] + b_C) \tag{10}$$

In Equation (10), W_C means the cell matrix and b_C refers to the cell vector. From this, a new hidden layer state can be obtained, as shown in Equation (11).

$$h_t = o_t * \tanh(C_t) \tag{11}$$

In Equation (11), h_t indicates the updated hidden layer state, and o_t stands for the information of the output gate. Similar to LSTM, there is a Gated Recurrent Unit (GRU) in the RNN variant that controls the flow of information in a gated manner [8]. The GRU includes reset and update gates, and the calculation method for the update gate is shown in Equation (12).

$$z_t = \delta(W_z[h_{t-1}, x_t]) \tag{12}$$

In Equation (12), δ means the weight, and W_z indicates the matrix of the update gate. Resetting the gate can

determine the quantity of forgotten information, and the calculation method is shown in Equation (13).

$$r_t = \delta(W_r[h_{t-1}, x_t]) \tag{13}$$

In Equation (13), W_r denotes the matrix of the reset gate. The memory state of the GRU is updated based on the values of the update and reset gates, as shown in Equation (14).

$$\begin{cases} h'_t = \tanh(W[r_t h_{t-1}, x_t]) \\ h_t = (1 - z_t)h_{t-1} + z_t h'_t \end{cases} \tag{14}$$

In Equation (14), h_{t-1} means the implicit state of the t-1 layer; h'_t expresses the memory state before the update; h_t refers to the memory state after the update. The schematic diagrams of LSTM and GRU models are expressed in Figure 5.

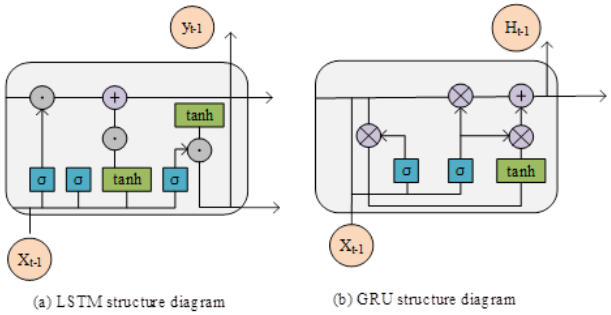


Figure 5: Structure diagram of LSTM and GRU

Compared to the LSTM model, the update gate in GRU is equivalent to the forget and input gates, while the reset gate is equivalent to the output gate, simplifying the training steps and achieving faster training speed. The calculation method for Bi-LSTM is denoted in Equation (15).

$$\begin{cases} h_{t1} = \vartheta(W_{xh_1}x_t + W_{h_1h_1}h_{(t-1)1} + b_{h_1}) \\ h_{t2} = \vartheta(W_{xh_2}x_t + W_{h_2h_2}h_{(t-1)2} + b_{h_2}) \\ o_t = W_{h_1o}h_{t1} + W_{h_2o}h_{t2} + b_0 \\ y_t = \delta(W_{oy}o_t + b_y) \end{cases} \tag{15}$$

In Equation (15), ϑ denotes the weight; x_t expresses the sequence of pollution paths; h_1 refers to the set of forward hidden layer vectors, and $h_1 = \{h_{11}, h_{21}, \cdots, h_{(t-1)1}, h_{t1}\}$; h_2 stands fro the reverse hidden layer vector, $h_2 = \{h_{12}, h_{22}, \cdots, h_{(t-1)2}, h_{t2}\}$; o represents the set of fully connected layer output vectors, $o = \{o_1, o_2, \cdots, o_{(t-1)}, o_t\}$; y stands for the output vector, $y = \{y_1, y_2, \cdots, y_{(t-1)}, y_t\}$.

4 Model Performance Testing and Result Analysis

To test the model's performance, this chapter is separated into two parts for testing. The first part and the second part test the vulnerability detection model based on STA and DL, respectively.

4.1 Performance Testing and Analysis of Vulnerability Detection Based on STA

To test the performance of vulnerability extraction based on STA, accuracy, recall, precision, FPR, false negative rate (FNR), and F1 values were used for evaluation. Adopting the open source framework Joern, with its fuzzy parsing method, it was suitable for machine learning applications. The Learning rate alpha of the Bi-LSTM model was set to 0.0025, the batch value was set to 128, the iteration number epoch was 5, and the Dropout value was set to 0.5. Other experimental running environments are displayed in Table 1.

Table 1: Experimental operating environment

Experimental environment	Experimental configuration
Processor	Intel(R)Corei7-8750H 2.20GHz
Memory	8GB
Graphics card	NVIDIA GeForce GXT 850M
System	Ubuntu 18.04
Static analysis framework	Joern1.0.141

It compared the vulnerability detection model based on slicing with the model based on STA. The CWE-119 dataset was selected, with a total sample size of 39753, 10440 samples with vulnerabilities, and 29313 samples without vulnerabilities. 5 experiments were conducted on two models to get the average value, and the test results are shown in Figure 6.

In Figure 6(a), the average accuracy of the vulnerability detection model based on slicing and STA were 81.15% and 89.78%, respectively. In Figure 6(b), the recall rate of the vulnerability detection model based on slicing and STA were 81.25% and 92.08%, respectively. In Figure 6(c), the average precision of the vulnerability detection model based on slicing and STA were 79.52% and 98.23%, respectively. In Figure 6(d), the average FPR of the vulnerability detection model based on slicing and STA were 34.73% and 13.55%, respectively. In Figure 6(e), the average FNR of the vulnerability detection model based on slicing and STA were 55.67% and 22.05%, respectively. In Figure 6(f), the F1 average value of the vulnerability detection model based on slicing and STA were 51.59% and 86.88%, respectively. The vulnerability detection model based on STA had higher accuracy, recall, precision, and F1 value, which could better classify pollution paths and achieve better performance. To test the impact of checking the legitimacy of the taint path on model performance, the F1 value and accuracy

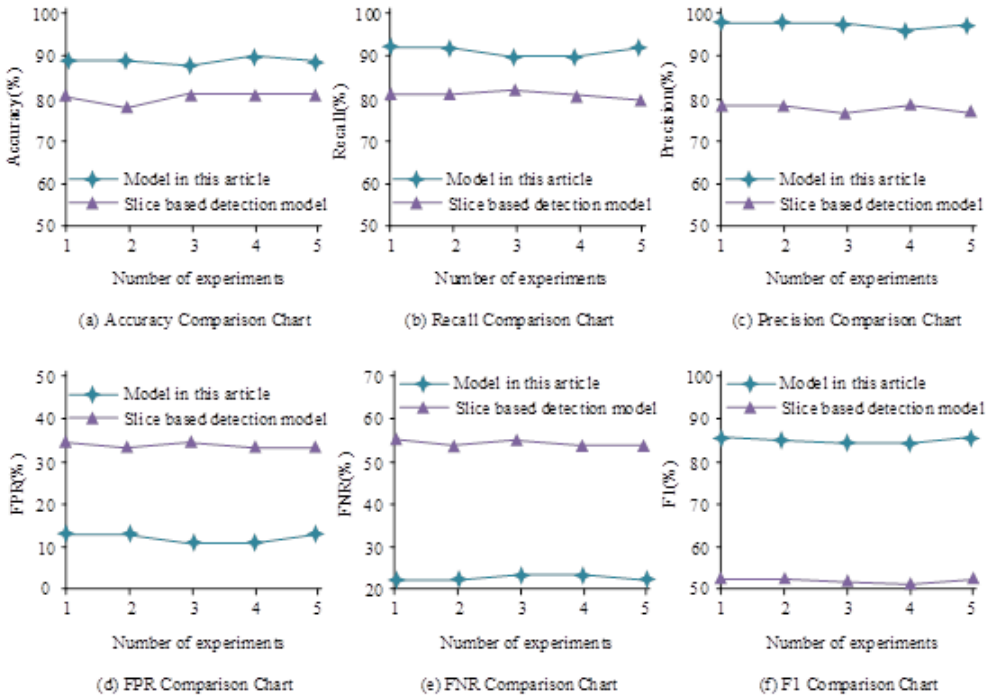


Figure 6: Indicator test results of two models

of patch-based taint analyses method, cross function taint analyses method, and analysis method after checking the legitimacy of the taint were compared. The test results are shown in Figure 7.

Figure 7(a) shows the accuracy comparison of three methods. The average accuracy of patch based and cross functional taint analyses methods were 89.25% and 85.64%, respectively. Figure 7(b) shows the comparison of F1 values among three methods. The patch based and cross functional taint analyses methods had average F1 value of 88.15% and 81.43%, respectively. The legality check F1 value for taints had an average F1 value of 92.35%. The legality check performance of taints was better, reducing the impact of noise on the experiment. 4.2 Performance Testing and Analysis of Vulnerability Detection Models Based on DL It selected the GRU and LSTM models for comparative testing. There were 1000 pieces of data in the dataset, with the first 600 pieces used for model training and the last 400 pieces used for model testing. It selected epoch as 20 and batch as 256. The test results are shown in Figure 8.

From Figure 8, the initial loss values of the two models were basically the same, and as the number of iterations increased, the model's loss value decreased. The GRU and LSTM models converged after 14 and 17 iterations, respectively, with a loss value of 0.8 and 1.2, respectively. Compared with LSTM model, GRU model had faster rate of convergence and lower loss value, so replacing LSTM with GRU in the model made the model have faster test speed. The support vector machine (VM), LSTM and GRU networks were compared and tested on five different datasets. The outcomes are expressed in Table 2.

From Table 2, under five different data sets, the mean squared error (MSE) values of GRU were 0.0552, 0.0364, 0.0219, 0.0413 and 0.0322 respectively, and the accuracy values were 94.01%, 96.15%, 97.01%, 95.24% and 96.23% respectively. Compared with the other two networks, it had lower mean squared error value and higher accuracy rate. It was effective to select GRU as the composition of vulnerability detection model. To test the performance of the LSTM and Bi-LSTM, the Juiet dataset was selected as the test data. The results are shown in Figure 9.

Figure 9(a) shows the accuracy curves of the LSTM and the Bi-LSTM models. The average accuracy of the Bi-LSTM, LSTM models were 93.85% and 90.46%, respectively. Figure 9(b) is the comparison curves of F1 values between the two models. The average F1 value of the Bi-LSTM and LSTM models were 93.79% and 92.52%, respectively. Bi-LSTM took into account bidirectional sequence information, and its feature learning performance was significantly superior to the unidirectional LSTM model. The vulnerability detection model based on Bi-RNN was compared with the existing VulDeePecker model based on DL. The results are shown in Figure 10.

Figure 10(a) shows the accuracy comparison curve of the two models. The average accuracy of VulDeePecker and vulnerability detection model based on Bi-RNN models were 86.55% and 93.08%, respectively. Figure 10(b) is the comparison curve of the recall rate of the two models. The average recall rate of VulDeePecker and vulnerability detection model based on Bi-RNN models were 87.13% and 97.22%, respectively. The vulnerability de-

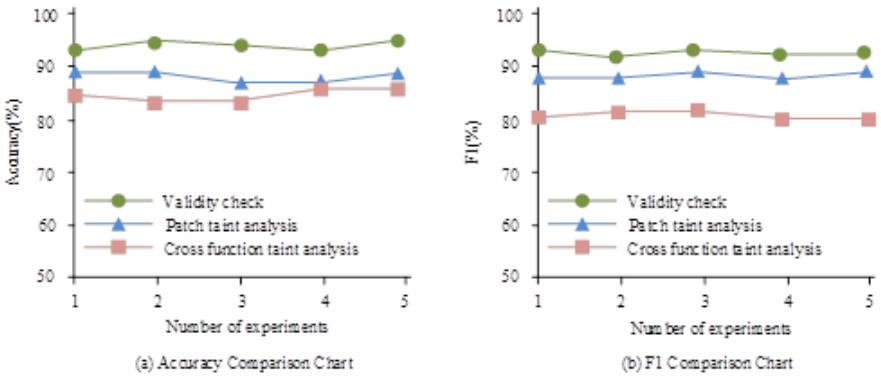


Figure 7: Comparison of F1 values and accuracy of three methods

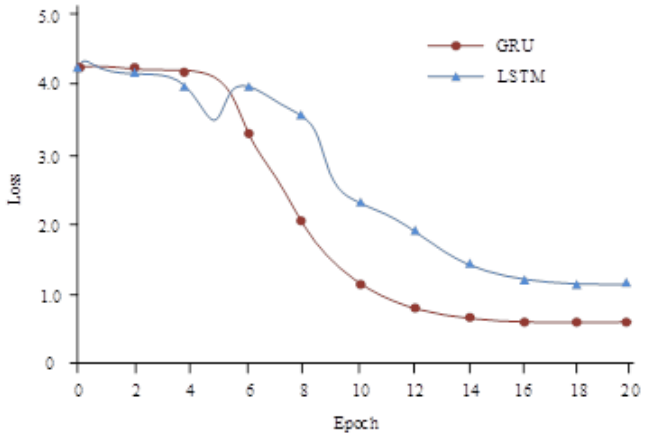


Figure 8: Iteration curves of two models

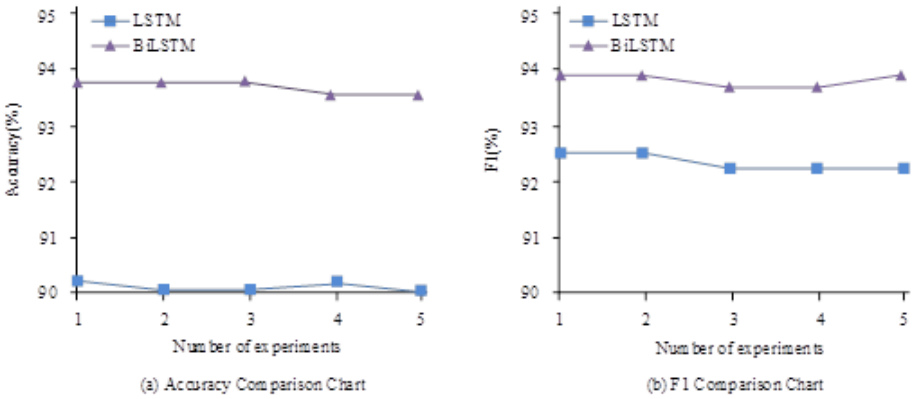


Figure 9: Comparison of accuracy and F1 values of three models

Table 2: Test results of three models under different datasets

Training methods	Number of training sets	Number of test sets	MSE of VM	VM accuracy	LSTM MSE	LSTM accuracy	GRU MSE	GRU accuracy
1st type	6000	650	0.1112	88.49%	0.0612	93.75%	0.0552	94.01%
2nd type	7100	800	0.0829	90.59%	0.0342	96.02%	0.0364	96.15%
3rd type	13500	1500	0.0495	94.28%	0.0289	95.89%	0.0219	97.01%
4th type	6600	7900	0.1359	86.12%	0.0418	95.03%	0.0413	95.24%
5th type	7900	6700	0.1475	85.11%	0.0409	95.02%	0.0322	96.23%

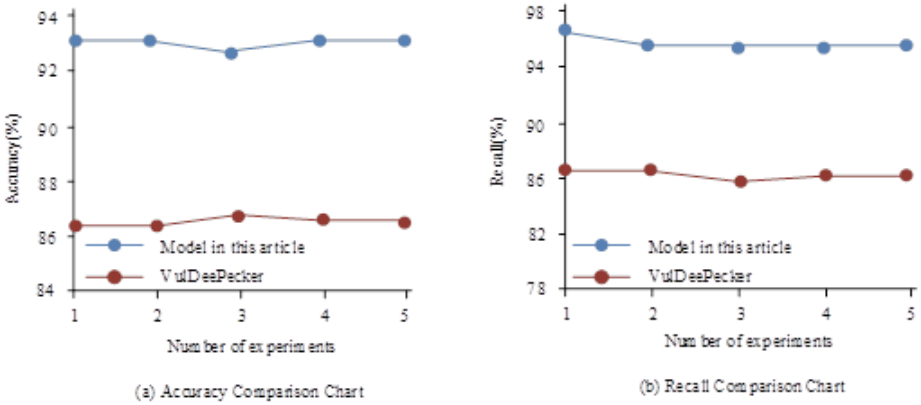


Figure 10: Comparison curves of accuracy and recall for different models

tection model based on Bi-RNN had higher detection accuracy and recall rate, and the classification effect of pollution paths was better than VulDeePecker model.

5 Conclusion

To detect vulnerabilities in software, a vulnerability detection model was constructed based on STA and Bi-RNNs. The results showed that the accuracy, recall, precision, FPR, FNR, and F1 average values of the vulnerability detection model based on slicing were 81.15%, 92.08%, 79.52%, 55.67%, 34.73%, and 51.59%, respectively. The accuracy, recall, precision, FPR, FNR, and F1 average values based on STA were 89.78%, 92.08%, 98.23%, 22.05%, 13.55%, and 86.88%, respectively. Therefore, compared to other models, the vulnerability detection model based on STA had higher accuracy, recall, precision, and F1 value, which can better classify pollution paths and perform better. The GRU model converged after 14 iterations with a loss value of 0.8, while the LSTM model converged after 17 iterations. Compared to the other two networks, the GRU model had lower mean square error values and higher accuracy, proving the effectiveness of selecting GRU as the composition of the vulnerability detection model. The vulnerability detection model based on Bi-RNN was compared and tested with the VulDeePecker model. The average accuracy of the VulDeePecker model was 86.55%, while the average accuracy of the vulnerability detection model based on Bi-RNN was 93.08%. The average recall rate of VulDeePecker model was 87.13%, and the average recall rate of vulnerability detection model based on Bi-RNN was 97.22%. The vulnerability detection model based on Bi-RNN had higher detection accuracy and recall rate, and its classification performance for contaminated paths was better than the VulDeePecker model. Through experiments, it has been proven that the two models constructed in this study have good performance in vulnerability detection and can achieve precise detection of vulnerabilities in software.

Acknowledgments

This research is supported by the general project of education and teaching reform in 2024, titled "Research on the Precise Teaching Model of Computer Basic Courses in Vocational Colleges Empowered by Data" (Project No.: Hnjg2024ZC-196). The key project of education and teaching reform in 2023, named "Research on the Application of Big Data in Finance and Commerce Majors of Vocational Undergraduate Students in the Era of Digital Intelligence" (Project No.: Hnjg2023ZD-62).

References

[1] S. Afrose, Y. Xiao, S. Rahaman, "Miller P.B. Evaluation of static vulnerability detection tools with Java cryptographic API benchmarks," *IEEE Transactions on Software Engineering*, vol. 49, no. 2, pp. 485-497, 2022.

[2] M. Alenezi, M. Zagane, Y. Javed, "Efficient deep features learning for vulnerability detection using character n- gram embedding," *Jordanian Journal of Computers and Information Technology*, vol. 7, no. 1, pp. 25-39, 2021.

[3] A. S. Ami, K. Kafle, K. Moran, A. Nadlkarni, D. Poshyvanyk, "Systematic mutation-based evaluation of the soundness of security-focused android static analysis techniques," *ACM Transactions on Privacy and Security*, vol. 24, no. 3, pp. 1-37, 2021.

[4] C. Batur Şahin, L. Abualigah, "A novel deep learning-based feature selection model for improving the static analysis of vulnerability detection," *Neural Computing and Applications*, vol. 33, no. 20, pp. 14049-14067, 2021.

[5] M. Cheung, J. J. Campbell, L. Whitby, J. Robert, J. Brabrook, J. Petzing, "Current trends in flow cytometry automated data analysis software," *Cytometry Part A*, vol. 99, pp. 10, pp. 1007-1021, 2021.

[6] A. C. Eberendu, V. I. Udegbe, E. O. Ezennorom, "A systematic literature review of software vulnerability detection," *European Journal of Computer Science and Information Technology*, vol. 10, no. 1, pp. 23-37, 2022.

[7] E. Iannone, R. Guadagni, F. Ferrucci, "The secret life of software vulnerabilities: A large-scale empirical study," *IEEE Transactions on Software Engineering*, vol. 49, no. 1, pp. 44-63, 2022.

[8] J. Iqbal, T. Firdous, A. K. Shrivastava, "Modelling and predicting software vulnerabilities using a sigmoid function," *International Journal of Information Technology*, vol. 14, no. 2, pp. 649-655, 2022.

[9] G. Landini, G. Martinelli, F. Piccinini, "Colour deconvolution: stain unmixing in histological imaging," *Bioinformatics*, vol. 37, no. 10, pp. 1485-1487, 2021.

[10] G. Lin, S. Wen, Q. L. Han, J. Zhang, Y. Xiang, "Software vulnerability detection using deep neural networks: A survey," *Proceedings of the IEEE*, vol. 108, no. 10, pp. 1825-1848, 2020.

[11] S. Liu, G. Lin, Q. L. Han, S. Wen, J. Zhang, Y. Xiang, "DeepBalance: Deep-learning and fuzzy oversampling for vulnerability detection," *IEEE Transactions on Fuzzy Systems*, vol. 28, no. 7, pp. 1329-1343, 2020.

[12] T. Mahmood, Z. Ali, "Prioritized muirhead mean aggregation operators under the complex single-valued neutrosophic settings and their application in multiattribute decision-making," *Journal of Computational and Cognitive Engineering*, vol. 1, no. 2, pp. 56-73, 2022.

[13] R. Manne, S. Kantheti, S. Kantheti, "Classification of Skin cancer using deep learning, convolutional neural networks -opportunities and vulnerabilities-a systematic review," *International Journal for Modern Trends in Science and Technology*, vol. 6, no. 11, pp. 101-108, 2020.

[14] S. Ren, G. Zhao, "A new formulation of continuous transverse shear stress field for static and dynamic analysis of sandwich beams with soft core," *International Journal for Numerical Methods in Engineering*, vol. 121, no. 8, pp. 1847-1876, 2020.

[15] S. Robert, L. G. Ungerleider, M. Vaziri-Pashkam, "Disentangling object category representations driven by dynamic and static visual input," *Journal of Neuroscience*, vol. 43, no. 4, pp. 621-634, 2023.

[16] X. Sun, X. Chen, L. Li, H. Cai, J. Grundy, J. Samhi, "Demystifying hidden sensitive operations in android apps," *ACM Transactions on Software Engineering and Methodology*, vol. 32, no. 2, pp. 1-30, 2023.

[17] M. B. Tamboli, N. R. Moparthi, "Deep learning model for intrusion identification," *Journal of Advanced Research in Dynamical and Control Systems*, vol. 12, no. 5, pp. 388-395, 2020.

[18] X. Wang, M. Cheng, J. Eaton, C. J. Hsieh, S. F. Wu, "Fake node attacks on graph convolutional networks," *Journal of Computational and Cognitive Engineering*, vol. 1, no. 4, pp. 165-173, 2022.

[19] J. Zhang, Q. Guo, T. Zhang, Z. Feng, X. Li, "Toward understanding and testing deep learning information flow in deep learning-based android apps," *International Journal of Computer and Systems Engineering*, vol. 17, no. 3, pp. 171-179, 2023.

[20] J. Zhang, Y. Wang, L. Qiu, J. Rubin, "Analyzing android taint analysis tools: FlowDroid, Amandroid, and DroidSafe," *IEEE Transactions on Software Engineering*, vol. 48, no. 10, pp. 4014-4040, 2021.

Biography

Li Luo received his bachelor's degree in Computer Science and Technology from Hunan University in 2005 and obtained his master's degree in Software Engineering from Chongqing University in 2014. Currently, he serves as a computer teacher in the Information Management Center of Hainan Health Vocational College. He has presided over 5 provincial-level research projects and participated in 6 school-level and provincial-level projects. He has also published 6 papers in provincial-level journals. His main research directions are computer networks and modern educational technology.

Honghua Zhu graduated from Hunan Normal University with a master's degree in Sports Human Science in 2007. After graduation, she served as a teacher in the Physical Education College of Qiongtai Normal University. She has participated in 3 provincial-level research projects of the university and published 2 papers. Her research directions are human anatomy, sports physiology, and health education.

Identifying and Intercepting Telecommunications Fraud Numbers on the Internet Through Big Data Technolog

Hui You[1] and Tuo Shi[2]

(Corresponding author: Hui You)

Cybersecurity and Protection Department, Beijing Police College[1]

Beijing Police College, Nanjian Road, Changping District, Beijing 102202, China

Beijing Police College, Beijing 102202, China[2]

Email: yhui1984@outlook.com

(Received July 30, 2023; Revised and Accepted July 28, 2024; First Online Aug. 17, 2024)

Abstract

With the advancement of network technology, methods of telecommunication fraud have become increasingly diverse. The identification and interception of fraudulent numbers are particularly crucial. This study utilized the network operator's big data as the basis and conducted Synthetic Minority Over-sampling Technique (SMOTE) processing and feature selection on the original data. The eXtreme Gradient Boosting (XGBoost) algorithm was employed to identify fraudulent numbers and enable early interception. Additionally, an improved Sparrow Search Algorithm (ISSA) algorithm was designed to optimize the parameters of the XGBoost algorithm, resulting in the development of the ISSA-XGBoost algorithm. Experiments were conducted using a collected dataset. The results indicated that after SMOTE processing and feature selection, the recognition effectiveness of the ISSA-XGBoost algorithm for fraudulent numbers was improved. Furthermore, compared to parameter optimization methods such as Grid Search and Bayesian Optimization (BO), the ISSA algorithm demonstrated superior performance. It achieved a precision of 0.945, a recall rate of 0.816, and an F1 value of 0.876. The method also exhibited a low resumption rate of 21.12% in practical applications. These findings validate the effectiveness of the proposed method for identifying and intercepting network telecommunications fraud numbers, further supporting its potential practical applications.

Keywords: Call Data; Number Identification; Parameter Optimization; Telecommunications Fraud; XGBoost

1 Introduction

As technological advancements continue, the scale, precision, intelligence, and specialization of telecommunications fraud have also improved. Criminal methods have become more complex and diverse [17], resulting in far-reaching threats and losses. The number of annual fraud cases and the associated monetary amounts are consistently rising [1]. Telecommunications fraud now constitutes a significant portion of network security concerns [5]. Contactless methods are predominantly used in telecom fraud, making it difficult for law enforcement agencies to gather evidence and make arrests due to their strong concealment, wide range, and prevalence of cross-border fraud [19].

Telecom operators possess valuable big data that can be utilized for anti-fraud purposes, enabling the identification of fraudulent numbers and early interception through data analysis and processing. With the support of big data technology, extensive research has been conducted on the prevention of telecom fraud [16]. Wang *et al.* [18] proposed a feature difference-aware graphical neural network, compared it with seven other baseline methods using real telecom datasets, and found an improvement of approximately 5%. Amin *et al.* [2] employed the spline classifier for telecom fraud detection, achieving improved classification performance while reducing feature dimensions and attaining 97.44% accuracy.

Ji *et al.* [9] introduced a multi-range gated graph neural network to model social networks as directed graphs for telecom fraud detection and obtained the most advanced outcomes in experimental testing. Amuji *et al.* [3] conducted a study on optimal classifiers for telecom fraud, finding that the optimal classifier had a posterior probability of 0.7368. They suggested that essential parameters to consider are the number of calls per hour and call duration. This paper used a method based on the eXtreme Gradient Boosting (XGBoost) algorithm to identify and intercept network telecom fraud numbers. Through experimental analyses using big data from telecom operators, the reliability of the proposed method was verified, making it suitable for practical telecom anti-fraud efforts

and achieving interception of fraudulent numbers.

2 Network Telecommunications Fraud and Big Data

2.1 Network Telecommunications Fraud

Network telecommunication fraud refers fraud using telephone, text messaging, the Internet, and other technologies to illegally occupy public and private property. Under the influence of the continuous development of network technology, the current telecom fraud presents the following characteristics.

Rapid updates of the modus operandi:
New technologies have provided criminals with novel methods to perpetrate fraud [14]. Moreover, criminals adapt their fraudulent tactics based on changes in social trends, news, policies, and other factors. Depending on the targeted victims, they use different strategies to gain trust and increase their success rate.

Non-contact:
Criminals engage in their illicit activities through various means such as telephone and Internet. The majority of their interactions occur through online channels, making it increasingly challenging to apprehend them.

Increased corporatization of crime:
Presently, criminals are increasingly organizing themselves into gangs and operating under the guise of legitimate businesses. They adopt corporate management modes to oversee their criminal activities, leading to the expansion and professionalization of criminal organizations. This trend has resulted in gradually establishing a professional criminal industry chain that shows a high level of organization.

2.2 Feasibility of Anti-fraud Through Big Data Technology

In the face of the ever-increasing prevalence of telecom fraud, solely relying on post-fraud crackdown is insufficient to combat this issue effectively. It is imperative to proactively intercept and provide early warnings in order to enhance the overall governance of telecom fraud [11]. For network telecommunication fraud, anti-fraud can be realized through big data technology, and its feasibility is analyzed below.

Accessibility of big data:
The digital transformation within public security agencies facilitates the availability of substantial foundational information about individuals and events, such as identity information and behavioral patterns of fraud gangs. By analyzing data from previous cases involving criminal gangs, it is possible to provide public security agencies with valuable insights for identifying perpetrators. Furthermore, apart from leveraging internal data within public security agencies, collaboration with telecommunication carriers and payment enterprises and utilizing big data from relevant departments enables identifying and freezing numbers and accounts associated with fraudulent activities.

Applicability of data mining techniques:
In the process of telecom anti-fraud, various data collected can be analyzed using data mining techniques to uncover patterns, trends, and relationships within the crime data. Techniques such as clustering, correlation analysis, and neural networks [13] can be employed to identify hidden relationships between different data objects and identify telecom fraud.

2.3 Analysis of Big Data in Telecommunications Operators

This research primarily relies on the big data provided by the telecommunication carrier to investigate the identification and interception of fraudulent numbers. The dataset used consists of actual user data obtained from a telecom operator, covering a period from July 2023 to December 2023. The dataset include 6,216 users, and the data has been desensitized to address privacy concerns. The fraudulent numbers within the dataset, labeled as 1, are sourced from the public security system and amount to 1,626 in total. The normal numbers, labeled as 0, are used by 4,590 ordinary users. The data includes the user's basic information, call records, text message data, and traffic data, as illustrated in Table 1.

The data in Table 1 are all quantitative variables. Outliers were eliminated, and for missing values, the mean value was selected to fill in. Then, max-min transformation is performed on the original data to obtain new value x' in the interval of [0,1]:

$$x' = \frac{x - \min}{\max - \min}$$

where max and min are the maximum and minimum values of x.

2.4 Imbalance Processing and Feature Selection

There are 1,626 fraudulent numbers in the collected data, accounting for 26.16%, and 4,590 normal numbers, accounting for 73.84%. The data is unbalanced, which may lead to a decline in the recognition effect. Therefore, this paper uses the Synthetic Minority Over-sampling Technique (SMOTE) algorithm [6] to deal with the samples. Firstly, a sample called x_i is randomly selected from a minority class. Then, N samples (x_{z_i}) are randomly selected

Table 1: Big data from the telecom operator

Category	Element
Basic data	Toll encryption number
	Quantity of numbers under each name
	Average monthly consumption
	Number of months with null consumption
	Fraudulent number label
Call data	Number of months with zero calls
	Average number of outgoing calls per month
	Average duration of outgoing calls per month
	Average number of people called per month
Text message data	Number of months with zero text messages
	Monthly average number of ascending text messages
	Monthly average number of ascending text message recipients
	Monthly average number of descending text messages
	Monthly average number of descending text message recipients
Traffic data	Number of months with zero traffic
	Average monthly traffic
	Total number of apps used per month

from k neighbors of x_i to generate new samples:

$$x_n = x_i + \delta \times (x_{z_i} - x_i)$$

where δ is a random number in [0,1]. After SMOTE processing, the distribution of fraudulent and normal numbers is shown in Table 2.

In order to further filter the data in Table 1 and improve the identification of fraudulent numbers, Pearson correlation [7] is calculated:

$$r = \frac{\sum(X - \bar{X})(Y - \bar{Y})}{\sqrt{\sum(X - \bar{X})^2}\sqrt{\sum(Y - \bar{Y})^2}}$$

The correlation coefficient between each feature and whether it is fraudulent is calculated, and the results are shown in Table 3.

Ten features from Table 3 that have a strong correlation with fraud are selected for fraud number identification and interception. It can be found that there is a strong correlation between the user's basic data, call data, and whether a number is fraudulent. The three features with the highest correlation coefficients are the number of months with zero calls, the average number of people called per month, and the average number of outgoing calls per month. These findings align with the actual situation. Normal numbers usually do not have zero call records. If there are numerous months with no call, it indicates a high likelihood of fraudulent numbers. Additionally, normal numbers usually have less people to call and less outgoing calls than fraudulent numbers. Fraudulent numbers tend to have more outgoing calls, as they are utilized for fraudulent activities.

3 Fraudulent Number Identification and Interception Based on XGBoost

3.1 XGBoost Algorithm

XGBoost, a machine learning algorithm, is an enhanced version of the gradient boosting decision tree that offers improved performance and higher speed. It has favorable performance in various recognition and classification tasks [4]. Therefore, this study explores the identification and interception of fraudulent numbers using the XGBoost algorithm.

In dataset $D = \{(x_i, y_i)\}$ composed of n-dimensional samples and d-dimensional features, the XGBoost algorithm obtains the output through integrating k decision trees:

$$\hat{y_i} = \sum_{k=1}^{K} f_k(x_i)$$

The objective function of the algorithm at the t-th iteration is:

$$\varsigma^{(t)} = \sum_{j=1}^{T}[w_j(\sum_{i\in I_j} g_i) + \frac{1}{2}w_j^2(\sum_{i\in I_j} h_i + \vartheta)] + \gamma T$$

where g_i is the first-order gradient, h_i is the second-order gradient, ϑ and γ are hyperparameters. The following parameters are defined.

The cumulative sum of the first-order partial derivatives of leaf node j:

$$G_j = \sum_{i\in I_j} g_i$$

Table 2: Distribution of samples after SMOTE processing

	Fraudulent number	Normal number
Original data	1,626 (26.16%)	4,590 (73.84%)
After SMOTE processing	4,590 (50%)	4,590 (50%)

Table 3: Correlation coefficient results (Bolded: r value between 0.5 and 0.8, indicating a stong correlation)

	Correlation coefficient
Quantity of numbers under each name	0.684
Average monthly consumption	0.587
Number of months with null consumption	0.615
Number of months with zero calls	0.774
Average number of outgoing calls per month	0.732
Average duration of outgoing calls per month	0.604
Average number of people called per month	0.756
Number of months with zero text messages	0.146
Monthly average number of ascending text messages	0.564
Monthly average number of ascending text message recipients	0.336
Monthly average number of descending text messages	0.125
Monthly average number of descending text message recipients	0.104
Number of months with zero traffic	0.133
Average monthly traffic	0.541
Total number of apps used per month	0.533

The cumulative sum of the second-order partial derivatives of leaf node j:

$$H_j = \sum_{i \in I_j} h_i$$

After substitution, the following objective function is obtained:

$$\zeta^{(t)} = \sum_{j=1}^{T} [w_j G_j + \frac{1}{2} w_j^2 (H_j + \vartheta)] + \gamma T$$

The derivative of w_j is calculated using $\zeta^{(t)}$:

$$w_j = -\frac{G_j}{H_j + \vartheta}$$

The simplified objective function is obtained:

$$\zeta^{(t)} = -\frac{1}{2} \sum_{j=1}^{T} \frac{G_j^2}{H_j + \vartheta} + \gamma T$$

3.2 Parameter Optimization-based Approach

Some parameters in the XGBoost algorithm need to be set by humans. However, it is often difficult to find the optimal value for setting the parameters manually, leading to the model's poor performance. In order to obtain better results for fraudulent number identification and interception, this paper optimizes the critical parameters in the XGBoost algorithm by combining with the Sparrow Search Algorithm (SSA), an algorithm based on the foraging behavior of sparrows [20], where individuals within the population are divided into searchers and followers. The searchers are responsible for searching for food, and their position updating formula is:

$$s_{i,j}^{t+1} = \begin{cases} s_{i,j}^t \cdot exp(-\frac{i}{\alpha t_{\max}}), & \text{if } R < ST \\ s_{i,j}^t \cdot Q \cdot L, & \text{if } R \geq ST \end{cases}$$

where $s_{i,j}^t$ is the position of the i-th individual at the j-th dimension during the t-th iteration, α is a random number between 0 and 1, $t_{\max}$ is the maximum number of iterations, R is the early warning value, ST is the security value, Q is a random numbers that follow a normal distribution, and L is a $1 \times d$ matrix whose elements are all 1. The follower's position updating formula is:

$$s_{i,j}^{t+1} = \begin{cases} Q \cdot exp(s_w - s_{i,j}^t), & \text{if } i > \frac{n}{2} \\ s_s^{t+1} + |s_{i,j}^t - s_s^{t+1}| \cdot A^+ \cdot L, & \text{otherwise} \end{cases}$$

where s_s^{t+1} is the optimal individual position, s_w is the poorest group position, and A^+:

$$A^+ = A^T (A A^T)^{-1}$$

A is a $(1 \times d)$ matrix.

In case of danger, the individual position updating formula can be written as:

$$s_{i,j}^{t+1} = \begin{cases} s_b^t + \beta \cdot |s_{i,j}^t - s_b^t|, & \text{if } f_i \neq f_j \\ s_{i,j}^t + K \cdot [\frac{s_{i,j}^t - s_w^t}{(f_i - f_w) + \epsilon}], & \text{if } f_1 = f_j \end{cases}$$

where s_b^t is the optimal group position, β is the parameter of step length, K is a random number in [-1,1], f_i is the current individual fitness value, f_w and f_j are the current worst and best fitness values. In order to obtain a higher-quality initial population, this paper proposes an improved SSA (ISSA), which uses the circle mapping [12] to achieve the population initialization:

$$x_{k+1} = \text{mod } [x_k + b - (\frac{1}{2\pi} \sin(2\pi x_k), 1]$$

where mod is the modulus function, $a = 0.5$, $b = 0.2$. Based on circle mapping, more uniform population can be obtained, thus improving the optimization effect. Based on the ISSA, three key parameters in the XGBoost algorithm, i.e., the number of trees, the maximum tree depth, and the learning rate, are optimized, and then the optimal parameters are used to construct the XGBoost model for identification and interception of fraudulent numbers.

4 Results and Analysis

4.1 Parameter Settings and Evaluation Methods

The ISSA-XGBoost model was built on the Windows 10 operating system based on MATLAB environment. The population size of the ISSA was 50, and the maximum number of iterations was 20. After optimization, the optimal number of trees (n_estimators) was 50, the maximum tree depth (max_deoth) was 5, and the learning_rate was 0.1. The rest of the parameters took default values. A five-fold cross-test was used for the experiments, and the final results were averaged. The following indicators are used in the evaluation of the fraudulent number recognition effect:

$$\begin{aligned} precision &= \frac{TP}{TP + FP} \\ recall &= \frac{TP}{TP + FN} \\ F_1 &= \frac{2precision}{precision + recall} \end{aligned}$$

In the above equations, TP is the number of normal numbers recognized as normal, FP is the number of normal numbers identified as fraudulent numbers, FN is the number of fraudulent numbers identified as normal numbers. The ISSA-XGBoost algorithm was deployed in real application environments for interception of identified fraudulent numbers. After the fraudulent number is restricted from use if the number is not fraudulent, the user will contact the operator for resumption of the phone number; therefore, the resumption rate was used to determine the model's interception effect: resumption rate = number of restored numbers/number of closed numbers.

4.2 Result Analysis

The impact of imbalance processing and feature screening was analyzed. The results of fraudulent number identification using the ISSA-XGBoost algorithm are presented in Table 4.

Table 4 shows that when conducting experiments on the original data, the ISSA-XGBoost achieved a high precision of 0.846. However, it had a lower recall rate of 0.516 and an F1 value of 0.641. This result indicated that the imbalanced data significantly impacted the recognition effectiveness for fraudulent numbers. After applying SMOTE processing, the precision of the ISSA-XGBoost improved by 0.066, reaching 0.912. The recall rate also significantly improved by 0.271, reaching 0.787. The F1 value increased by 0.204, reaching 0.845. These improvements demonstrated the beneficial role of SMOTE. Furthermore, even better results were achieved after performing feature screening and utilizing strong correlation features as inputs for the ISSA-XGBoost. The precision increased to 0.945, the recall rate improved to 0.816, and the F1 value reached 0.876. These improvements highlighted the importance of feature screening. For parameter optimization of the XGBoost algorithm, the ISSA was compared with the following optimization algorithms:

1) Grid Search (GS) [15],

2) Bayesian Optimization (BO) [10],

3) Particle Swarm Optimization (PSO) [8],

4) SSA.

The results are presented in Figure 1.

Figure 1 shows that the recognition effect of the XGBoost algorithm were improved to some extent after parameter optimization. It can be observed that the improvement in the fraudulent number recognition effect was relatively minor for the XGBoost algorithm optimized by GS and BO compared to the two swarm intelligence algorithms, PSO and SSA. Moreover, in the comparison between the PSO and SSA, it is evident that the SSA outperformed the PSO algorithm, achieving an F1 value of 0.871, an increase of 0.004 compared to the PSO algorithm. Additionally, after further enhancement through circle mapping, the F1 value of the ISSA improved by 0.005 compared to the SSA, further highlighting the effectiveness of the ISSA for enhancing XGBoost performance. To verify the effectiveness of the proposed method for interception of fraudulent numbers, it was compared with the random forest (RF) algorithm. Both algorithms were applied in practice for six months, and the resumption rates were compared in Figure 2.

Table 4: Effect of data processing on the results

	Precision	Recall rate	F1 value
Original data	0.846	0.516	0.641
SMOTE processed data	0.912	0.787	0.845
SMOTE processed data + feature screening	0.945	0.816	0.876

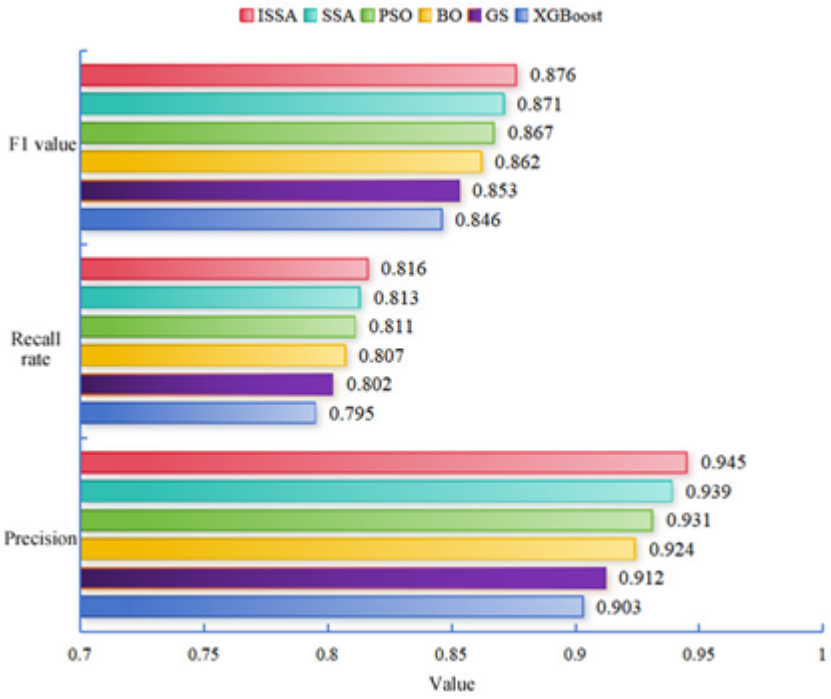


Figure 1: Recognition results under different parameter optimizations

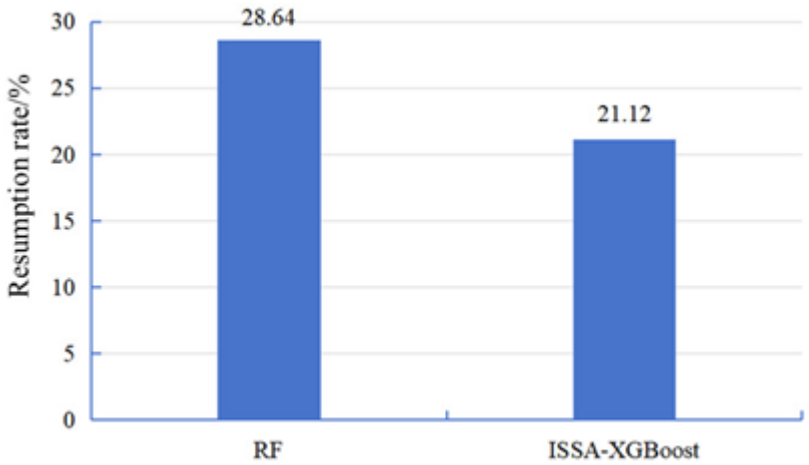


Figure 2: Comparison results of resumption rate

Figure 2 shows that when using the RF algorithm for identification and interception, the resumption rate within a six-month period was 28.64%. However, when applying the ISSA-XGBoost method, the resumption rate within the same six-month period was 21.12%, which showed a significant reduction of 7.52% compared to the RF. The results indicated that ISSA-XGBoost method had a more favorable practical application effect.

5 Conclusion

This paper designed an ISSA-XGBoost algorithm for the identification and interception of fraudulent numbers. Experimental analyses of real data revealed that the recognition effectiveness for fraudulent numbers was improved to some extent by applying SMOTE and feature filtering techniques. Moreover, compared to parameter optimization methods like BO, the ISSA-XGBoost algorithm demonstrated better parameter optimization results and achieved superior recognition outcomes. The precision reached 0.945, while the F1 value reached 0.876. Additionally, the resumption rate obtained through a practical application over six months was 21.12%, further validating its effectiveness in fraudulent number recognition and interception. These results highlight the potential for further promotion and practical application of the ISSA-XGBoost algorithm.

Acknowledgments

This study was supported by Beijing Police College Organized Scientific Research: Research on Precise Deterrence Strategies for Telecommunications and Internet Fraud Victims from an Interactive Perspective (project number: 2023KYZZ01-4).

References

[1] M. A. Ali, M. A. Azad, M. P. Centeno, F. Hao, A. van Moorsel, "Consumer-facing technology fraud: economics, attack methods and potential solutions," *Future Generation Computer Systems*, vol. 100, pp. 408-427, 2019.

[2] M. M. Amin, A. Zainal, N. F. M. Azmi, N. A. Ali, "Feature selection using multivariate adaptive regression splines in telecommunication fraud detection," *IOP Conference Series Materials Science and Engineering*, vol. 864, pp. 1-6, 2020.

[3] H. O. Amuji, E. Chukwuemeka, E. M. Ogbuagu, "Optimal Classifier for Fraud Detection in Telecommunication Industry," *Open Journal of Optimization*, vol. 08, no. 1, pp. 15-31, 2019.

[4] S. Bhattacharya, S. S. Ramakrishnan, M. P. K. Reddy, R. Kaluri, S. Singh, G. T. Reddy, M. Alazab, U. Tariq, "A Novel PCA-Firefly Based XGBoost Classification Model for Intrusion Detection in Networks Using GPU," *Electronics*, vol. 9, no. 2, pp. 1-17, 2020.

[5] Y. C. Chang, K. T. Lai, S. C. T. Chou, W. C. Chiang, Y. C. Lin, "Who is the boss? identifying key roles in telecom fraud network via centrality-guided deep random walk," *Data Technologies and Applications*, vol. 55, no. 1, pp. 1-18, 2021.

[6] F. Y. Chin, C. A. Lim, K. H. Lem, "Handling leukaemia imbalanced data using synthetic minority oversampling technique (SMOTE)," *Journal of Physics: Conference Series*, vol. 1988, no. 1, pp.1-7, 2021.

[7] T. Fu, X. Tang, Z. Cai, Y. Zuo, Y. Tang, X. Zhao, "Correlation research of phase angle variation and coating performance by means of Pearson's correlation coefficient," *Progress in Organic Coatings*, vol. 139, pp. 1-9, 2020.

[8] V. Goodarzimehr, F. Omidinasab, N. Taghizadieh, "Optimum design of space structures using hybrid particle swarm optimization and genetic algorithm," *World Journal of Engineering*, vol. 20, no. 3, pp. 591-608, 2023.

[9] S. Ji, J. Li, Q. Yuan , J. Lu, "Multi-range gated graph neural network for telecommunication fraud detection," in *International Joint Conference on Neural Networks (IJCNN'20)*, Glasgow, UK, pp. 1-6, 2020.

[10] R. Kirubahari, S. M. J. Amali, "An improved restricted Boltzmann Machine using Bayesian Optimization for Recommender Systems," *Evolving Systems*, vol. 15, no. 3, pp. 1099-1111, 2024.

[11] G. Li, Y. Wen, S. Zhong, "Research on the Detection Countermeasures of Telecommunication Network Fraud Based on Big Data for Killing Pigs and Plates," *Journal of Robotics*, vol. 2022, no. Pt.1, pp. 1.1-1.11, 2022.

[12] J. Liu, Z. Wang, "A Hybrid Sparrow Search Algorithm Based on Constructing Similarity," *IEEE Access*, vol. 9, pp. 117581-117595, 2021.

[13] C. Ma, J. Hu, J. Gao, S. Liu, L. Yan, "Research and analysis of the potential correlation of fraud types in telecommunications fraud data," *Journal of Nonlinear and Convex Analysis*, vol. 23, no. 10, pp. 2227-2236, 2022.

[14] A. Masrub, M. Alahemar, "SIM Boxing Problem: ALMADAR ALJADID Case Study," in *International Conference on Electrical Engineering (ICEE'20)*, Istanbul, Turkey, pp. 1-5, 2020.

[15] M. N. Pal, M. Banerjee, "Retinal vessel segmentation using a strip wise classification approach with grid search-based parameter selection," *International Journal of Computational Vision and Robotics*, vol. 12, no. 2, pp. 194-218, 2022.

[16] N. Ruan, Z. Wei, J. Liu, "Cooperative fraud detection model with privacy-preserving in real CDR datasets," *IEEE Access*, vol. 7, pp. 115261-115272, 2019.

[17] O. A. Starostenko, "Nature and methods of committing fraud using information-telecommunication technologies," *Bulletin of Udmurt University Series Economics and Law*, vol. 30, no. 4, pp. 576-582, 2020.
[18] Y. Wang, H. Chen, S. Liu, X. Li, Y. Hu, "Feature difference-aware graph neural network for telecommunication fraud detection," *Journal of Intelligent & Fuzzy Systems: Applications in Engineering and Technology*, vol. 45, no. 5, pp. 8973-8988, 2023.
[19] R. Yao, F. Wang, S. Chen, S. Zhao, "Assisting telecommunication fraud prediction: detect individuals carrying multiple phones based on trajectory data mining," in *Information Communication Technologies Conference (ICTC'20)*, Nanjing, China, pp. 158-165, 2020.
[20] L. Zhang, C. Wang, M. Fang, W. Xu, "Spectral Reflectance Reconstruction Based on BP Neural Network and the Improved Sparrow Search Algorithm," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E105.A, no. 8, pp. 1175-1179, 2022.

Biography

Hui You, born in June 1984, graduated from Beijing Normal University with a doctor's degree in January 2020. She is currently working at Beijing Police College as an associate professor. She is interested in cyber security and big data.

Tuo Shi, born in August 1988, graduated from Communication University of China with a doctor's degree in June 2018. She is currently working at Beijing Police College as an associate professor. She is interested in big data technology and intelligence analysis.

Security Management of Students' English Education Information based on Blockchain Technology

Ying Lei[1] and Nanchang Zeng[2]
(Corresponding author: Nanchang Zeng)
School of Applied Foreign Languages, Guangdong Polytechnic of Industry and Commerce[1]
Information Center, Guangdong Polytechnic of Industry and Commerce[2]
Guangzhou, Guangdong 510515, China
Email: zengnanc@hotmail.com
(Received July 30, 2023; Revised and Accepted July 28, 2024; First Online Aug. 17, 2024)

Abstract

This paper provides a concise introduction to blockchain and the Practical Byzantine Fault Tolerance (PBFT) algorithm. The PBFT algorithm was optimized and applied to the security management of students' English education information. Subsequently, simulation experiments were conducted on the blockchain-based management mode for students' performance in English education. The results demonstrated that this management mode enabled normal operations such as entry, query, modification, and deletion of students' education information. Additionally, this mode ensured the synchronization of backup data across other nodes. Moreover, the increase in consensus nodes led to a decrease in the throughput of the management mode, while an increase in the number of tasks resulted in an increase in the throughput. The management mode, which adopted the improved PBFT algorithm, achieved a higher throughput. Furthermore, this algorithm guaranteed stable throughput and resilience against third-party attacks.

Keywords: Blockchain; Education Information; PBFT

1 Introduction

With the continuous expansion of higher education and the ongoing reform and enhancement of teaching modes, the size of higher education institutions has significantly increased, resulting in a substantial rise in student enrollment [1]. This increase in student numbers has made college management more challenging. Throughout the college education process, it is necessary to input each student's relevant information, including academic qualifications, majors, grades, etc., into the education information management system for storage [3].

The educational information of students is important information about their education and assessment. The information is a vital resume for future job applications and a significant reference for enterprises. Therefore, ensuring proper recording, security, and authenticity of students' education information is essential. In the traditional education information management system, student information is centralized, with data stored in local databases. However, this centralized storage approach has inherent flaws, such as excessive administrator privileges, difficult tracking for data modifications, and challenges in data recovery after loss. However, the emergence of blockchain technology offers a new and secure way to store students' educational information [11].

Related studies have explored the application of blockchain in various domains. For instance, Zhang *et al.* [13] proposed a blockchain-based decentralized supply chain system, ensuring secure information sharing and reliable product origin records without fully trusted intermediaries. Yi [12] introduced blockchain in logistics security to protect personal privacy and constructed a logistics blockchain model and verifying its efficiency and security in a distributed platform.

Mao *et al.* [9] proposed a blockchain-based credit evaluation system and assessed its performance. This paper briefly introduces blockchain technology and the Practical Byzantine Fault Tolerance (PBFT) algorithm. The algorithm was optimized and applied to securely manage students' English education information. Additionally, simulation experiments were conducted on the blockchain-based management model for students' English education performance.

2 Blockchain-based Student Education Information Management

2.1 Blockchain

The traditional education information management system for student data storage is centralized, granting excessive authority to system administrators and making it susceptible to data tampering. Furthermore, the system cannot trace data operations [8]. Additionally, if the system is attacked and data loss occurs without any data backup, it would be difficult to recover the lost data. However, the emergence of blockchain technology addresses these issues. As a distributed ledger technology, blockchain links data blocks in chronological order using cryptographic algorithms, forming a continuous chain of data [14]. Each data block in the chain contains transaction information, i.e., educational information operation data in this paper, a timestamp, and a pointer to the previous data block. When blockchain is applied to student information management, student-related data is stored decentralized across multiple nodes to prevent data loss in case of a single node failure. The data is encrypted and verified through cryptographic algorithms during transmission within the blockchain [6]. The consensus mechanism maintains consistency among multiple nodes, ensuring data integrity and preventing tampering. Additionally, the presence of timestamps enables the traceability of data operations.

2.2 Consensus Algorithm

The consensus algorithm is a mechanism used to maintain the consistency of nodes in the blockchain. The principle of the consensus algorithm can be understood as voting to verify the validity of a new block, and it can be uploaded to the blockchain after it passes. The PBFT consensus algorithm [10] is an algorithm that makes the blockchain nodes reach consensus in a non-trusted environment. It can ensure that the blockchain reaches a consensus when the relationship between the total number of nodes (n) and the number of unreliable nodes (f) is $3f+1 \leq n$. The basic process is as follows. The client initiates a request to the master node, and the master node broadcasts it to the other nodes of the blockchain for verification. Once the data passes the verification of most nodes, it is added to the blockchain and backed up by other nodes. However, this consensus algorithm requires the establishment of a master node, and there is a possibility that the master node may crash. Therefore, view switch is required to change the master node. The view switch process temporarily disrupts the consensus response of the blockchain and results in additional communication traffic.

Therefore, this paper proposes several improvements to the PBFT algorithm. The nodes are categorized into consensus, candidate, and ordinary nodes. Consensus nodes are responsible for verifying the consensus of the upload request [7], while candidate nodes respond to read requests for stored data and perform initial verification of the upload request. The candidate node responds to the read request of the stored data and carries out the preliminary verification of the upload request. In contrast, the ordinary node is only responsible for the transaction, i.e., backing up the data. The specific steps are as follows.

1) The client submits a "transaction request" to the candidate node.

2) Candidate node α that receives the "transaction request" verifies its validity and broadcasts it to other candidate nodes after it passes. If the "transaction request" is to read data, the candidate node returns the result to the client [5], and the client receives $2f+1$ results. The result is recognized as the data reading result. If the "transaction request" is to write data, other candidate nodes will verify its validity and return the result to candidate node α.

3) When candidate node α receives $2f+1$ results that the "transaction request" passes the verification, it generates a multi-signature. Then, it broadcasts the proposal with the multi-signature and the transaction content to the consensus node.

4) Upon receiving the proposal, the consensus node first performs multi-signature verification. The transaction content is stored in the local transaction pool if the verification is passed. However, if the verification fails, the proposal is discarded, and the processing result is returned to the client.

5) When the number of transaction proposals in the local pool reaches a certain number, master node c generates a pre-preparation message and broadcasts it along with the transaction proposal to other consensus nodes [4].

6) When other consensus node d receives the pre-preparation message, the validity of the message is verified, and the result of this verification is recorded in the node's local list. Then, the consensus node examines whether the operation logic recorded in the block is standard; if it is not normal, it stops uploading and returns to the client, and if it is normal, it generates a preparation message to broadcast to other consensus nodes.

7) When other consensus node d receives the preparation message, it records it in the local log and list. When d receives $2f+1$ preparation messages, it generates a commit message and broadcasts it to other consensus nodes.

8) Other consensus node d receives the commit message and verifies it. When d receives $2f+1$ valid commit messages, it broadcasts the new block to the candidate nodes as well as to the ordinary nodes and then adds it into the node's local blockchain.

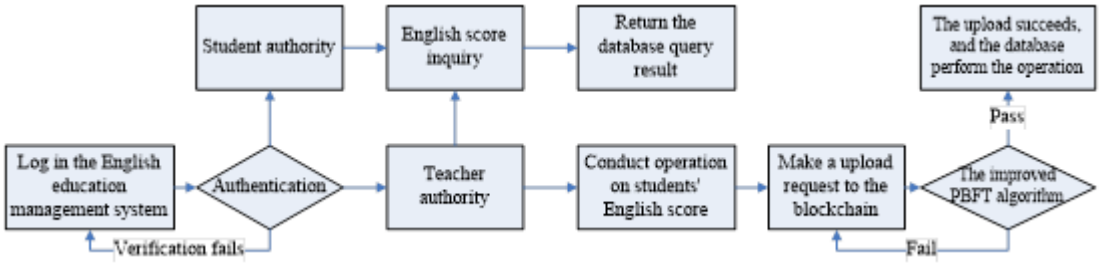


Figure 1: The security management process of students' English performance information combined with blockchain

2.3 Security Management of Students' English Education Information

Blockchain is applied to the security management of students' English education information. There are various kinds of educational information related to students, so this paper takes the security management of students' English performance as the object. The security management process of students' English performance combined with blockchain is shown in Figure 1 [15], and its specific steps are as follows.

1) The user initiates a login process in the English education management system to validate the user's identity. If the authentication fails, the user is redirected to the login screen.

2) When the identity is confirmed as a student, the user's permission is only to query English scores. When querying the English score, the processing method is the same as that of the transaction request of reading described previously. The user initiates a "read request" to the candidate node, and the candidate nodes execute the request in the respective local database and return the read result to the user. When the user receives $2f+1$ same results, the return result is the final read result.

3) If the identity is confirmed as a teacher, the user is granted privileges to query students' English scores and perform operations on the information in the database, such as entry, modification, and deletion. The process of querying students' English score information is the same as that of student users. When teachers perform operations on students' English performance information, the operation details are packaged into information blocks, and an upload request is sent to the blockchain.

4) For the upload request of the operation content, the improved PBFT consensus algorithm is utilized to validate the request. If the validation fails, the process returns to Step 3. However, if the validation passes, the upload request is deemed successful [**?**]. Each node's local blockchain adds a new block, and the local data in the node executes the operation content contained in the new block, i.e., entering, modifying, or deleting the English scores in the database.

3 Simulation Experiments

3.1 Experimental Environment

Simulation experiments were conducted using servers from a lab. The blockchain network was established using Ethernet's virtual machines, which acted as the blockchain nodes. Specifically, server 1 functioned as the local database node, server 2 served as the client, and server 3 was designated as the third-party attacker. The maximum number of virtual machine nodes was 50. When configuring the consensus nodes, server 1 was always given priority as the primary consensus node. The remaining consensus nodes, candidate nodes, and normal nodes were set up according to the specific requirements of the tests.

3.2 Experimental Setup

1) Functional testing of the blockchain-based information management mode for English performance

   Step 1: Server 2 entered the students' English performance information as a teacher.

   Step 2: Server 2 queried the entered English performance information as a student and teacher respectively.

   Step 3: Server 2 changed the entered English score as a teacher.

   Step 4: Server 2 deleted the entered English score as a teacher.

   During the execution of each of the above steps, the databases of other nodes were monitored in the background.

2) Performance impact of traditional and improved PBFT consensus algorithms on the blockchain-based information management mode for English performance The number of consensus nodes was set to 3, 5, 7, 9, and 11, while the number of operation requests for uploading was set to 300, 600, 900, 1,200, and

Table 1: Functional test results of blockchain-based information management of English performance

Steps	Database monitoring results of server 1	Database monitoring results of the other nodes
Step 1	Students' English performance were successfully entered.	Students' English scores were successfully entered and kept aligned with server 1.
Step 2	Both students and teachers obtained the same query results.	-
Step 3	Students' English scores were successfully modified.	Students' English scores were modified, and the final scores were consistent with that in server 1.
Step 4	Students' English scores were successfully deleted.	Students' English scores were deleted.

1,500, respectively. The performance of the mode under the traditional and the improved PBFT consensus algorithms was tested, with varying numbers of consensus nodes and operation requests. Throughput was used as the performance metric.

3) Security testing The number of consensus nodes was 7, and the number of operation requests for uploading was 900. During uploading, after five consensus rounds, two consensus nodes were turned into failure nodes. The throughput of the management mode under the traditional and improved PBFT consensus algorithms was tested. Additionally, the English scores stored on server 1 were tampered with through server 3 before and after the consensus nodes became failure ones. Server 2 then queried the English scores.

3.3 Experimental Results

The information entry, query, modification, and deletion functions of the blockchain-based mode were tested, and the results are shown in Table 1. The table shows that the entry, query, modification, and deletion functions of the mode were functioning normally. Additionally, except for the "query" operation, which did not affect the database, when the other three operations (entry, modification, and deletion) were performed on the database of server 1, the databases of the other nodes also underwent the same adjustments.

The throughput of the blockchain-based management mode under the two consensus algorithms when facing different numbers of consensus nodes and tasks is shown in Figure 2. It can be observed that as the number of consensus nodes increased, the throughput of the mode decreased under both consensus algorithms. This is because the increase in the number of consensus nodes leads to more consensus rounds, reducing the actual data transmission efficiency. Furthermore, with the same number of consensus nodes, more tasks resulted in a higher throughput. Additionally, the throughput of the mode utilizing the improved algorithm was higher than the traditional algorithm.

The change in throughput of the management mode in the face of node failure under the two consensus algorithms is shown in Figure 3. It can be observed that when the consensus round was five and the consensus nodes became failure ones, the management mode under the improved algorithm experienced a significant drop in throughput. However, the throughput then returned to normal in the following consensus round. Importantly, when the throughput stabilized, the mode utilizing the improved algorithm maintained a higher throughput than the traditional algorithm.

Table 2 shows the query results from server 2 after a third-party attacker tampered with the English scores stored on server 1 before and after the node failure under the two algorithms. It was found that the third-party attacker's tampering with the local database failed under both the traditional and improved algorithms and server 2 obtained the accurate query data.

4 Conclusion

This paper provides a brief introduction to blockchain and the PBFT consensus algorithm. The PBFT algorithm was optimized and and applied to enhance the security management of students' English education information. Additionally, simulation experiments were performed on a blockchain-based management mode for students' English education performance. The entry, query, modification, and deletion functions in the mode were all normal. All functions, except for query, synchronized the databases of other nodes with the adjustments made on server 1. Increasing the number of consensus nodes decreased the throughput of the management mode. Increasing the number of tasks enhanced the throughput. The management mode with the improved PBFT consensus algorithm achieved higher throughput. The improved PBFT consensus algorithm ensured quick stabilization of throughput in the presence of failure node. Both consensus algorithms effectively resist third-party attacks on the blockchain.

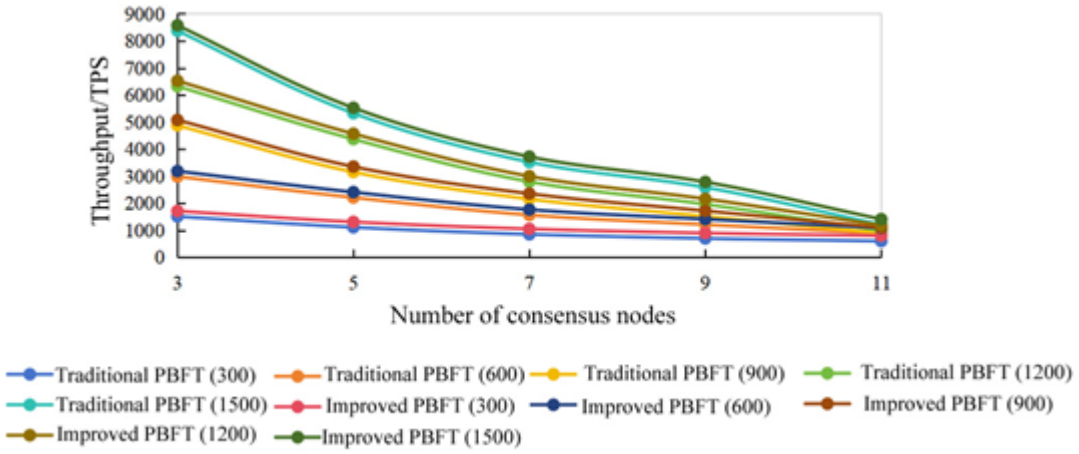


Figure 2: Throughput of the management mode when facing different numbers of consensus nodes and tasks under different consensus algorithms

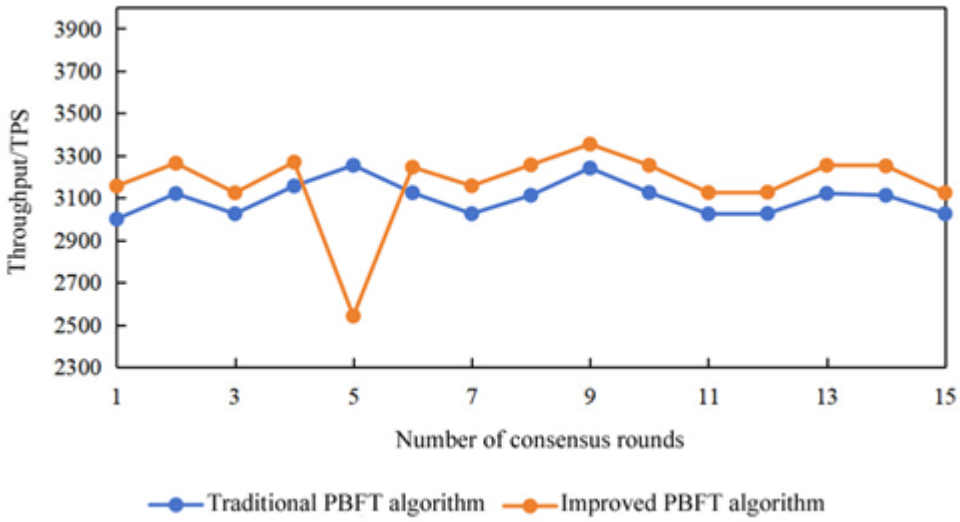


Figure 3: Throughput of the management mode when facing failure nodes under different consensus algorithms

Table 2: Results of third-party attacks before and after node failure under different consensus algorithms

	Third-party attack results before node failure	Third-party attack results after node failure
The traditional PBFT consensus algorithm	No change in English scores queried before and after the attack	No change in English scores queried before and after the attack
The improved PBFT consensus algorithm	No change in English scores queried before and after the attack	No change in English scores queried before and after the attack

References

[1] M. Cherepniov, "Decentralized scheme for secure database creation and storage," *International Journal of Open Information Technologies*, vol. 8, pp. 109-115, 2020.

[2] P. Dangayach, "Pharmaceutical supply chain tracking system based on blockchain technology and radio frequency identification tags," *International Journal of Business Research*, vol. 19, no. 4, pp. 37-44, 2019.

[3] R. Jabbar, N. Fetais, M. Kharbeche, M. Krichen, K. Barkaoui, M. Shinoy, "Blockchain for The Internet of Vehicles: How to use Blockchain to secure Vehicle-to-Everything (V2X) Communication and Payment?," *IEEE Sensors Journal*, vol. 21, no. 14, pp. 15807-15823, 2021.

[4] S. S. Kamble, A. Gunasekaran, M. Goswami, J. Manda, "A systematic perspective on the applications of big data analytics in healthcare management," *International Journal of Healthcare Management*, vol. 12, no. 3, pp. 226-240, 2019.

[5] L. Koh, A. Dolgui, J. Sarkis, "Blockchain in transport and logistics – paradigms and transitions," *International Journal of Production Research*, vol. 58, no. 7, pp. 2054-2062, 2020.

[6] L. Kong, "A fast encryption method for enterprise financial data based on blockchain," *Web Intelligence and Agent Systems*, vol. 20, no. 2, pp. 133-141, 2022.

[7] M. Li, S. Shao, Q. Ye, G. Xu, G. Q. Huang, "Blockchain-enabled logistics finance execution platform for capital-constrained E-commerce retail," *Robotics and Computer Integrated Manufacturing: An International Journal of Manufacturing and Product and Process Development*, vol. 65, pp. 1-14, 2020.

[8] Z. Ma, W. Huang, H. Gao, "Secure DRM Scheme Based on Blockchain with High Credibility," *Chinese Journal of Electronics*, vol. 27, no. 05, pp. 141-152, 2018.

[9] D. Mao, F. Wang, Z. Hao, H. Li, "Credit Evaluation System Based on Blockchain for Multiple Stakeholders in the Food Supply Chain," *International Journal of Environmental Research & Public Health*, vol. 15, no. 8, pp. 1-21, 2018.

[10] H. Wu, N. Su, C. Ma, P. Liao, D. Li, "A privacy protection solution based on NLPCA for blockchain supply chain financial system," *International Journal of Financial Engineering*, vol. 07, no. 3, pp. 1-22, 2020.

[11] Y. Ye, Z. Ren, X. Luo, J. Zhang, W. Wu, "Garou: An Efficient and Secure Off-Blockchain Multi-Party Payment Hub," *IEEE Transactions on Network and Service Management*, vol. 18, no. 4, pp. 4450-4461, 2021.

[12] H. Yi, "A secure logistics model based on blockchain," *Enterprise Information Systems*, vol.2019, no. 4, pp. 1-17.

[13] G. Zhang, Z. Yang, W. Liu, "Blockchain-based decentralized supply chain system with secure information sharing," *Computers & Industrial Engineering*, vol. 182, pp. 1-9, 2023.

[14] C. D. Zhu, "Fraud detections for online businesses:a perspective from blockchain technology," *Financial Innovation*, vol. 2, no. 1, pp. 256-265, 2016.

[15] Y. Ziegler, V. Uli, F. Keller, A. Kramer, "The Impact of Blockchain Networks on Logistics: an Update," in *mednarodna konferenca o razvoju organizacijskih znanosti*, 2019.

Biography

Ying Lei, born in August 1978, graduated from Jinan University with a PhD degree. She is working in Guangdong Polytechnic of Industry and Commerce as a business English teacher. Her research focuses on English education, English translation and communication of Chinese culture.

Nanchang Zeng, born in August 1991, graduated from Jiaying University in 2014 with a Bachelor of Engineering. He is workfing in Guangdong Polytechnic of Industry and Commerce as an senior engineer. His research focuses on information technology theory and practice.

Spatial Database Indexing Optimization Method Based on NVD-GkNN Algorithm

Huili Xia

(Corresponding author: Huili Xia)

College of Big Data and Artificial Intelligence, Zhengzhou University of Economics and Business
Zhengzhou 451191, China
Email: xiahuili80@163.com

(Received July 11, 2023; Revised and Accepted May 20, 2024; First Online Aug. 17, 2024)

Abstract

Spatial data query methods are essential in spatial databases and intelligent transportation. The nearest neighbor query is one of the most widely used methods for spatial data query. However, the nearest neighbor query and its variants suffer from low query efficiency, inefficient processing, and high computational cost. Therefore, to address the shortcomings of the nearest neighbor query and its variants in the obstacle space and road network environments, the study proposes to add Voronoi diagrams and grid Voronoi diagrams to these methods, using the filtering function of Voronoi diagrams to decrease the quantity of queried points and enhance the query efficiency. The results show that the Network Voronoi Diagram-group k Nearest Neighbor (NVD-GkNN) algorithm in the road network environment combined with the grid Voronoi diagram has a CPU running time as long as the Timing Algorithm (TA) and Incremental Euclidean Restriction (IER) CPU running time. The Voronoi graph-based nearest neighbor and its variants are designed to improve the query efficiency, which also provides a way to improve the efficiency of other spatial data query methods. The research aims to improve the query efficiency of spatial databases, maintain data security of databases, and assist in network security.

Keywords: Database; Indexing; kNN; Road Network Environment; Voronoi Diagram

1 Introduction

With the technology boost in China, the fields of geographic information systems and intelligent transportation systems are also developing rapidly. Spatial data query methods play an important role in these fields [3]. When dealing with complex spatial data, traditional data query methods are not effective, so new data query methods need to be studied [13]. Nearest neighbor query is one of the most widely used techniques for spatial data query, and many variants of it have been developed, like k-nearest neighbor query and reverse nearest neighbor query [19]. However, most of these query methods suffer from low query efficiency, inefficient processing and high computational cost. In addition, nearest neighbor query methods can be divided into two types according to the data environment, one is based on the Euclidean space and the other is based on the road network environment, and different query methods are applicable to different data environments [14]. Voronoi diagrams are mainly used to partition the space, and their filtering function can decrease queried points' quantity and enhance the data query efficiency, which is effective in spatial query and multimedia database [1]. The Voronoi diagram can be used to decrease query points' quantity and enhance the data querying. In addition, Voronoi diagrams can also reduce development and maintenance costs. Based on the above background, the research innovatively combines Voronoi diagrams and grid Voronoi diagrams with nearest neighbor query and its variant query, and then applies these algorithms to data query in obstacle space and road network environments respectively. The Network Voronoi Diagram-group k Nearest Neighbor (NVD-GkNN) algorithm improves the query efficiency of group k-nearest neighbor in road network environment, and the reverse nearest neighbor in road network environment combined with grid Voronoi diagrams is also enhanced. The Network Voronoi Diagram-Reverse Nearest Neighbor (NVD-RNN) algorithm performs better in certain situations. The k Nearest Neighbors-Obstacle (kNN-Obs) algorithm combined with Voronoi diagrams in obstacle space can improve its own query efficiency under certain circumstances. The study aims to enhance the query efficiency and diminish the query cost of nearest neighbor queries and their variants. The research aims to improve the query efficiency of spatial databases, maintain data security of databases, and assist in network security.

2 Related Works

The querying of data in spatial databases is a popular research area that has been explored by many researchers. Eiter and other researchers used local dynamic maps to set up spatial flow query responses in collaborative intelligent transportation systems, which could complete data flow queries oriented towards semantic concepts and spatial relationships. In addition, the semantic enhancement method of this study was conducted in ontology based query responses. The experimental results showed that the method designed by the research institute is feasible and can improve the efficiency of queries [6]. Sarode and Reshmi proposed a group search optimization algorithm based on neural networks to study data aggregation models, and developed a query based data aggregation model based on this algorithm. The query ranking in this study was determined based on latency and throughput. The research results indicated that the data aggregation model could improve throughput while reducing latency, and its performance was significantly better than traditional data aggregation models [16]. Jakob and Guthe proposed to use GPUs for exact domain problems in point clouds in order to optimise the query performance of kNNs on point clouds. The results showed that this approach gave real-time capabilities for large queries on very large point clouds, and the speed of the queries was greatly improved [9]. Chen et al. proposed a HorseIR-based approach to better querying of databases, combining database queries with compiled code based on dynamic arrays. Experimental results proved that this method was usable for database queries [4]. Scholars such as Shi and Yang proposed a spectral clustering based method to classify the climate in China. This method first analyzed the correlation between meteorological variables, and then constructed a similarity matrix graph based on k-nearest neighbor and sparse subspace representations. Finally, the study used a method of determining the number of clusters to conduct sensitivity analysis on different parameters. The research results showed that this classification method has high accuracy [18]. Experts such as Jang et al. proposed an input variable initialization method based on the k-nearest neighbor method to achieve the optimal input of neural networks. This method required finding the input that made the output very close to the target output and processing it as the initial input variable. The experimental results indicated that the method designed by the research institute is significantly superior to random initialization [10].

Shi and other researchers designed a flutter identification method based on enhanced k-nearest neighbors to identify flutter. This method compared the information of different sensors based on a large number of experiments, and extracted the features. Finally, the enhanced k-nearest neighbor method was used to identify chatter under different cutting conditions. The experimental results showed that the method designed by the research institute could effectively identify flutter with high accuracy [17]. Jg et al. proposed a generalised fragment allocation strategy to avoid the current database fragment allocation strategy that was prone to low-quality allocation plans, which implemented multiple candidate allocation schemes based on cost through PostgreSQL improved genetic algorithms evaluation. The results of the study showed that the performance of this strategy was improved by a factor of 2-4 with good robustness and scalability [11]. Ding and other experts proposed an R-KWS method in view of the evaluation of combinatorial candidate networks in order to improve the efficiency of existing database query methods, which could share the overlapping parts between candidate networks. The experiment indicated that this method can enhance the query efficiency without losing the quality of the query results [5]. Gulzar et al. proposed an optimised and incomplete framework for Skyline that simplified the Skyline process for databases with missing data, in order to avoid the situation of Skyline query methods' error when there was missing data in the database. The results showed the superiority of the framework and the number of control tests required to retrieve the skyline [8]. In order to improve the performance of distance metric learning, scholars such as Ruan et al. designed a nearest neighbor search model for distance metric learning. This model could construct metric optimization constraints by searching for the optimal nearest neighbor number. In addition, the study also designed a k-free nearest neighbor model based on a support vector machine solver. The experimental results showed that the performance of the nearest neighbor search model designed by the research institute for distance metric learning was significantly superior to existing baseline methods [15].

In summary, there have been many studies on spatial database data query methods, but most of them suffered from low query efficiency, less efficient processing and high computational cost. Based on these problems, the study innovatively combines Voronoi diagrams and nearest neighbor queries, aiming to make up for the shortcomings of existing methods in different data environments.

3 Spatial Database Optimization Based on NVD-GkNN Algorithm

3.1 Optimal Design of Spatial Databases Under Different kNN Algorithms

The Nearest Neighbor (NN) method plays an important role in spatial data querying [20]. The traditional k-Nearest Neighbor (kNN) algorithm uses the second-order Ming's distance as a measure of similarity between samples. By noting the observations of samples i and samples j as $x_i = (x_{i1}, x_{i2}, \cdots, x_{ip})$ and $x_j = (x_{j1}, x_{j2}, \cdots, x_{jp})$, where each sample has a different variable, the Ming's

distance is described in Equation (1).

$$d_{ij}(q) = (\sum_{k=1}^{p} |x_{ik} - x_{jk}|^q)^{1/q} \tag{1}$$

In Equation (1), d_{ij} serves as the range between samples i and samples j, q represents the order, k denotes the first k variable for each sample, and k takes values in the range $[1, p]$. Depending on the value of q, the Ming's distance can be classified as Manhattan distance, Euclidean distance and Chebyshev distance [12]. When the value of q is 1, $1/q$ is 1. The first-order Ming's distance can be called the Manhattan distance, as shown in Equation (2).

$$d_{ij}(1) = \sum_{k=1}^{p} |x_{ik} - x_{jk}| \tag{2}$$

When the value of q is 2, $1/q$ is $1/2$ and the second order Minkowski distance can be called the Euclidean distance, as in Equation (3).

$$d_{ij}(2) = \sqrt{\sum_{k=1}^{p} (x_{ik} - x_{jk})^2} \tag{3}$$

The change in the Ming's distance is greater when the q value $\to \infty$. At this point the Ming's distance has been transformed into the Chebyshev distance, as in Equation (4).

$$\begin{aligned} d_{ij}(\infty) &= \lim_{p\to\infty} (\sum_{k=1}^{p} |x_{ik} - x_{jk}|^q)^{1/q} \\ &= \max_{1\le k\le p} |x_{ik} - x_{jk}| \end{aligned} \tag{4}$$

The traditional kNN has relatively good query performance, but when the number of samples is relatively large, it suffers from computational complexity and memory consumption problems. Based on these problems, the study combines a kNN with optimised weights on the basis of the third-order Ming's distance and applies it to the filling of missing values. Equation (5) shows the details.

$$\hat{x}_{ik}^m = \sum_{j=1}^{K} \frac{D_{ij}^{-1}}{\sum_{V=1}^{K} D_{iv}^{-1}} x_{jk} \tag{5}$$

In Equation (5), $x_i = (x_i^c, x_i^m)$ is the vector to be filled, x_i^c represents the complete part of the vector, x_i^m represents the missing part of the vector, D_{ij} serves as the range in the vector i and the vector j, then $\hat{x}_{ik}^m$ is the filled value of x_i^m. When the samples are close to each other, the fill value obtained by the kNN algorithm satisfying and the algorithm needs to be improved, as in Equation (6).

$$x_{ik}^m = \frac{K}{(K-1)^2} \sum_{j=1}^{K} (1 - \frac{D_{ij}}{\sum_{V=1}^{K} D_{ij}} x_{jk} \tag{6}$$

In Equation (6), x_{ik}^m denotes the new filled value, D_{ij} is the third-order Ming's distance formula.

$$D_{ij} = d_{ij}(3) = (\sum_{k=1}^{p} |x_{ik} - x_{jk}|^3)^{1/3}$$

K denotes the number of similar samples. The Euclidean distance used in the traditional kNN algorithm does not work well in mixed attribute data, compared to grey correlation analysis which is more appropriate. It replaces the Euclidean distance with a grey distance. In addition, to deal with missing values in mixed attributes, a grey weighted kNN filling method combining iterative kNN and grey distances can be used. Before carrying out grey correlation analysis on the data, the data needs to be pre-processed with datacentering, as in Equation (7).

$$\begin{cases} \hat{x_{ij}} = x_{ij} - \bar{x_i} \\ \hat{x_{ij}} = x_{ij} - \bar{x_j} \end{cases} \tag{7}$$

In Equation (7), $\hat{x_{ij}} = x_{ij} - \bar{x_i}$ is sample-centered, $\hat{x_{ij}} = x_{ij} - \bar{x_j}$ is attribute-centered, i takes the values of $[1, p]$ and j takes the values of $[1, p]$. x_{ij} and $\hat{x_{ij}}$ each represent the value of the ith sample in the jth attribute before and after standardisation, $\bar{x_i}$ represents the mean of all attributes in the ith sample, and $\bar{x_j}$ represents the mean of all samples within the jth attribute.

$$\begin{cases} \hat{x_{ij}} = \frac{x_{ij} - \bar{x_i}}{\sqrt{\sum_{j=1}^{q} (x_{ij} - \bar{x_i})^2}} \\ \hat{x_{ij}} = \frac{x_{ij} - \bar{x_j}}{\sqrt{\sum_{i=1}^{q} (x_{ij} - \bar{x_j})^2}} \end{cases} \tag{8}$$

Equation (8) shows the outlier normalisation, which is done by dividing the data normalised by the outlier after datacentering. Equation (9) is data regularisation, which is normalised by the standard deviation (SD).

$$\begin{cases} \hat{x_{ij}} = \frac{x_{ij} - \bar{x_i}}{\sqrt{\sum_{j=1}^{q} (x_{ij} - \bar{x_i})^2 / (p-1)}} \\ \hat{x_{ij}} = \frac{x_{ij} - \bar{x_j}}{\sqrt{\sum_{i=1}^{q} (x_{ij} - \bar{x_j})^2 / (q-1)}} \end{cases} \tag{9}$$

In addition, the data are pre-processed by means of Z-score normalisation, which starts with the mean and SD of the attributes, as shown in Equation (10).

$$\delta_m = \sqrt{(\sum_{i=1}^{n} (x_{im} - \bar{m})^2) / (n-1)} \tag{10}$$

In Equation (10), δ_m serves as the SD of the attribute m, n serves as the data points' quantity in the data set, x_{im} represents the value of the attribute m for the ith sample, and $\bar{m}$ represents the mean of the attribute m. The Min-Max normalisation method uses the linear variation of the original data as the entry point for data processing, one of which is the upper bound validity measure, as in Equation (11).

$$\acute{x_p}(j) = \frac{x_p(j) - \min_\forall x_i(j)}{\max_\forall x_i(j) - \min_\forall x_i(j)} \tag{11}$$

In Equation (11), $\max_\forall x_i(j)$ serves as the maximum value in the data, $\min_\forall x_i(j)$ represents the minimum value, and $\acute{x_p}(j)$ serves as the value of the sample p after pre-processing on the attribute j. When the data attribute is extremely small, a lower bound validity measure

is used, as in Equation (12).

$$x_p^{'}(j) = \frac{\max_{\forall} x_i(j) - x_p(j)}{\max_{\forall} x_i(j) - \min_{\forall} x_i(j)} \tag{12}$$

Equation (12) is called the lower bound validity measure, where very small data are also called cost-based indicators. In addition, the Min-Max standardisation method involves a moderate validity measure, as in Equation (13).

$$x_p^{'}(j) = \frac{|x_p(j) - x_{specified}|}{\max_{\forall} x_i(j) - \min_{\forall} x_i(j)} \tag{13}$$

In Equation (13), $x_{specified}$ is a pre-set value, as the upper and lower bound validity measures have relatively similar results, the moderate validity measure can also be used when selecting the data processing method. Voronoi diagrams are mainly used for spatial partitioning and the Voronoi diagram is constructed as shown in Figure 1 [7].

There are multiple methods for constructing Voronoi diagrams and one of them was used for the study. As shown in Figure 1(a), it is a part of a Voronoi diagram that must be partitioned from the region where p_k is located. The region is divided by the vertical bisector of p_{k+1} and p_j, and forms two parts. The boundary between the vertical bisector and the area where p_k is located intersects at two points and one of these points needs to replace the other. This behaviour is repeated until the dashed polygon shown in Figure 1(a) is formed. Figure 1(b) forms the Voronoi polygon of p_{k+1} by deleting the polygon vertices and the middle edge. A grid Voronoi diagram is also a type of Voronoi, an example of which is shown in Figure 2 [2].

Figure 2(a) shows the initial road network diagram with the generation points from p_1 to p_3 and the intersections of the road network from p_4 to p_{16}. It is connected via L. Figure 2(b) is a grid Voronoi diagram with each line having a Voronoi link set corresponding to the generation point. The links may be in the generating point V_{link} only, or in different V_{link}. When a link is in a different V_{link}, any one of the lines that does not pass through L may be connected to it.

3.2 Design of the GkNN Spatial Database Optimization Method Based on Voronoi Diagrams

In different data environments, nearest neighbor query methods can be separated into two kinds, the first is based on the Euclidean space and the other is based on the road network environment. For obstacle queries in Euclidean space, the study used the kNN query method k Nearest Neighbors-Obstacle (kNN-Obs) based on Voronoi diagram with the algorithm k NN-Obs $q, k, P, O)$, where q denotes the query point, P is the dataset, O represents the set of obstacles and k is the value of the kNN query. This method involves two aspects, the filtering process and the refinement process. The filtering process mainly generates a kNN candidate set, while the refining process refines the objects in it through Voronoi diagrams, and prunes the objects that do not meet the criteria. Before using the kNN algorithm, the originality of the data needs to be maintained, so the randomly distributed data is processed, as shown in Figure 3.

Figure 3(a) is a spatial data Euclidean diagram showing the distribution of the data in the spatial dimension. Figure 3(b) is a data redistribution diagram, showing the differences and characteristics of the data before and after the distribution. After redistribution of the data, the original local Voronoi diagram is then generated from this data, as shown in Figure 4(a). However, the original local Voronoi diagram is not particularly accurate, as it does not take into account the relationship between the server boundary points, so this Voronoi diagram needs to be optimised and the result is shown in Figure 4(b).

Both the multi-core technique and the optimised scanline algorithm are applied to the generation of the original local Voronoi diagram. In Figure 4(a), L1 to L11 are the original generation points and S1 to S3 are the servers. The Voronoi cells are the polygons where the original generation points are located and the boundary points of the S1 and S2 servers are L2 and L3. Optimisation of the original local Voronoi diagram requires finding the clustering centers above each server. Then the closest location data to the clustering centers needs to be find, which is used as the server's clustering result, and then it is finally optimised. For the nearest collar variant query in the road network environment, the study uses two types of methods. The first one is the NVD-RNN based on the grid Voronoi diagram, the algorithm is NVD-RNN $q, S, k)$, where q is the query point, S is the dataset and k is the value of the RNN query. The steps in this method involve filtering and refining, which results in a candidate set in which all possible points are stored as results. Refinement involves computing the points in the candidate set and finding the final result, then filtering and refinement are used recursively for each query. The Voronoi diagram of the grid involved in this method is shown in Figure 5.

In Figure 5, the grid graph NVP (p_2) has the query point q, so its 1-RNN is p_2. The resulting candidate set after filtering is $\{p_7, p_8, p_9\}$, and then the refinement process calculates the nearest distance between the query points q and $p + 7$, p_8, and p_9. The filtering process involves two while loops, the first of which terminates after all the data in the dataset are looped through once, while the second while loop requires all the data in the candidate set to be looped through once before aborting. The refinement process involves only one while loop and the data in it is finite. The loop terminates when this data has been looped through once. RNN is a variant of nearest neighbor, and its query is shown in Figure **??**(a).

In Figure **??**(a), NN(c) denotes the result obtained by NN query for point c and RNN(c) denotes the result obtained by RNN query for point c. In addition, the data used in both query methods are from a database or collection of data identified in advance. In Figure **??**(b), the

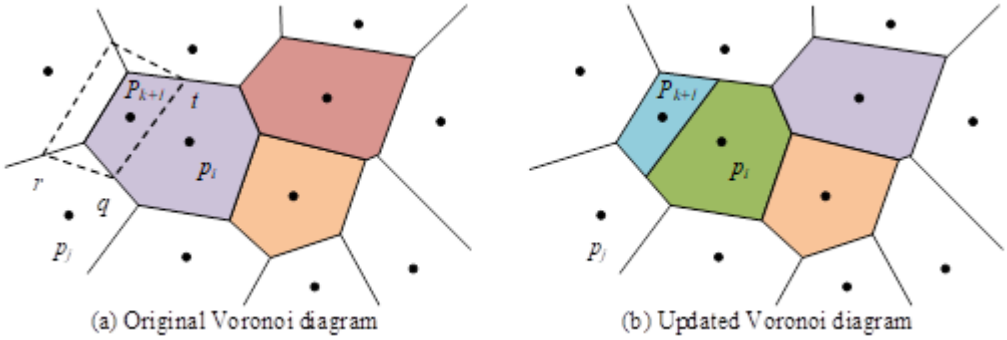


Figure 1: The process of updating the Voronoi diagram

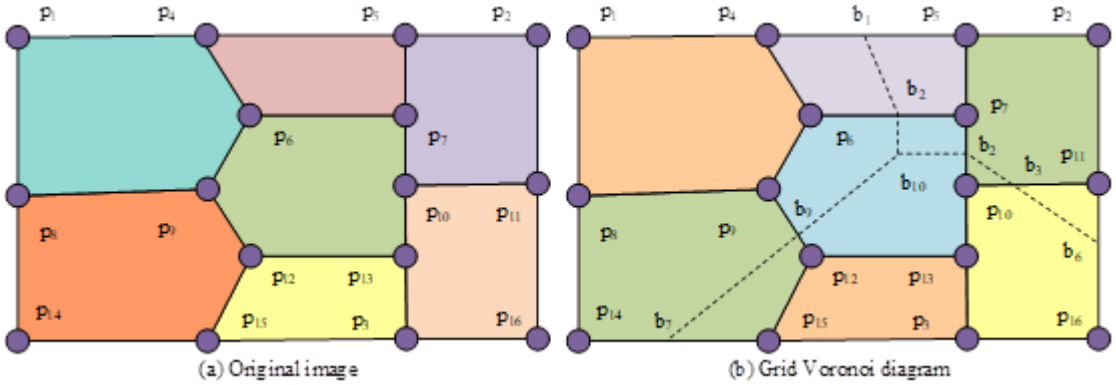


Figure 2: Instance of original and network Voronoi diagrams

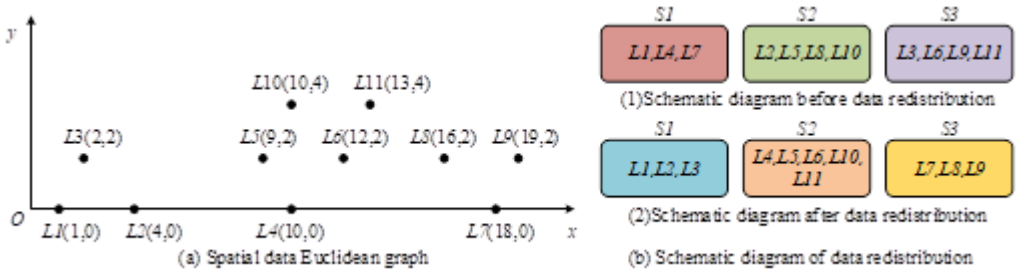


Figure 3: Spatial data Euclidean diagram and schematic diagram before and after data redistribution

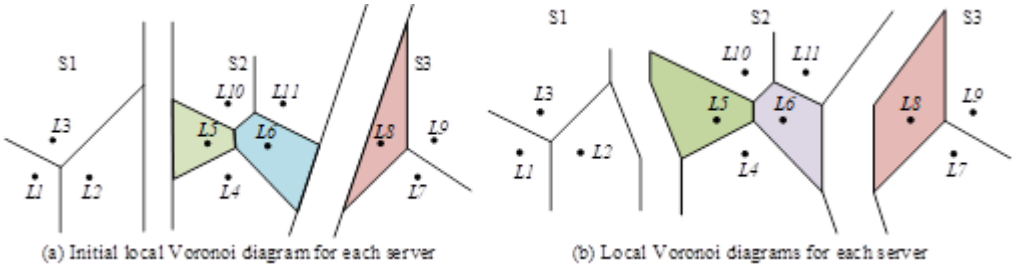


Figure 4: Original and improved local Voronoi diagrams for each server

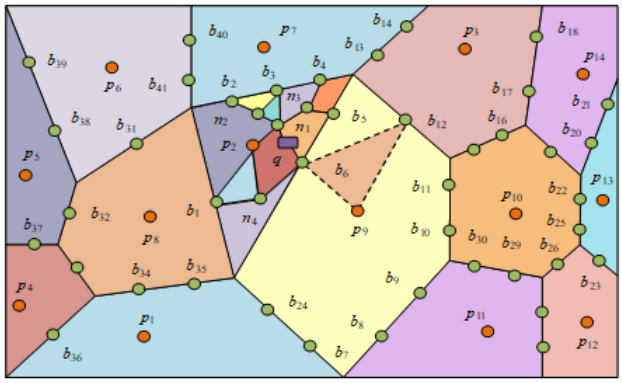


Figure 5: Instance of network Voronoi diagram

RNN query mainly involves a location query with overlapping parts among each circle. In the road network environment, the second nearest neighbor variant query is NVD-GkNN method. This method involves three aspects: dataset processing, filtering and refinement. The processing of the dataset is shown in Equation (14).

$$\begin{cases} \frac{\partial dist(q,Q)}{\partial x} = \sum_{i=1}^{n} \frac{x-x_i}{\sqrt{(x-x_1)^2+(y-y_i)^2}} = 0 \\ \frac{\partial dist(q,Q)}{\partial y} = \sum_{i=1}^{n} \frac{y-y_i}{\sqrt{(x-x_1)^2+(y-y_i)^2}} = 0 \end{cases} \tag{14}$$

(14) In Equation (14), q is the centre of mass of the dataset Q, (x, y) serves as the coordinates of q and (x_i, y_i) serves as the coordinates of any point in the dataset Q. However, when the value of n is taken to be 2, Equation (14) does not give a closed solution and only an estimated value can be obtained, and the centre of mass q can only be obtained as an approximation. Therefore, for obtaining a good estimate, the coordinates of the centre of mass need to be modified, as shown in Equation (15).

$$\begin{cases} x = x - \mu \frac{\partial dist(q,Q)}{\partial x} \\ y = y - \mu \frac{\partial dist(q,Q)}{\partial y} \end{cases} \tag{15}$$

In Equation (15), x represents the modified prime horizontal coordinates, y represents the modified prime vertical coordinates, and μ represents the step size. The algorithm for querying the GkNN of point sets in a road network environment is NVD-GkNN (Q, P, k), where Q represents the query point set, P represents the generated point set, and k is the value of the GkNN query. In addition, the study also analysed the impact of adding and removing points on GkNN. The algorithm for the effect of added points on GkNN is ADDNVD-GkNN (Q, P, k, w), where w is the added point in the generated point set, while the algorithm for the effect of deleted points on GkNN is DENVD-GkNN (Q, P, k, d), where d is the deleted point in the range $d(d \in P)$.

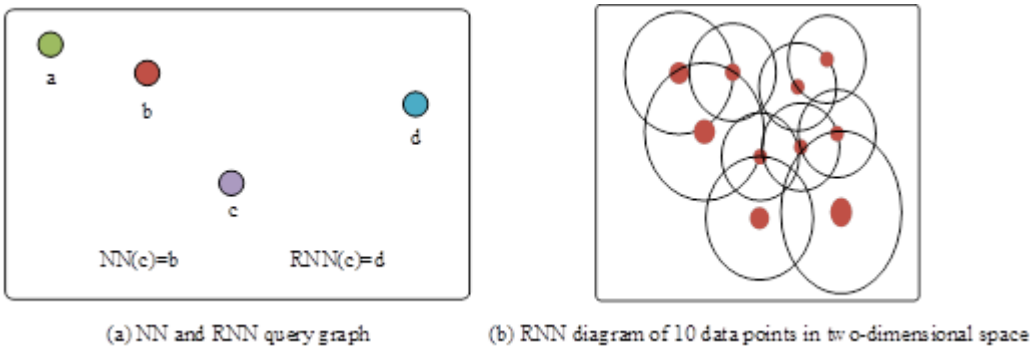
4 Analysis of Spatial Database Optimization Results Based on Voronoi Diagrams and kNN

4.1 Analysis of the Results of kNN-based Spatial Database Optimization

The study under the obstacle space mainly uses the kNN-Obs algorithm. The following contentl is a comparative analysis of the kNN-Obs algorithm and the Pruning algorithm (PostPruning) algorithm, in terms of both the k value and the obstacle dimension. Firstly, the impact of different k values on the Central Processing Unit (CPU) running time (RT) and page views was analysed, as shown in Figure 7.

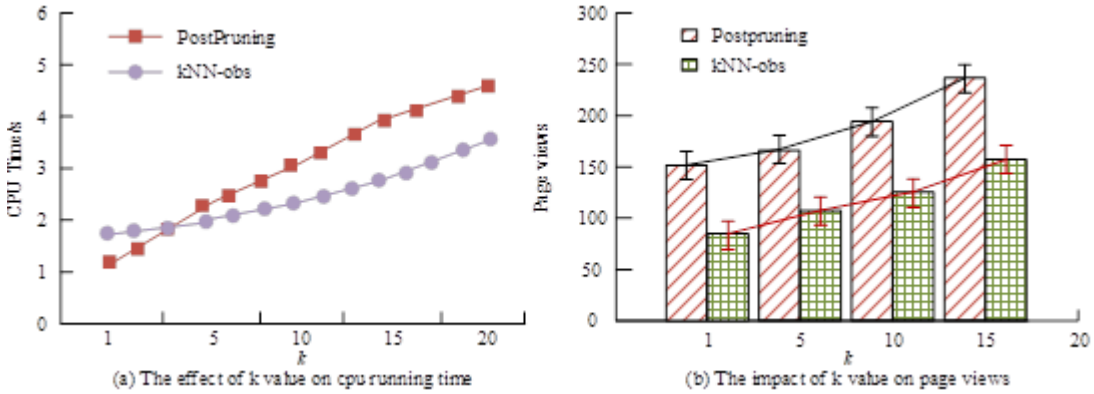
Figure 7(a) showed that the CPU runtime of the kNN-Obs and PostPruning algorithms increased with the value of k. The maximum value of the CPU runtime of the kNN-Obs algorithm was between 3s and 4s, and the minimum value was close to 2s. The maximum value of the CPU runtime of the PostPruning algorithm was between 4s and 5s, and the minimum value was around 1.1s. When the value of k was less than 4s, the PostPruning algorithm CPU runtime was shorter. When the value of k was greater than 4s, the kNN-Obs algorithm CPU runtime was shorter. Figure 7(b) illustrates that as the k values of the kNN-Obs and PostPruning algorithms gradually increased, their respective page views also increased. The maximum value of page views for the kNN-Obs algorithm was close to 150 and the minimum value was close to 100. The maximum value of page views for the PostPruning algorithm was close to 250 and the minimum value was around 150. The overall number of page views for the kNN-Obs algorithm was smaller than that for the PostPruning algorithm. Therefore, on the whole, the kNN-Obs algorithm outperformed the PostPruning algorithm. The number of different obstacles also had an impact on CPU runtime and page accesses, as shown in Figure 8.

As can be seen in Figure 8(a), the CPU runtime of the kNN-Obs and PostPruning algorithms gradually got larger as the quantity of obstacles increases. The kNN-Obs had a maximum CPU runtime of between 6s and 8s, with a minimum value close to 2s, while PostPruning had a maximum CPU runtime of between 10s and 12s, with a minimum value close to 4s. As can be seen in Figure 8(b), the number of page views for both the kNN-Obs and PostPruning algorithms increased as the quantity of obstacles increased. The kNN-Obs had a maximum value of around 175 page views and a minimum value of close to 100, while PostPruning had a extreme value around 150. In summary, the kNN-Obs algorithm outperformed the PostPruning algorithm. In order to better validate the performance of the kNN Obs algorithm, comparative analysis was conducted from the accuracy and F1 value of the algorithm. In the experimental environment, the CPU frequency was 2.0 GHz, the memory was 4GB, and



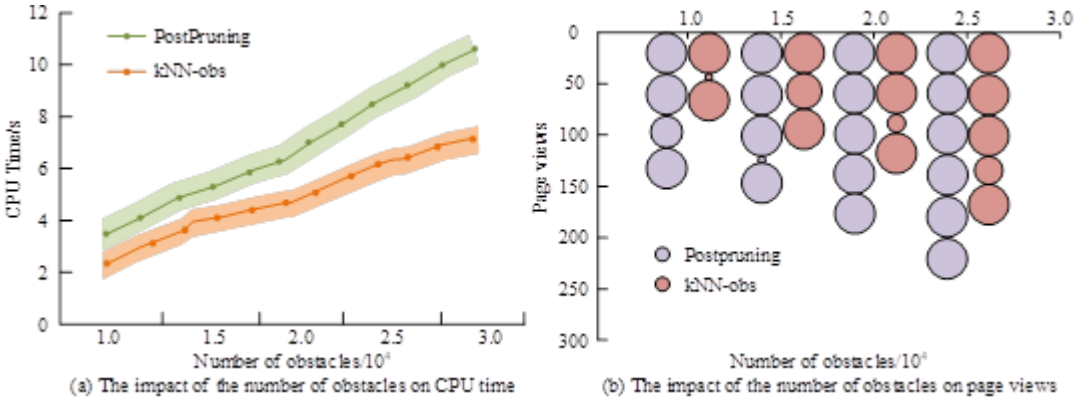
(a) NN and RNN query graph (b) RNN diagram of 10 data points in two-dimensional space

Figure 6: NN and RNN query graph and RNN graph for 10 data points in two dimensional space



(a) The effect of k value on cpu running time (b) The impact of k value on page views

Figure 7: The impact of k value on CPU runtime and page views



(a) The impact of the number of obstacles on CPU time (b) The impact of the number of obstacles on page views

Figure 8: The impact of the number of obstacles on CPU runtime and page views

the operating system was Windows 7. A total of 4 performance tests of the algorithm were conducted. The comparison of accuracy and F1 values between the kNN Obs algorithm and the PostPruning algorithm was shown in Table 1.

From Table 1, the maximum accuracy of the kNN Obs algorithm was 97.9%, while the minimum accuracy was 95.7%. The maximum accuracy of the PostPruning algorithm was 94.6%, and the minimum accuracy was 92.9%. The accuracy of the kNN Obs algorithm was generally about 3% higher than that of the PostPruning algorithm. The maximum F1 value of the kNN Obs algorithm was 0.981, and the minimum value was 0.964. The maximum F1 value of the PostPruning algorithm was 0.921, and the minimum value was 0.892. It can be seen that the F1 value of the kNN Obs algorithm was much higher than that of the PostPruning algorithm, which also indicated that the performance of the kNN Obs algorithm was superior to that of the PostPruning algorithm.

4.2 Analysis of Spatial Database Optimization Results Based on Voronoi Diagrams and kNN

The study employed a grid Voronoi graph-based reverse nearest neighbor algorithm, NVD-RNN, in a road network environment. The CPU RT of the NVD-RNN algorithm, the Expectation Propagation (EP) algorithm and the Auto-Regression k Nearest Neighbors (ARkNN) algorithm were compared and analysed. The outcomes of the comparison of the three algorithms were shown in Figure 9.

In Figure 9, k was the number of target nodes to be obtained and D represented the number of generated points. From Figure 9(a), it indicated that the CPU RT of the three algorithms increased gradually under the growth of k value. The maximum CPU RT of the NVD-RNN algorithm was close to 900s and the minimum value was around 400s. The maximum CPU RT of the EP algorithm was around 1100s and the minimum value was close to 300s. The maximum CPU RT of the ARkNN algorithm was over 1000s, with a minimum value of around 200s. The NVD-RNN algorithm CPU runtime started to be smaller than the EP algorithm when the value of k was greater than 7s, and the NVD-RNN algorithm CPU runtime started to be smaller than the ARkNN algorithm when the value of k was greater than 17s. Figure 9(b) showed that the CPU runtime of the three algorithms increased gradually as the value of D increased. The maximum CPU runtime of the NVD-RNN algorithm was about 3s and the minimum value was about 1.3s. The maximum CPU runtime of the EP algorithm was about 5.3s and the minimum value was close to 3s, while the maximum CPU runtime of the ARkNN algorithm was close to 4s and the minimum value was about 2.1s. The minimum value was about 2.1s. In summary, the NVD-RNN algorithm was greatly superior to the rest algorithms. In order to further validate the performance of the NVD-RNN algorithm, the study selected Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) as comparative indicators. The experimental environment was also under the Windows 7 system, and the number of experiments was also 4. The comparison of mean square error and root mean square error of NVD-RNN algorithm, ARkNN algorithm, and EP algorithm was shown in Table 2.

From Table 2, it can be seen that the maximum MSE value of the NVD-RNN algorithm was 1.25 and the minimum value was 1.13. The maximum MSE value of the ARkNN algorithm was 2.01, and the minimum value was 1.87. The maximum MSE value of the EP algorithm was 2.51, and the minimum value was 2.37. The maximum RMSE value of the NVD-RNN algorithm was 0.971, and the minimum value was 0.863. The maximum RMSE value of the ARkNN algorithm was 1.373, and the minimum value was 1.329. The RMSE of the EP algorithm had a maximum value of 1.572 and a minimum value of 1.542. From this, it can be seen that the NVD-RNN algorithm had significantly lower values in MES and RMSE than the ARkNN algorithm and EP algorithm, which also indicated that the NVD-RNN algorithm had better performance. The k-nearest neighbor NVD-GkNN algorithm for the Voronoi group under the road network was the second algorithm. The following was a comparative analysis of the NVD-GkNN algorithm, the Incremental Euclidean Restriction (IER) algorithm and the Timing Algorithm (TA) algorithm, as shown in Figure 10.

Figure 10(a) illustrated that the CPU runtime of all three algorithms increased as the value of k increased. The maximum CPU runtime of the NVD-GkNN algorithm was about 2s and the minimum value was about 0.6s. The maximum CPU runtime of the IER algorithm was about 3.2s and the minimum value was about 0.8s. The maximum value of the TA algorithm CPU runtime was about 4.6s and the minimum value was about 2.1s. On the whole, the NVD-GkNN algorithm CPU runtime was always below the other two algorithms' value. Figure 10(b) shows that as the value of k gradually grew, the page accesses of the three algorithms also gradually increased. The maximum value of page accesses for the NVD-GkNN algorithm was around 280 and the minimum value was around 130. The maximum value of page accesses for the IER algorithm was around 340 and the minimum value was around 150. The maximum value of page accesses for the TA algorithm was around 450 and the minimum value was around 280. On the whole, the page views of the NVD-GkNN algorithm were consistently lower than those of the other two algorithms. In summary, it can be seen that the NVD-GkNN algorithm had a clear advantage. In addition to comparing the impact of different k values for the three algorithms on CPU runtime and page accesses, the study also compared the influence of the query point set Q on CPU runtime and page accesses for the three algorithms, as shown in Figure 11.

Figure 11(a) demonstrated that the CPU runtime of

Table 1: Comparison of accuracy and F1 values between the kNN Obs algorithm and the PostPruning algorithm

Algorithm	Number of experiments							
	1		2		3		4	
	Precision	F1	Precision	F1	Precision	F1	Precision	F1
PostPruning	93.7%	0.903	94.6%	0.892	92.9%	0.913	93.2%	0.921
kNN-Obs	95.8%	0.964	96.4%	0.976	95.7%	0.981	97.9%	0.975

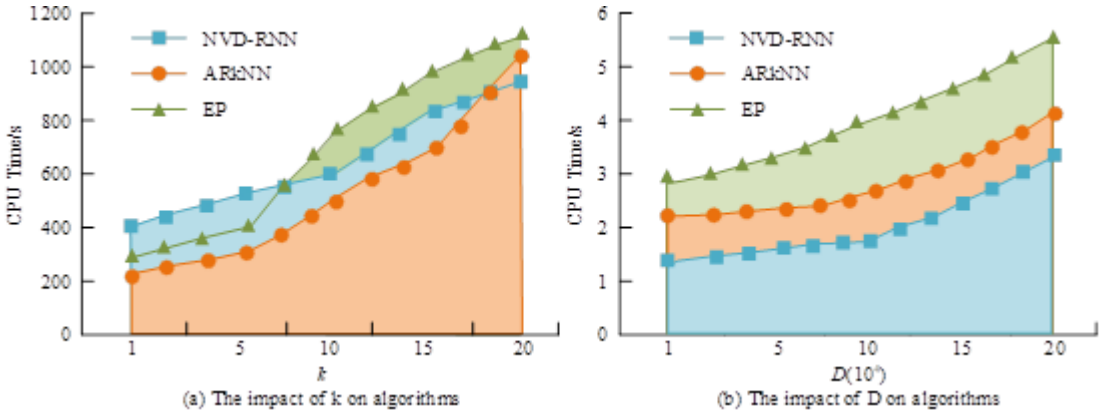


Figure 9: The impact of changes in k and D values on CPU running time

Table 2: Comparison of mean square error and root mean square error of NVD-RNN algorithm, ARkNN algorithm and EP algorithm

Algorithm	Number of experiments							
	1		2		3		4	
	MSE	RMSE	MSE	RMSE	MSE	RMSE	MSE	RMSE
NVD-RNN	1.21	0.971	1.17	0.863	1.25	0.954	1.13	0.874
ARkNN	1.87	1.373	1.97	1.329	2.01	1.356	1.89	1.337
EP	2.37	1.542	2.45	1.568	2.51	1.572	2.47	1.556

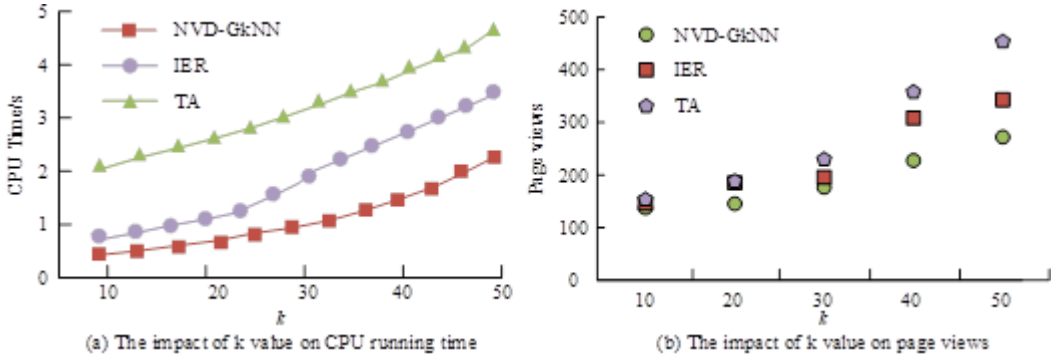


Figure 10: The impact of k value on CPU runtime and page views

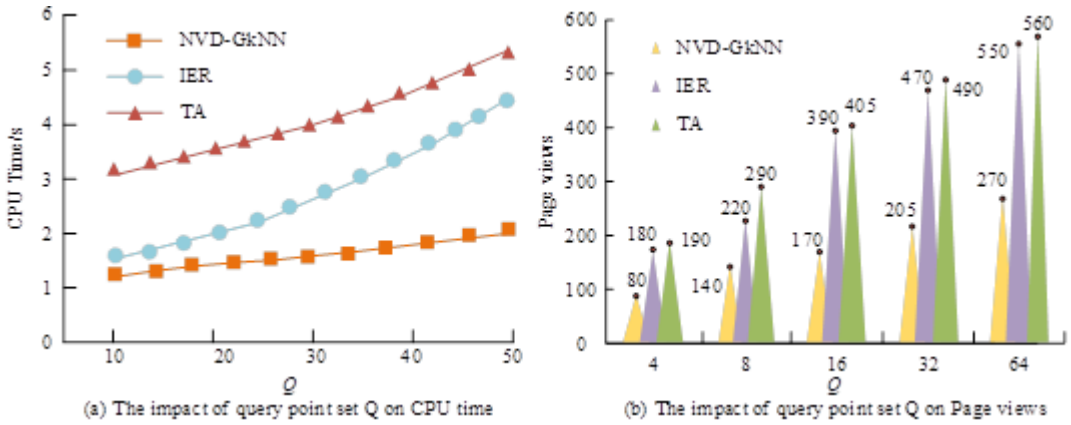


Figure 11: The impact of query point set Q on CPU running time and page views

all three algorithms grew as the value of Q increased. The maximum CPU runtime of the NVD-GkNN algorithm was about 2.5s and the minimum value was about 1.2s. The maximum CPU runtime of the IER algorithm was about 4.3s and the minimum value was about 1.7s. The maximum value of the TA algorithm CPU runtime was about 5.2s and the minimum value was about 3.1s. As can be seen from Figure 11(b), the page views of all three algorithms increased with the value of Q. The maximum value of page views for the NVD-GkNN algorithm was about 270 and the minimum value was about 80. The maximum value of page views for the IER algorithm was about 550 and the minimum value was about 180. The maximum value of page views for the TA algorithm was about 560 and the minimum value was about 190. In summary, it can be seen that the NVD-GkNN algorithm had more obvious advantages. In the following, the dynamic update algorithm, TA algorithm and IER algorithm were compared, as shown in Table 3 and Table 4.

Table 3: Performance comparison of dynamic update algorithm (ADDNVD-GkNN) with TA and IER algorithms

Algorithm	P	
	1000	1001
ADDNVD-GkNN	1.262	1.972
TA	3.102	6.165
IER	1.857	3.283

As can be seen from Table 3 and Table 4, the comparison in the ADDNVD-GkNN algorithm, the DENVD-GkNN algorithm and the other two algorithms was performed mainly in two dimensions: execution time and page accesses. When the number of page views was 1000, the execution time of the ADDNVD-GkNN algorithm was 1.84s and 0.595s faster than the TA algorithm and the IER algorithm's value, respectively. When the number of page views was 1001, the execution time of the ADDNVD-GkNN algorithm was 4.193s and 1.311s faster than the TA

Table 4: Performance comparison of dynamic update algorithm (DENVD-GkNN) with TA and IER algorithms

Algorithm	P	
	2000	1999
DENVD-GkNN	2.009	2.168
TA	4.326	8.186
IER	2.634	5.154

algorithm and the IER algorithm's value, respectively. In summary, it can be seen that ADDNVD-GkNN algorithm outperformed the TA algorithm and the IER algorithm. When the quantity of page views was 2000, the execution time of the DENVD-GkNN algorithm was 2.317s and 0.625s faster than the TA and IER algorithms' value, respectively. When the number of page views was 1999, the execution time of the DENVD-GkNN algorithm was 6.018s and 2.986s faster than the TA and IER algorithms' value, respectively. It can be seen that the DENVD-GkNN algorithm was faster than the TA and IER algorithms. GkNN algorithm had an advantage over the TA algorithm and the IER algorithm.

5 Conclusion

To address the shortcomings of the nearest neighbor query and its variants in the obstacle and road network environments, the study combines Voronoi diagrams and grid Voronoi diagrams with the nearest neighbor query and its variants. These methods are then used to query spatial data in the obstacle space and road network environments and compared with existing spatial database query methods. The results show that after a value of k greater than 4s, the kNN-Obs algorithm's minimum CPU runtime was 0.2s less than the PostPruning algorithm, and the maximum runtime was 1.2s less. The kNN-Obs has a minimum of 50 and a maximum of around 100 fewer page

views than the PostPruning algorithm. After a value of k greater than 7s, the CPU runtime of the NVD-RNN algorithm is at least about 100s less than that of the EP algorithm, and at most about 180s less. After a value of k greater than 17s, the CPU runtime of the NVD-RNN algorithm is at least about 50s less than that of the ARkNN algorithm. Under the influence of the k value, the CPU runtime of the NVD-GkNN algorithm is at least 0.1s and 1.2s less than that of the IER and TA algorithms, and at most 1.2s and 1.4s less. The page views under the NVD-GkNN algorithm are at least 10 and 20 less than that of the IER and TA algorithms, and at most 90 and 110 less. Under the influence of the Q set, the CPU RT of the NVD-GkNN algorithm is at least 0.3s and 1.9s less than that of the IER and TA algorithms, and at most 2.8s and 3.5s less than that of the IER and TA algorithms. The page views under the NVD-GkNN algorithm are at least 50 and 65 less than that of the IER and TA algorithms, and at most 290 and 300 less. The research results have improved the query efficiency of spatial databases, maintained data security of the database, and also contributed to network security. Future research can continue to explore the spatial data query method based on Voronoi diagram, improve the existing query method and enhance the query efficiency.

Acknowledgments

The research is supported by: the Henan Provincial Key R & D and Promotion Special (Science and Technology Tackling) Project, China. Project name: Research on the consistency checking algorithm of spatial direction relation in spatial database (No.232102210085).

References

[1] F. Baccelli, S. S. Kalamkar, "On point processes defined by angular conditions on Delaunay neighbors in the Poisson-Voronoi tessellation," *Journal of Applied Probability*, vol. 58, no. 4, pp. 952-965, 2021.

[2] D. Bonnet, S. Cabello, B. Mohar, H. Pérez-Rosés, "The inverse voronoi problem in graphs I: Hardness," *Algorithmica*, vol. 82, no. 10, pp. 3018-3040, 2020.

[3] H. Cai, Y. Zhu, J. Li, J. Yu, "A profit-maximizing mechanism for query-based data trading with personalized differential privacy," *The Computer Journal*, vol. 64, no. 2, pp. 264-280, 2020.

[4] H. Chen, J. V. D'Silva, H. Chen, B. Kemme, L. Hendren, "HorseIR: bringing array programming languages together with database query processing," *ACM SIGPLAN Notices*, vol. 53, no. 8, pp. 37-49, 2020.

[5] G. Ding, H. Sun, J. Li, C. Li, Y. Fei, "An efficient relational database keyword search scheme based on combined candidate network evaluation," *IEEE Access*, no. 99, pp. 30863-30872, 2020.

[6] T. Eiter, R. Ichise, J. X. Parreira, P. Schneider, L. Zhao, "Deploying spatial-stream query answering in C-ITS scenarios," *Semantic Web*, vol. 12, no. 1, pp. 41-77, 2021.

[7] F. Feng, S. Xiong, Z. Liu, Z. Xian, Y. Zhou, H. Kobayashi, "Cellular topology optimization on differentiable Voronoi diagrams," *International Journal for Numerical Methods in Engineering*, vol. 124, no. 1, pp. 282-304, 2023.

[8] Y. Gulzar, A. A. Alwan, S. Turaev, "Optimizing skyline query processing in incomplete data," *IEEE Access*, vol. 7, no. 1, pp. 178121-178138, 2019.

[9] J. Jakob, M. Guthe, "Optimizing LBVH offshore construction and Hierarchy placeholder raversal to accelerate kNN Queries on Point Clouds using the GPU," *Computer Graphics Forum*, vol. 40, no. 8, pp. 124-137, 2020.

[10] S. Jang, Y. E. Jang, Y. J. Kim, H. Yu, "Input initialization for inversion of neural networks using k-nearest neighbor approach," *Information Sciences*, vol. 519, no. 9, pp. 229-242, 2020.

[11] A. Jg, A. Wl, A. Zl, Z. B. Jian, S. B. Li, "A general fragments allocation method for join query in distributed database-ScienceDirect," *Information Sciences*, vol. 512, no. 13, pp. 1249-1263, 2020.

[12] K. Koufos, H. S. Dhillon, M. Dianati, C. P. Dettmann, "On the kk nearest-neighbor path distance from the typical intersection in the Manhattan Poisson line cox process," *IEEE transactions on mobile computing*, vol. 22, no. 3, pp. 1659-1671, 2023.

[13] T. Liu, S. Wang, Z. Lei, J. Zhang, X. Zhang, "Trajectory risk cognition of ship collision accident based on fusion of multi-model spatial data," *Journal of Navigation*, vol. 75, no. 2, pp. 299-318, 2022.

[14] J, Park, H, L. Dong, "Parallelly running k-nearest neighbor classification over semantically secure encrypted data in outsourced environments," *IEEE Access*, vol. 8, pp. 64617-64633, 2020.

[15] Y. Ruan, Y. Xiao, Z. Hao, B. Liu, "A nearest-neighbor search model for distance metric learning," *Information Sciences*, vol. 552, no. 3, pp. 261-277, 2020.

[16] P. Sarode, T. R. Reshmi, "Optimized query ordering data aggregation model using neural networks and group search optimization in wireless sensor network," *Ad Hoc & Sensor Wireless Networks*, vol. 46, no.3, pp. 189-214, 2020.

[17] F. Shi, H. Cao, X. Zhang, X. Chen, "A reinforced k-nearest neighbors method with application to chatter identification in high-speed milling," *IEEE Transactions on Industrial Electronics*, vol. 67, no. 12, pp. 10844-10855, 2020.

[18] J. Shi, L. Yang, "A climate classification of China through k-Nearest-neighbor and sparse subspace representation," *Journal of Climate*, vol. 33, no. 1, pp. 243-262, 2020.

[19] J. Wang, J. Shen, "Fast spectral analysis for approximate nearest neighbor search," *Machine Learning*, vol. 111, no. 6, pp. 2297-2322, 2022.

[20] A. Yl, C. Rxab, W. A. Xin, A. Hl, Z. Shi, D. Xin, "GLDH: Toward more efficient global low-density locality-sensitive hashing for high dimensions," *Information Sciences*, vol. 533, pp. 43-59, 2020.

Biography

Huili Xia, born in November 1980, female, from Wugang City, Henan Province, Han ethnicity. She obtained her Bachelor's degree in Computer Science and Technology from Zhengzhou University of Light Industry in 2005 and her Master's degree in Computer Application Technology from Hubei University of Technology in 2008. Since 2008, she has been a full-time teacher at Zhengzhou University of Economics and Business. She published fifteen academic papers, including seven papers in Chinese core journals and one paper indexed by EI. She edited two textbooks, hosted one Henan Provincial Key R & D and Promotion Special (Science and Technology Tackling) Project, and participated in four provincial-level scientific research projects, has been granted one utility model patent by the China National Intellectual Property Administration.

Design of Machine Learning Method for Network Security Situation Awareness

Wei Li, Xuefeng Jiang, Huan Le, Zhenmin Miao, and Hui Shao
(Corresponding author: Hui Shao)

Information Center, China Tobacco Zhejiang Industrial Co., Ltd
Hangzhou 310001, China
Email: shaohui8181@163.com

(Received July 12, 2023; Revised and Accepted Apr. 6, 2024; First Online Aug. 17, 2024)

Abstract

In recent years, benefiting from the mature application of communications, big data, cloud computing, and other technologies," the Internet " has been widely popularized in people's livelihoods, the economy, government affairs, and other aspects. However, the increasingly complex network environment and Internet data have brought huge hidden dangers to network security. In order to effectively respond to attacks on network systems, this study proposes a network security situational awareness system that combines the Bayesian algorithm and hidden semi-Markov model. In the process, Bayes-HsMM was first used to construct a network security situation model, HsMM was used to initialize the data, and the parameters were updated after obtaining an effective initial parameter set. Then rough set theory was used to extract features in the network connection and fuse the aggregation. Classes jointly build a network security situational awareness system-MixID. Comparing the research method with seven other methods, the results show that the average detection accuracy of MixID for network attacks is about 95.7%; the average detection accuracy of the other seven algorithms for network attacks is about 93.4%; the average detection accuracy of MixID for known network attacks is about 93.4%. The F1-measure of network attacks is about 97.3%; among the remaining seven algorithms, the highest F1-measure of known network attacks is about 92.5%; in addition, in dynamic networks, seven algorithms including CANN, FBSLAIDS, and HG-GA-SVM The updated network attack detection accuracy reaches a maximum of approximately 93.8%MixID's's detection accuracy is 94.2%. Comparing the above results, it can be seen that the detection accuracy of the research method is higher, the adaptability of the dynamic network is better, and it has better applicability and innovation.

Keywords: Bayesian Algorithm; Network Security; Network Security Situational Awareness; Semi Hidden Markov Model

1 Introduction

With the continuous maturity of network, the Internet had a great influence on human society. Especially after entering the era of "Internet plus", it can be said that people cannot live without the Internet. This has also led to a surge in network users and various online behaviors, but meanwhile, the network environment has become increasingly complex. Many criminals use the internet to directly or indirectly steal or destroy users' secrets, privacy, etc., bringing incalculable losses to users. Therefore, the network security arouses lots of attentions. Due to the rapid update and development of network attack methods, traditional network security protection technologies are becoming increasingly difficult to ensure network security. Therefore, network security situational awareness (NSSA) technology has emerged. The NSSA system is composed of multiple systems such as antivirus software, intrusion detection system, firewall, and security audit system, which can detect and defend against network attacks. It can also predict the network security situation [8, 9, 17]. Machine learning (ML), as the core of artificial intelligence, can acquire knowledge through self-learning and continuously improve its performance. In addition, ML can also obtain hidden and understandable knowledge from massive data, achieving effective utilization of information. It is widely used to solve complex problems in engineering and scientific fields. Typical ML algorithms include Bayes and Hidden Semi Markov Model (HMM). Bayes has the advantages of stable classification efficiency and fast processing speed for massive data; The classification accuracy of HsMM is high, and it can also be directly predicted. Based on the above characteristics, ML has certain advantages in NSSA. In the 1980s, ML began to be applied to network security, but due to technological limitations at that time, this method was not taken seriously. With the advancement of technology, applying ML to network security has once again become a new trend. To achieve accurate perception of network security situations, research attempts to integrate ML al-

gorithms into NSSA systems. Then, this study combines Bayes algorithm and hidden semi Markov model to establish a new hybrid situational awareness system, to compensate for the poor reliability of traditional situational awareness systems.

The innovation points of the research can be mainly divided into two aspects. First, since it is cumbersome to collect actual network security data, a unified security information collection model is proposed to achieve orderly storage of security events. Second, Bayesian algorithm and semi-hidden Markov are integrated to describe the change patterns within the network security protocol fields and the status relationships between fields, and ultimately achieve the correct perception of the network security situation. The research can be divided into four main parts. The first part is mainly a summary of the current status of Bayesian algorithm, semi-hidden Markov algorithm and network situation awareness technology at home and abroad; the second part is an introduction to Bayesian algorithm, and finally a machine learning algorithm based on Network situational awareness-MixID system. The third part is an analysis of the performance and practical application effects of the constructed system; the fourth part is a summary statement of the research content of the entire article.

2 Related Works

In the "Internet plus" era, Internet technology is utilized in many respects of production and life, so the importance of network security is becoming increasingly obvious, and NSSA technology is born from this. Yu proposed a feature extraction method based on ML to address the issue of delayed and accurate NSSA. This method extracts scene features through ML algorithms and analyzes the security of network information backend. The test indicates that it can reduce the noise and redundancy of network traffic data, and improve the accuracy of network security monitoring [20]. Al-Haija and Ishtaiwi proposed an intelligent classification model based on shallow neural networks (SNNs) for the classification of traffic packet data in firewalls. This model classifies the attributes of data packets through SNN and determines the next operation of the firewall.

According to the test, the classification precision is about 98.5%, the loss of cross entropy is only 0.022, and the false positive rate and false negative rate are low. The classification performance of this model is superior to other models [1]. Liu and his team presented an important node recognition algorithm based on K-shell decomposition algorithm and improved structural holes for the identification of important security nodes in complex networks. This algorithm analyzes the impact of nodes on network security based on the maximum connectivity coefficient, network efficiency, and Kendall coefficient of the network. Experiments have showcased that this algorithm improves the recognition accuracy of important nodes [12]. Chang presented a network security job service model based on rough set data analysis to address the issue of single star technology in pain producing network security job services. This model can simultaneously collect and process data, effectively respond to different intrusion methods, and has fast response speed and low network burden [4]. Yi et al. proposed a risk assessment model based on fuzzy theory, particle swarm optimization algorithm and RBF neural network for the problem of network security risk assessment. This model evaluates security risks by mining patterns in historical data and combining them with current network conditions. After testing, the accuracy of this model in security risk assessment is higher than that of traditional models [18].

With the continuous upgrading and widespread application of computer technology, machine learning algorithms, as a representative algorithm, have attracted the attention of many scholars due to their good self-learning ability and massive data processing capabilities. Lim Ong proposed an investment portfolio model based on unsupervised time series clustering to address the issue of achieving portfolio diversification. This model uses shape clustering method to cluster at different time lengths and perform many-to-one distance comparisons. This avoids the limitations of long feature vectors and scalability. The test indicates that the average annual return of the investment portfolio provided by the model has increased by 598 basis points, and the performance indicators have increased by 337% [11]. Hernández and his team proposed a bankruptcy prediction model based on rough sets and See5 to address the issue of insurance company bankruptcy prediction. The test results show that this model has higher prediction accuracy than traditional models, and it can also provide an easy understand decision model [6]. Sridevi and Arun proposed a medical model based on Gaussian Naive Bayes and Support Vector Machines to address the difficulty in developing treatment and care measures for Alzheimer's disease.

This model studies a large amount of data on Alzheimer's disease patients and develops comprehensive treatment and care measures. After testing, the treatment and nursing strategies developed by this model can effectively reduce the trauma faced by patients and healthcare workers [16]. Pal and Sharma P proposed a land modeling model based on ML to address the issue of land modeling. The test indicates that the land model established by this model has high resolution and low uncertainty [13]. Bheemalingaiah and his team proposed a prediction model based on Boolean ML algorithm to accurately predict the incidence rate of cardiovascular diseases. The test illustrates that the precision of this model in predicting cardiovascular diseases can reach 86% [2].

To sum up, there are currently many machine learning algorithms used in the field of network security and defense, and many scholars have analyzed the application of machine learning algorithms in the field of network security. However, there are currently few studies that combine Bayesian algorithms and semi-hidden Markov algo-

rithms to jointly deal with the field of network situational awareness and defend against attacks on network systems. In view of this, this study proposes a network security situational awareness system based on Baye-HsMM in order to achieve accurate detection of network attacks.

3 Research on Network Security Situation Awareness Model Based on Machine Learning Method

3.1 Network Security Situation Awareness Model Based on Bayes-HsMM

With the popularization of the Internet, the quantity of network users is also rapidly increasing, including some criminals. They use various attack methods to disrupt the network information security system, causing great harm to the interests of the country or individuals. Therefore, the protection of network security is urgent. ML is applied in NSSA due to its massive information processing ability, powerful learning ability, and classification recognition ability. Based on this new approach, a hybrid detection system MixID, a NSSA model integrating Bayes-HsMM, is proposed. The Bayesian algorithm can train and classify massive amounts of data very efficiently in network security situational awareness; at the same time, compared with other classifiers, it can usually give relatively better results when the data set is smaller fitting is not easy to occur. On the other hand, hidden Markov models are particularly suitable for processing time series data and provide researchers with more information about the data in certain applications, ultimately achieving good defense against network security attacks. The application of ML in network security defense is showcased in Figure 1.

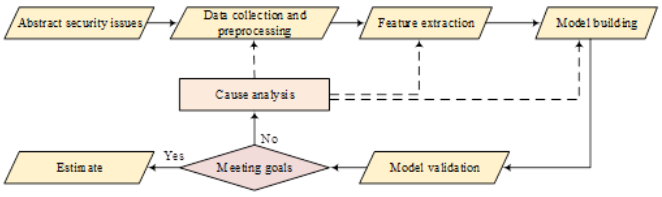


Figure 1: Application Process of ML In Network Security Defense

Figure 1 shows that the application of ML in network security defense can be separated into five steps, namely security problem abstraction, data collection and preprocessing, feature extraction, model construction, validation and evaluation. Security data collection mainly involves collecting and encapsulating various types of security data on heterogeneous security nodes, and then filtering, integrating, and calculating the correlation of the encapsulated data. The criteria for information filtering discrimination are indicated in Equation (1).

$$|a_p - H| < \delta \tag{1}$$

In Equation (1), a represents security information; a_p represents the safety information parameter; H represents the normal value; δ represents the size of the deviation. For two different events, if their basic eigenvalues are equal, the similarity calculation formula is indicated in Equation (2).

$$S(e,e') = \begin{cases} 1 & v_{ej} = v_{e'j}, \forall j \in [1,n] \\ 0 & v_{ej} \neq v_{e'j}, \forall j \in [1,n] \end{cases} \tag{2}$$

In Equation (2), v_{ej} and $v_{e'j}$ respectively represent the basic characteristics of two events; $S(e,e')$ represents event similarity; n represents the number of basic features. If $S(e,e') = 1$, it indicates that two events are similar and can be integrated; If $S(e,e') = 0$, then the two events are not similar and cannot be integrated. The formula for determining the repetition of similar events is illustrated in Equation (3).

$$\langle a, a, L, a \rangle^W \Longrightarrow [a]_C^W \tag{3}$$

In Equation (3), W represents the time interval window; C represents the minimum number of occurrences of similar events. The integration process of security information filtering is illustrated in Figure 2.

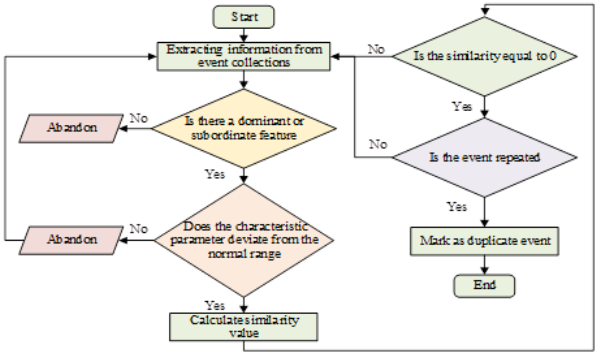


Figure 2: Security Information Filtering and Integration Process

After filtering and integrating security information, the next step is to perform event correlation analysis by calculating event dissimilarity. Due to the flexibility of the Bayes algorithm in data processing and its ability to classify different types of data information, this study used the Bayes algorithm to analyze event correlation [10,14,15]. The Bayes theorem formula is illustrated in Equation (4).

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)} \tag{4}$$

In Equation (4), $P(H|X)$ represents posterior probability; X represents the condition; H represents an event. Assuming the primary safety event set $E = \{e_1, e_2, L, e_n\}$,

the formula for calculating the character feature dissimilarity between any two events e_i and e_j in the set is demonstrated in Equation (5).

$$D^S(e_i, e_j) = \sum_{i=1}^{p}(\frac{1}{n_{f_{il}}} + \frac{1}{n_{f_{jl}}})\sigma(f_{il}, f_{jl}) \tag{5}$$

In Equation (5), $D^S(e_i, e_j)$ represents the degree of character feature dissimilarity; $n_{f_{il}}$ and $n_{f_{jl}}$ represent the number of common features l of events with values of f_{il} and f_{jl}, respectively; p represents the number of character type features; When $f_{il} = f_{jl}$, $\sigma(f_{il}, f_{jl}) = 0$, otherwise $\sigma(f_{il}, f_{jl}) = 1$. The equation for calculating the dissimilarity of digital features is demonstrated in Equation (6).

$$D^N(e_i, e_j) = \sum_{i=1}^{q}(f_{il} - f_{jl})^2 \tag{6}$$

In Equation (6), $D^N(e_i, e_j)$ represents the degree of dissimilarity of digital features; q represents the number of numerical type features. The formula for calculating the degree of dissimilarity of hybrid safety events is demonstrated in Equation (7).

$$D(e_i, e_j) = (\frac{D^N(e_i, e_j)^2 + D^S(e_i, e_j)^2}{D^N(e_i, e_j) + D^S(e_i, e_j)})^{\frac{1}{2}} \tag{7}$$

In Equation (7), $D(e_i, e_j)$ represents the degree of heterozygous heterogeneity. The higher the similarity between e_i and e_j, the closer $D(e_i, e_j)$ approaches zero. The security event correlation algorithm based on Bayes algorithm is demonstrated in Figure 3.

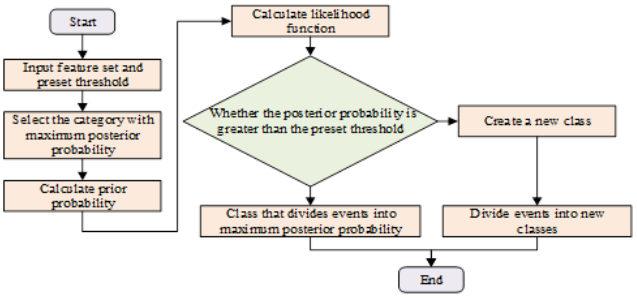


Figure 3: Security Event Correlation Algorithm Flow

After collecting and processing security data, the next step is feature extraction. However, the methods and technologies of network attacks have undergone rapid updates, which has made it difficult to extract features from network attack data. Therefore, the research proposes an unknown protocol parsing method called Rebuilder based on HsMM. The structure of HsMM is demonstrated in Figure 4.

Figure 4 shows that the duration period of state $s(i)$ is $d(i)$. When it transitions with the probability of a_{ij}, the duration period towards states $s(j)$ and $s(j)$ is $d(j)$. At this point, the hidden state sequence $S_{[t+1,t+d_j]}$ of $s(j)$ is

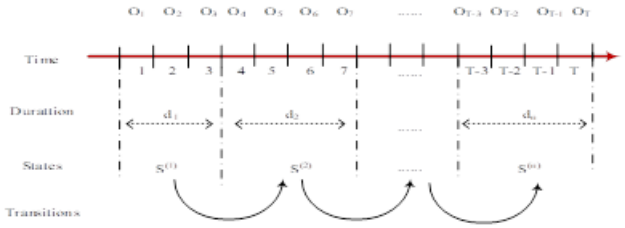


Figure 4: Schematic diagram of HsMM

emitted with a probability of b_j, $d_j(O_{[t+1,t+d_j]})$, generating the observation sequence $O_{[t+1,t+d_j]}$; Then it sequentially shifts backwards until the last observation value, and then estimates the parameter set using the observation sequence [3, 5, 7, 19]. The formula for calculating the probability of maximizing the scale observation sequence is demonstrated in Equation (8).

$$\hat{\lambda} = \arg_\lambda \max P(O|\lambda) \tag{8}$$

In Equation (8), λ represents the parameter set; $P(O|\lambda)$ represents the probability of the observation sequence. The calculation formulas for the forward, backward, and intermediate variables of HsMM are depicted in Equation (9).

$$\begin{cases} \alpha_t(i, d_i)Ⓢ P[O_{1:t}, s_{t-d+1:t}] = S^{(i)}|\lambda], s^{(i)} \in S, d_i \in D \\ \beta_t(j, d_j)Ⓢ P[O_{t+1:T}, s_{t-d+1:t}] = S^{(j)}|\lambda], s^{(i)} \in S, d_j \in D \\ \eta_t(i, d_i)Ⓢ P[O_{1:T}, s_{t-d+1:t}] = S^{(i)}|\lambda] = \alpha_t(i, d_i)\beta_t(i, d_i) \end{cases} \tag{9}$$

In Equation (9), $\alpha_t(i, d_i)$, $\beta_t(j, d_j)$, and $\eta_t(i, d_i)$ represent the forward, backward, and intermediate variables of the model, respectively; t represents time. The learning algorithm process of HsMM is depicted in Figure 5.

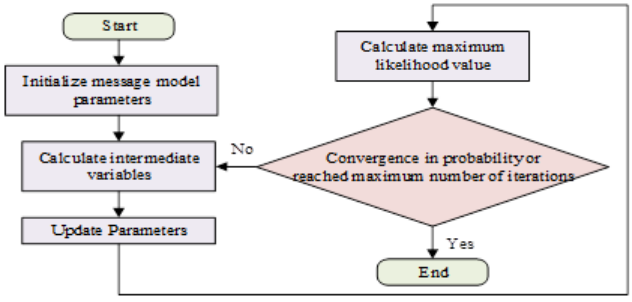


Figure 5: HsMM Algorithm Flow

Figure 5 shows that HsMM first initializes the data to obtain the initial parameter set; It then calculates intermediate variables and updates parameters; It repeats the above steps, iteratively updates the parameters, calculates the maximum likelihood probability, and records it. When the probability converges or reaches the maximum number of iterations, the algorithm terminates.

3.2 Network Security Situation Awareness System-MixID

To effectively respond to network attacks, NSSA technology is highly valued. There is a significant disparity between normal network connection instances and abnormal network connection instances, but there are similarities within them. Therefore, clustering algorithms can be used to classify network connections. Before classifying network connections, feature selection is the first step. The research uses the rough set theory to select the characteristics of network connection. Assuming that the fuzzy information system of the network connection instance is $IMS = (U, C \cup D, V, f)$, and there is a subset $B \subseteq C$, then the calculation formula of mutual information between B and D is depicted in Equation (10).

$$\begin{cases} I^{\%}(B,D) = H^{\%}(D) - H^{\%}(D|B) \\ H^{\%}(D) = -\frac{1}{n}\sum_{i=1}^{n} \log \frac{|[x_i]_D|}{n} \\ H^{\%}(D|B) = -\frac{1}{n}\sum_{i=1}^{n} \log \frac{|[x_i]_B \cap [x_i]_D|}{|[x_i]_B|} \end{cases} \tag{10}$$

In Equation (10), $I^{\%}(B,D)$ represents mutual information; $H^{\%}(D)$ represents the information amount of fuzzy equivalence relation; $[x_i]_D$ represents the fuzzy equivalent set containing divided by D on U; $|[x_i]_D| = \sum_{j=1}^{n} r_{ij}$ represents the cardinality of the set; $H^{\%}(D|B)$ represents the conditional entropy of D under condition B. After feature extraction is completed, a Gaussian mixture model is used for clustering. Assuming that the mixed dataset $X = \{X_1, X_2, L, X_n\}$ contains n samples and is generated by a model containing K Gaussian distributions, the distribution function of the mixed data samples is depicted in Equation (11).

$$p(x) = \sum_{i=1}^{K} \pi_i N_i(x, \mu_i, \sum_i) \tag{11}$$

In Equation (11), μ_i represents the mean of the -th Gaussian distribution $N_i x, \mu_i, \sum_i)$; $\sum_i$ serves as the variance of the i-th Gaussian distribution $N_i(x, \mu_i, \sum_i)$; π_i represents the weight of the i-th Gaussian distribution, and $\sum_{i=1}^{K} \pi_i = 1$. The calculation formula of probability density function of Gaussian function is depicted in Equation (12).

$$\begin{cases} N_i(x, \mu_i, \sum_i) = (2\pi)^{-\frac{d}{2}} |\sum_i|^{\frac{1}{2}} \exp\{-\frac{1}{2}(x-\mu_i)^T \sum_i^{-1}(x-\mu_i)\} \\ p(x|\Theta) = \sum_{i=1}^{K} \pi_i p(x|\Theta_i) \\ = \sum_{i=1}^{K} \pi_i (2\pi)^{-\frac{d}{2}} |\sum_i|^{\frac{1}{2}} \exp\{-\frac{1}{2}(x-\mu_i)^T \sum_i^{-1}(x-\mu_i)\} \end{cases}$$

In Equation (12), d represents the data dimension; Θ represents the set of parameters to be estimated, $\Theta Ⓢ \langle K, \pi, \mu, \sum \rangle$. After clustering, the variance matrix is obtained, and the optimization objective calculation formula is depicted in Equation (13).

$$J(\Theta) = \frac{\sum_{i=1}^{K} \sum_{p \in C_i} (p-\mu_i)(p-\mu_i)^T}{\sum_{i=1}^{K} (\mu_i - \mu)(\mu_i - \mu)^T} \tag{13}$$

In Equation (13), C_i represents the i-th cluster; p represents the sample in C_i; μ_i represents the mean of C_i. The closer the distance between samples within the visible cluster, the farther the distance between samples between clusters, the smaller $J(\Theta)$, and the better the clustering effect. To avoid duplicate calculations on the data, incremental updates are required, and the calculation formula is depicted in Equation (14).

$$\begin{cases} \pi_l^{new} = \frac{1}{|X_{new}|}[|X_{old}| g\pi_l^1 + \sum_{i=1}^{|X'|} p(l|x_i, \Theta^{old})] \\ \mu_l^{new} = \frac{|X^{old}| g\pi_l^1 g\sigma_l^1 + \sum_{i=1}^{|X'|} p(l|x_i, \Theta^{old})}{|X^{old}| g\pi_l^1 + \sum_{i=1}^{|X'|} p(l|x_i, \Theta^{old})} \\ \sum_l^{new} = \frac{|X^{old}| g\pi_l^1 g(\sigma_l^1 - \mu_l^{new})(\sigma_l^1 - \mu_l^{new})^T}{|X^{old}| g\pi_l^1 + \sum_{i=1}^{|X'|} p(l, x_i, \Theta^{old})} \\ \qquad + \frac{\sum_{i=1}^{|X'|} p(l|x_i, \Theta^{old})(x_i - \mu_l^{new})^T}{|X^{old}| g\pi_l^1 + \sum_{i=1}^{|X'|} p(l, x_i, \Theta^{old})} \end{cases} \tag{14}$$

After clustering is completed, the matching degree between the detected mode and the normal mode or attack mode is determined by calculating cosine similarity. The formula for calculating cosine similarity is showcased in Equation (15).

$$d(x, y) = \frac{x^T y}{||x|| ||y||} \tag{15}$$

In Equation (15), x and y represent different samples. The greater the cosine similarity, the higher the matching degree. When the detection mode matches normal or attack mode, the health value of the detection mode increases and other modes in the library decrease; When the life value of any mode is less than the preset threshold λ, the record is deleted. The above algorithms are combined to form an adaptive network intrusion detection system, which is combined with Bayes algorithm and HsMM model to form a NSSA system. The NSSA system consists of four parts: data collection, feature processing, detection response, and database; The data collection is based on Bayes algorithm, feature processing is based on HsMM model, and detection response is based on adaptive network intrusion detection system. The MixID structure is showcased in Figure 6.

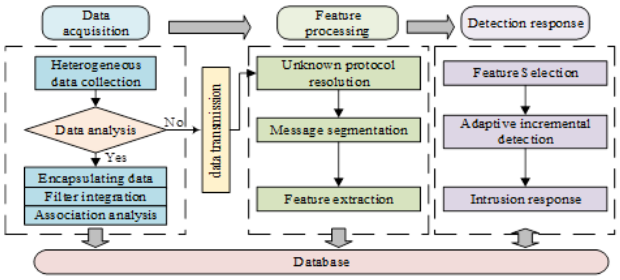


Figure 6: MixID Structure Diagram

Figure 6 shows that MixID first collects and analyzes data based on preset rules. If it can be parsed, data encapsulation, filtering integration, and correlation analysis should be carried out, and placed in the database; If it cannot be parsed, data transmission will be carried out, marked as unknown network behavior, and message segmentation and feature extraction will be performed before

Table 1: Device Configuration Information

Name	Configuration	Role
Terminal 0	4-core processor, 4G memory, Win7 system, network testing software	Network testing equipment
Terminal 1/2	4-core processor, 4G memory, Win7 system, Simulated attack software	Attack equipment
Terminal 3/4	4-core processor, 4G memory, Win7 system, Collection Agent	Simulating host operation, target machine
The server	4-core processor, 4G memory, Win Serverr16 system	Provide network access services, target machine
Terminal 5/6	8-core processor, 16G memory, Win7 system	System operation, managing hosts
Firewall	Set according to actual situation	Provide security information, monitor the network

being placed in the database. Finally, the extracted features mentioned above are used as inputs to detect and respond to network attacks through feature selection, clustering, and incremental updates. MixID mainly analyzes events based on experience to obtain the characteristics of known networks. This feature ensures the process accuracy of MixID for known network behavior. Feature processing ensures the detection performance of MixID for unknown network behavior by parsing it. The detection response process ensures the adaptability and generalization ability of MixID to dynamic networks through adaptive and incremental updates. The mutual compensation between data collection, feature extraction, and detection response ensures the ability of MIxID to resist unknown network risks.

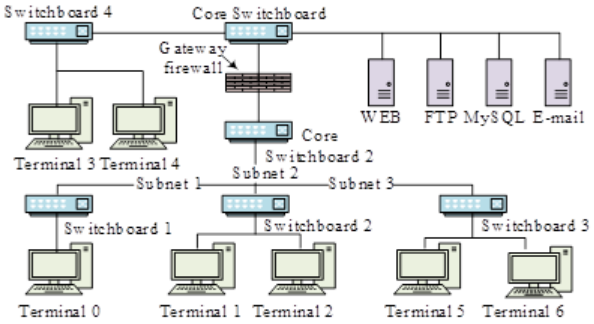


Figure 7: Simulation Network Topology Structure of MixID System

4 Analysis and Application of 3 MixID System Test Results

To test the network detection performance of the MixID system, the study will compare MixID, CANN, FBSLAIDS, HG-GA-SVM, SVM-SA-DT, Firefly C4.5, Firefly Bn, and CAI. Then, this study measures the performance of the above system using indicators such as accuracy, false positive rate, false positive rate, and F1 measure. The simulation network topology of the MixID system is showcased in Figure 7.

For ensuring the accuracy of the experiment, the performance tests of the above systems were conducted under the same conditions. The training set contains a total of 50000 network instances, of which 10000 are attack instances; The test set contains a total of 10000 network instances, of which 3500 are attack instances and 500 are unknown attacks. The configuration of the simulation network is showcased in Table 1.

Table 1 indicates that device terminal 0 is a testing device; Terminals 1 and 2 are the attack hosts; Terminals 3, 4, and servers are target machines, and both the target machine and firewall are equipped with collection agents to send network attack information to the management host for security analysis. Terminal 5 is the security information receiving end and feature extraction end; Terminal 6 is responsible for running intrusion algorithms and monitoring network status. The convergence of various algorithms is showcased in Figure 8.

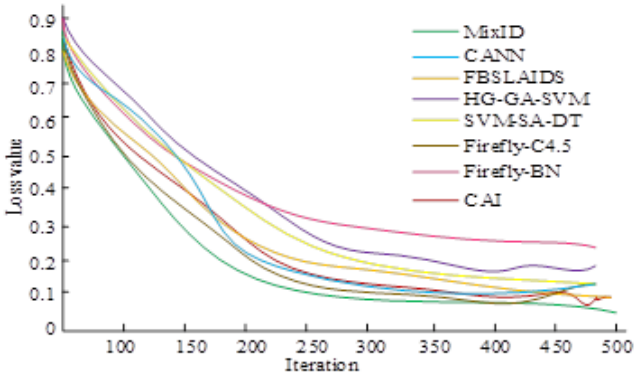


Figure 8: Loss Curves of Various Algorithms

Figure 8 shows that MixID converges after approxi-

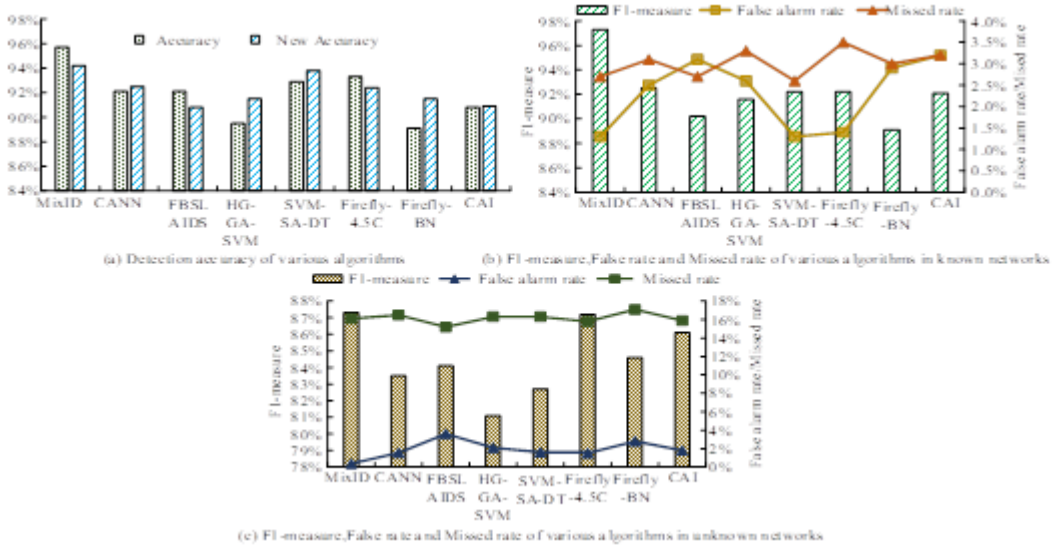


Figure 9: The Detection Performance of Various Algorithms for Known and Unknown Network Attacks

mately 200 iterations; CANN, FBSLAIDS, HG-GA-SVM, SVM-SA-DT, Firefly C4.5, Firefly Bn, and CAI began to converge after approximately 250, 280, 330, 300, 250, 300, and 250 iterations, respectively. It can be seen that MixID has better convergence than the other seven algorithms. The comparison of detection performance between known and unknown network attacks by MixID, CANN, FBSLAIDS, HG-GA-SVM, SVN-SA-DT, Firefly C4.5, Firefly Bn, and CAI is showcased in Figure 9.

Figure 9(a) shows that the detection accuracy of MixID for network attacks is about 95.7%, and the updated detection accuracy in dynamic networks is about 94.2%; The detection accuracy of CANN, FBSLAIDS, HG-GA-SVM, SVM-SA-DT, Firefly C4.5, Firefly Bn, and CAI for network attacks is about 92.1%, 92.1%, 89.5%, 92.9%, 93.4%, 89.1%, and 90.8%, respectively; The updated detection accuracy in dynamic networks is approximately 92.5%, 90.8%, 91.5%, 93.8%, 92.4%, 91.5%, and 90.9%, respectively. This indicates that MixID has the highest accuracy in detecting network attacks. Figure 9(b) shows that the false positive rate, false negative rate, and F1 measure of MixID for known network attacks are approximately 1.3%, 2.5%, and 97.3%, respectively. The false positive rates of known network attacks for CANN, FBSLAIDS, HG-GA-SVM, SVM-SA-DT, Firefly C4.5, Firefly Bn, and CAI are around 2.5%, 3.1%, 2.6%, 1.3%, 1.4%, 2.9%, and 3.2%, respectively; The underreporting rates for known network attacks are approximately 3.1%, 2.7%, 3.3%, 2.6%, 3.5%, 3%, and 4.1%, respectively; The F1 measures against known network attacks are approximately 92.5%, 90.2%, 91.6%, 92.2%, 92.2%, 89%, and 92.1%, respectively. It can be seen that MixID has the highest F1 measure against known network attacks, with lower false positive and false negative rates. Figure 9(c) shows that the false positive rate, false negative rate, and F1 measure of MixID for unknown network attacks are around 0.4%, 16%, and 87.3%, respectively. The false positive rates of CANN, FBSLAIDS, HG-GA-SVM, SVM-SA-DT, Firefly C4.5, Firefly Bn, and CAI for unknown network attacks are approximately 1.5%, 3.6%, 2.1%, 1.6%, 1.5%, 2.8%, and 1.8%, respectively; The underreporting rates for unknown network attacks are around 16.5%, 15.2%, 16.3%, 16.3%, 15.8%, 17.1%, and 15.9%, respectively; The F1 means for unknown networks are approximately 83.5%, 84.1%, 81.1%, 82.7%, 87.2%, 84.6%, and 86.1%, respectively. From this, it can be seen that MixID has the lowest F1 mean and false alarm rate for unknown network attacks, with a lower false alarm rate and better detection performance for unknown network attacks. The recall rates and P-R curves of MixID, CANN, FBSLAIDS, HG-GA-SVM, SVM-SA-DT, Firefly C4.5, Firefly Bn, and CAI are showcased in Figure 10.

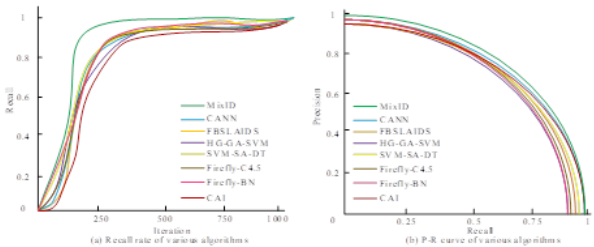


Figure 10: Recall rates and P-R curves of various algorithms

Figure 10(a) shows that the average recall rate of MixID is approximately 97.5%; The average recall rates of CANN, FBSLAIDS, HG-GA-SVM, SVM-SA-DT, Firefly C4.5, Firefly Bn, and CAI are approximately 96.8%, 97.2%, 96.6%, 97.3%, 96.4%, 96.9%, and 95.8%, with

MixID having the highest recall rate. Figure 10(b) indicates that the P-R curve of MixID almost completely envelops the P-R curves of the other seven algorithms. This indicates that MixID has the best classification performance for network attacks. In practical applications, due to the randomness of network attack behavior, the network security environment is unpredictable. But analyzing a certain period of network security events can achieve the mining of network security situation patterns. Therefore, a NSSA system can be used to predict network security trends within a certain period. The situation prediction results of various network situational awareness models are showcased in Figure 11.

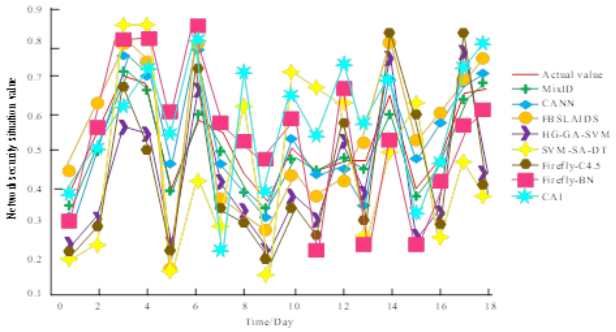


Figure 11: Network Security Situation Prediction Results of Different Models

Figure 12 shows that all eight models can handle historical data well, but the prediction performance of models other than MixID is not satisfactory. Especially for SVM-SA-DT and Firefly BN, the predicted results differ significantly from the true values, up to 0.35; The difference between the predicted results of CANN and the actual value has slightly decreased, with a maximum of about 0.3, but it is still hard to satisfy the needs; The predicted curve of MixID has a high degree of agreement with the actual situation value curve, with a maximum difference of about 0.05 and a minimum difference approaching 0. This indicates that MixID has good NSSA performance in practical applications, which can achieve accurate detection of network attacks and accurate prediction of security trends. The confidence levels and MAPE of eight algorithms such as MixID and CANN are showcased in Figure 12.

Figure 12(a) shows that the average confidence level of MixID is about 0.93, the maximum confidence level is about 0.97, and the minimum confidence level is about 0.9; Among the other algorithms, the average confidence value of CANN is the highest, about 0.86, the maximum confidence value of CANN is about 0.91, and the minimum confidence value is about 0.8; This is much lower than the confidence value of MixID, indicating that the confidence level of MixID is significantly higher than other algorithms. Figure 12(b) shows that among the other algorithms, the average MAPE value of SVM-SA-DT is the smallest, around 0.15, with a maximum MAPE value of

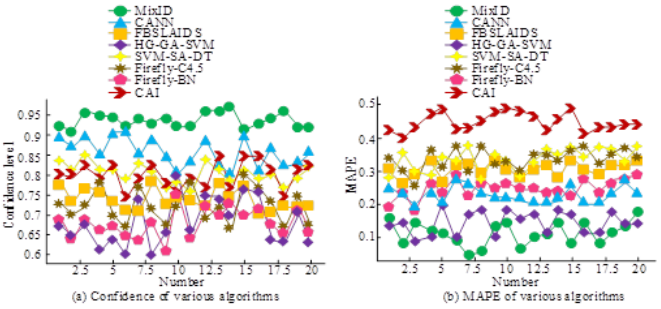


Figure 12: Confidence and MAPE of Various Algorithms

around 0.2 and a minimum MAPE value of about 0.1; The average MAPE value of MixID is about 0.1, the maximum MAPE value is about 0.15, and the minimum MAPE value is about 0.02. It can be seen that in most cases, the prediction error of MixID is smaller than that of other models. The results indicate that MixID performs relatively smoothly and has a high accuracy in predicting network security situations.

5 Conclusion

The NSSA system, as a means of network security defense, can detect network attack behavior. To ensure the accuracy of NSSA system in detecting network attacks, a research proposal is proposed to establish a NSSA system - MixID by combining Bayes and HsMM. The study conducted performance testing and comparison on eight systems including MixID, CANN, FBSLAIDS, HG-GA-SVM, SVM-SA-DT, Firefly C4.5, Firefly Bn, and CAI. The test indicates that the average recall rate of MixID is around 97.5%; Among the other seven systems, SVM-SA-DT has the highest average recall rate, approximately 97.3%, indicating that MixID has a higher recall rate than other models. The detection accuracy of the seven models other than MixID for network attacks is around 92.1%, 92.1%, 89.5%, 92.9%, 93.4%, 89.1%, and 90.8%, respectively; The detection accuracy of MixID is about 95.7%, which is higher than other models. The false alarm rate of MixID for known network attacks is about 1.3%, while the false alarm rates of SVM-SA-DT and Firefly C4.5 in the other seven models are about 1.3% and 1.4%, respectively, which are comparable to MixID and lower than the other models. The false positive rate of MixID for known network attacks is approximately 2.5%; Among the other seven models, FBSLAIDS and SVM-SA-DT are the lowest, at around 2.7% and 2.6%, respectively, but still slightly higher than MixID. Models such as CANN, FBSLAIDS, and HG-GA-SVM require a minimum of about 250 iterations to begin convergence, while MixID only needs about 200 iterations to begin convergence. The false alarm rate of MixID for unknown network construction is about 16%, while the false alarm rates of FBSLAIDS, Firefly C4.5, and CAI are about 15.2%, 15.8%, and 15.9%,

respectively, which are lower than MixID. In terms of network security situation prediction, the predicted results of MixID are highly consistent with the true values, with an average MAPE value of about 0.1 and an average confidence value of about 0.93; Among the other seven algorithms, CANN has the highest average confidence value, around 0.86; The MAPE value of SVM-SA-DT is the smallest, approximately 0.15; It can be seen that the confidence level of MixID is higher than other algorithms, and the error is smaller than other algorithms. The test showcases that MixID has higher network attack detection accuracy, lower false alarm rate, better classification performance, and convergence compared to other NSSA systems. It can achieve accurate perception and prediction of network security situations. However, MixID's performance in underreporting unknown network attacks is not ideal, and there is still room for improvement.

References

[1] Q. A. Al-Haija, A. Ishtaiwi, "Multi-class classification of firewall log files using shallow neural network for network security applications," *Advances in Intelligent Systems and Computing*, vol. 1370, no. 1, pp. 27-41, 2021.

[2] M. Bheemalingaiah, G. R. Swamy, P. Vishvapathi, P. V. Babu, P. N. Rao, "Detection of heart disease by using reliable boolean machine learning algorithm," *Journal of Theoretical and Applied Information Technology*, vol. 99, no. 15, pp. 3856-3880, 2021.

[3] B. Cai, L. Zhang, Y. Shi, "Control synthesis of hidden semi-Markov uncertain fuzzy systems via observations of hidden modes," *IEEE Transactions on Cybernetics*, vol. 50, no. 8, pp. 3709-3718, 2020.

[4] M. Chang, "Construction of network security job service model based on rough set data analysis algorithm," *Journal of Intelligent and Fuzzy Systems*, vol. 39, no. 2, pp. 4981-4987, 2020.

[5] Y. Du, C. Duan, T. Wu, "Lubricating oil deterioration modeling and remaining useful life prediction based on hidden semi-Markov modeling: Proceedings of the institution of mechanical engineers," *Journal of Engineering Tribology*, vol. 236, no. 5, pp. 916-923, 2022.

[6] J. F. Hernández, Z. Díaz, M. J. Segovia, E. M. D. Pozo, "Machine learning and statistical techniques: An application to the prediction of insolvency in Spanish non-life insurance companies," *The International Journal of Digital Accounting Research*, vol. 5, no. 9, pp. 1-45, 2020.

[7] A. Islam, F. Othman, "Prevention of shoulder-surfing attack using shifting condition with the digraph substitution rules," *Artificial Intelligence and Applications*, vol. 1, no. 1, pp. 58-68, 2023.

[8] A. K. Jain, S. R. Sahoo, J. Kaubiyal, "Online social networks security and privacy: comprehensive review and analysis," *Complex & Intelligent Systems*, vol. 7, no. 5, pp. 2157-2177, 2021.

[9] P. Li, A. A. Laghari, M. Rashid, J. Gao, T. R. Gadekallu, A. R. Javed, S. Yin, "A deep multimodal adversarial cycle-consistent network for smart enterprise system," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 1, pp. 693-702, 2022.

[10] Y. Li, Y. Zuo, H. Song, Z. Lv, "Deep learning in security of internet of things," *IEEE Internet of Things Journal*, vol. 9, no. 22, pp. 22133-22146, 2021.

[11] T. Lim, C. S. Ong, "Portfolio management: A financial application of unsupervised shape-based clustering-driven machine learning method," *International Journal of Computing and Digital Systems*, vol. 10, no. 1, pp. 235-243, 2021.

[12] Y. Liu, J. Wang, H. He, G. Huang, W. Shi, "Identifying important nodes affecting network security in complex networks," *International Journal of Distributed Sensor Networks*, vol. 17, no. 2, pp. 1560-1571, 2021.

[13] S. Pal, P. Sharma, "A review of machine learning applications in land surface modeling," *Earth*, vol. 2, no. 1, pp. 174-190, 2021.

[14] P. Praveen, "Journal of cardiovascular disease research a novel approach to predict cardio disease using naive bayes algorithm," *Journal of Cardiovascular Disease Research*, vol. 12, no. 13, pp. 2016-2021, 2021.

[15] F. Smarandache, "Plithogeny, plithogenic set, logic, probability and statistics: a short review," *Journal of Computational and Cognitive Engineering*, vol. 1, no. 2, pp. 47-50, 2022.

[16] M. Sridevi, K. Arun, "A framework for performance evaluation of machine learning techniques to predict the decision to choose palliative care in advanced stages of Alzheimer's disease," *Indian Journal of Computer Science and Engineering*, vol. 12, no. 1, pp. 35-46, 2021.

[17] X. Wang, S. Yin, M. Shafiq, A. A. Laghari, S. Karim, O. Cheikhrouhou, 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, pp. 1-14, 2022.

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

[19] S. Yin, H. Li, A. A. Laghari, S. Karim, A. K. Jumani, "A bagging strategy-based kernel extreme learning machine for complex network intrusion detection," *EAI Endorsed Transactions on Scalable Information Systems*, vol. 8, no. 33, pp. e8, 2021.

[20] G. Yu, "Research on computer network information security based on improved machine learning," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 3, pp. 6889-6900, 2020.

Biography

Wei Li, obtained his ME in Computer Application from Xi'an Jiaotong University in 2009. Presently, he is working as a deputy senior engineer in China Tobacco Zhejiang Industrial Co., Ltd. His areas of interest are computer communication technology, Intelligent network and network security.

Xuefeng Jiang, obtained his Bachelor of Science from Peking University in 1988, and his Master of Science from Hangzhou University in 1991. Presently, he is working as security minister in China Tobacco Zhejiang Industrial Co., Ltd. His areas of interest are digital security, big data security and security compliance.

Huan Le, graduated from Electrical Engineering and Automation from Shandong University in 2015. Presently, he is working as a technician engineer in China Tobacco Zhejiang Industrial Co., Ltd. His areas of interest are network security operation and maintenance, industrial network security and industrial control system security.

Zhenmin Miao, obtained his BE in Economic Information Management from Zhejiang Wanli University in 2007. Presently, he is working as a network engineer in China Tobacco Zhejiang Industrial Co., Ltd. His areas of interest are computer communication technology, intelligent network and network security.

Hui Shao, obtained her BE in Computer Science and Technology from the University of Electronic Science and Technology of China in 2019, and followed up with her ME in Computer Science and Technology from the same university in 2022. Presently, she is working in China Tobacco Zhejiang Industrial Co., Ltd. Her areas of interest are computer network security, intelligent networks, and visualization technology.

Construction of Early Warning and Defense Model for Distributed Network Viruses

Lin Wang and Chuang Wang
(Corresponding author: Lin Wang)

Xuchang Vocational Technical College, Xuchang 461000, China
Email: wlin16@126.com

(Received Aug. 8, 2023; Revised and Accepted June 2, 2024; First Online Aug. 17, 2024)

Abstract

With the development of Internet technology, information security has become a prospective cohort study in academia, which is of great significance for computer technology in China. First, a security warning system based on distributed network viruses is designed to improve network security further and provide early warning for network viruses. Based on this system, an optimized Apriori algorithm is introduced to mine association rules between data. A distributed network virus warning and defense system based on the Apriori algorithm is constructed. The application effectiveness of the proposed model is validated using different network virus attack scenarios as examples. Taking the results of the Dos dataset as an example, this model's DR, FDR, and MDR are 82.03%, 11.12%, and 18.75%, respectively. The amount of alarm data after defense significantly decreases, indicating that the system has a high blocking rate against network viruses. This indicates that the method can meet the requirements of virus attack scenarios. It has good applicability and application effects. This study can provide a certain guidance value for distributed network virus warning and defense.

Keywords: Apriori Algorithm; Distributed Network; Early Warning; Virus

1 Introduction

In the new era, countries have elevated their emphasis on politics, economy, culture, and military construction to an unprecedented level. The emergence of information technology has greatly influenced and changed people's lives. Especially e-commerce, e-banking and other information technologies that use the internet as a bridge have brought great convenience to people's lives [4, 6, 17]. The internet has become an indispensable and important component of economic and social development. However, with the rapid development of network technology, there are also some network security issues such as hacker attacks and customer privacy breaches. People are eager to improve network security through certain means. Therefore, various network security technologies have emerged, promoting the development and improvement of the network security system framework [5]. These network security technologies play a huge role in ensuring information security, which can to some extent resist virus intrusion [10]. Currently, to effectively defend against network viruses, intrusion prevention technology has received high attention. The distributed network early warning system achieves the data protection by timely tracking data, verifying data security and legality. Intrusion prevention technology is an important tool for maintaining information security and ensuring network order. Therefore, based on the improved Apriori algorithm, a network virus warning and defense model is constructed, hoping to achieve better network security maintenance through this method. The Apriori algorithm is an algorithm for mining frequent itemsets of association rules. The core is a recursive algorithm based on the two-stage frequency set idea. It has unique advantages in data feature mining. Therefore, to better ensure network security, the Apriori algorithm is introduced to mine association rules between intrusion data. In response to the shortcomings of the Apriori algorithm in data mining, the Apriori algorithm is improved. Based on the improved Apriori algorithm, a warning and defense model for network viruses is constructed, hoping to achieve better network security maintenance through this method.

The innovative points of the research are as follows. Firstly, research innovatively utilizes the Apriori algorithm to mine association rules between data, to better analyze potential relationships between data. Secondly, the research optimizes both the Apriori algorithm and the early warning system. The hash technology is applied to select and partition samples. In the early warning system, all alarm information is uniformly labeled to achieve unique identification of the information.

The research mainly includes four parts. The first part summarizes the current research status of Apriori algorithm and network information security. The second part designs a distributed network warning model based on the

improved Apriori algorithm. The third part verifies the performance of the proposed method in the study. The fourth part summarizes the research content and proposes future research directions.

2 Related Works

The Apriori algorithm is a fundamental and classic algorithm in association rule mining, which is widely used in the data mining. Zhou designed an information recommendation book management system based on an improved Apriori data mining algorithm to open book information to library staff and borrowers. The improved Apriori data mining algorithm is used to mine strongly correlated material rules with support and generate an association rule database. The experimental results show that the system can effectively recommend book related information. When 50 clients are running simultaneously, the CPU usage is only 6.47%, indicating good performance [22]. Protecting the diversity of visual landscapes has become an important component of planning decisions. Aur *et al.* analyzed the landscape features of a certain area using the data mining tool Apriori algorithm. Nine different types of landscape features are selected to examine people's consensus on the landscape and the relationship between perception parameters, such as mystery, typicality, vitality, safety, impression, silence, perspective, degradation, and protection worth. The high visual quality of the landscape indicates higher consistency among observers [3]. To address the pressure on SMS gateways caused by the increasing number of alarm messages, Zhu tested the evaluation module device based on an improved Apriori algorithm and confidence formula. The web log mining and mining system design are implemented in three steps. The results demonstrate that the method can effectively reduce the pressure on SMS gateways [23]. To improve the judgment and decision-making ability of sports training effectiveness, Wang *et al.* proposed a design method for a sports training decision support system based on the improved association rule Apriori algorithm. A spatial model based on decision support data association rule distribution is constructed. This method achieves adaptive scheduling and information fusion of motion training decision support data. It improves the judgment ability of sports training decision support, which has a good effect on sports training decision-making [19]. Natalia and Salvatore analyzed the causes of runway deviation accidents at airports using the Apriori algorithm. Different variables are related to different runway offsets. For all types of runway offsets, the most important variable is the level of the aircraft. Major and hazardous severity events are related to small aircraft, while catastrophic severity events are related to medium to large aircraft [12].

Network information security is a key concern in the development of Internet. In recent years, to better solve network security problems, many scholars have conducted in-depth research on corresponding network security detection methods. There are problems in the power system network security system, such as excessive reliance on manual labor, low detection accuracy, and poor real-time performance. Therefore, Liao *et al.* proposed a detection method to calculate the information entropy of flow detection and random forest classification. The results indicate that this method can quickly locate and detect abnormal locations of traffic. It reduces the workload of the power system security monitoring application platform. It effectively improves the reliability and early warning ability of power system network security [9]. The existing network information security automatic monitoring system has incomplete monitoring, slow response speed, and low accuracy. Therefore, Niu designed an automatic monitoring system for network information security in a cloud computing environment. On the basis of the overall system architecture, information collection, information transmission, and information security warning modules are designed. The acquisition, transmission, and integration of network information changes, as well as the risk warning of network anomalies, are achieved. The experimental results indicate that the system has high comprehensive monitoring capability and a response speed of less than 0.5s [14]. Li developed a GANs encryption algorithm based on neural networks. This encryption algorithm can improve the objective function and learning model, thereby achieving better algorithm security performance. The proposed method can achieve facial generation without artificial knowledge, while ensuring the information security [8]. Wang and Long designed a campus network security protection system using intrusion detection algorithms. By setting up firewalls, encrypting cloud data, using intrusion detection technology, and restoring data, existing network security protection systems are improved. The results indicate that this method can provide security guarantees for campus networks [18]. Cheng *et al.* established an evaluation model for the network security evaluation in the industrial internet using Evidence Reasoning (ER) algorithm and Confidence Rule Base (BRB) method. The results indicate that this method has strong applicability for evaluating the network security status of complex industrial internet systems. It can accurately reflect the actual network security status of industrial internet systems, providing safe and reliable suggestions for network administrators to take timely response measures [5].

In summary, the Apriori algorithm has been widely applied in the data mining, providing an effective method for data collection and analysis. With the development of internet technology, network security has received special attention. The existing network security technologies still have issues such as precision and low applicability, which cannot better meet the security management needs of different network purposes. Based on the advantages of Apriori algorithm in data processing, a warning and defense model for network viruses is constructed by combining this technology. It is hoped that this method can

better achieve network security maintenance, providing effective support for network virus intrusion detection.

3 Construction of a Warning and Defense Model Based on Distributed Network Viruses

3.1 Construction of Data Association Rules Based on Improved Apriori Algorithm

The distributed network security warning system can effectively detect network viruses, which plays a crucial role in network detection. It can achieve decentralized detection of network points and integrate data from various branches. Network security intruders are promptly identified [1,16]. Effective measures are taken to resist various attacks and ensure information security. The distributed network security early warning system can detect the status of the system in real-time to identify potential intrusion objects, and provide timely feedback on the detection situation to the system center. Possible network insecurity situations are analyzed and predicted to provide decision support for relevant personnel [2]. Before building a network security warning system, the Apriori algorithm is used to mine association rules between data.

The Apriori algorithm has multiple flaws. Therefore, it needs to be optimized and improved. Improvements can be divided into algorithm improvement and application improvement of early warning systems. The algorithm's self-improvement includes four processes, namely combining hash technology, partitioning, selecting samples, and counting dynamic itemsets. Combining hashing technology refers to the frequent addition of k-term set L_k to achieve set diversification. At the same time, sets that do not meet the threshold requirements are eliminated, greatly reducing the number of sets. Partition refers to dividing the database confidence into independent parts by scanning data. There is no overlap or overlap between the contents of each section. The sample selection process requires filtering frequent itemsets that meet the specifications in the original dataset, which can greatly improve the system's work efficiency. Dynamic itemset counting requires continuous scanning of database information, achieving data block partitioning of database information, and dynamically evaluating each data block. The application improvement of the early warning system involves three processes, namely filtering rules, introducing main attributes, and introducing interest. In the data mining, the information is enormous. Many data rules may conflict with system requirements. Therefore, these conflicting rules need to be deleted, which is called filtering rules. Reducing these rules can effectively improve the efficiency of system intrusion detection. When processing data, it mainly generates attribute information of the data. These information includes link establishment time, port specific information, link address, etc. Too many attributes will cause data redundancy and increase the work intensity. Therefore, attributes can be further subdivided into central and subordinate attributes. These two types of attributes can be further subdivided according to their respective importance when making value judgment [13]. There is high association between data rules. Traditional data association algorithms have significant advantages in mining these rules. Therefore, interest is introduced to improve the association degree of rules. The interest level is shown in Formula (1).

$$I_R = support(AB)/support(A)support(B) \tag{1}$$

A and B represent two different pieces of data, respectively. The importance of rules is determined by the interest level. When the interest degree is close to 1, it indicates that the practical significance of this rule is relatively small. When the interest degree is much greater than 1, it indicates that the rule has significant practical significance. The improved algorithm flowchart is shown in Figure 1.

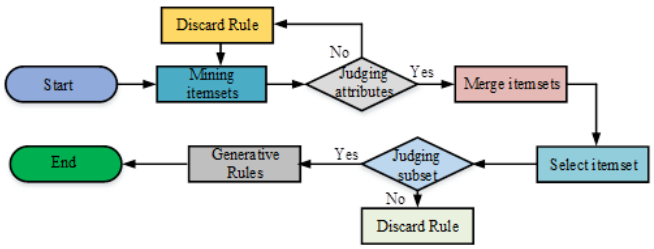


Figure 1: Optimization Apriori algorithm flowchart

The Apriori is used to determine frequent itemsets. The frequent itemsets in databases have strong correlation. These association rules need to satisfy both support and credibility. The support level is shown in Formula (2).

$$S = support(A \cup B) \tag{2}$$

The credibility can be expressed by Formula (3).

$$\begin{aligned} Confidence(A \Rightarrow B) &= P(AIB) \\ &= \frac{sup\_count(A \cup B)}{sup\_count(A)} \end{aligned} \tag{3}$$

In Formula (3), the corresponding number of transactions is included. The corresponding centralized explanations can be made based on the corresponding association rules. The corresponding frequent itemsets generate a large number of non empty subsets. Therefore, a large number of association rules can be generated during the mining process. The mining structure of association rules is shown in Figure 2.

The performance requirements of the system vary at each stage. During the learning, the system utilizes relevant technologies to mine data and generate corresponding system rules. By searching for data types and establishing connections between data, a set of association

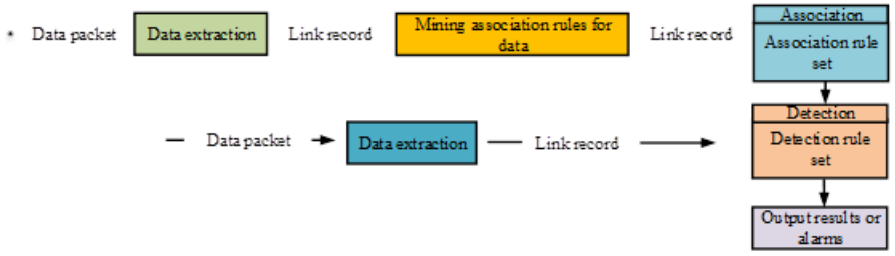


Figure 2: Basic process of association rule data mining

rules can be generated. After the association rule set is established, it is necessary to filter the association rule set. Then it is filtered to retain the relevant sets that meet the requirements. During detection, rules that meet the conditions are searched in the rule set.

3.2 Design of a Security Warning System Based on Distributed Network Viruses

The distributed network security early warning system can achieve various functions, including intelligent analysis of network security, network threat assessment, threat root causes analysis, and threat spread assessment. The framework of a distributed network security early warning system is shown in Figure 3.

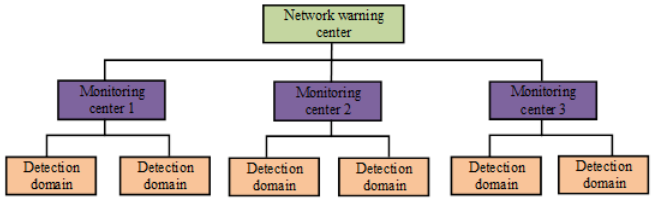


Figure 3: Structure of network warning center

Based on the above association rule mining, data is classified and connected in a distributed network security early warning system according to the mining rules. The distributed network security early warning system involves multiple monitoring centers. The main function of these monitoring centers is to achieve real-time collection of network data. Based on the data situation, the entire system is warned.

The overall design of the distributed network security early warning system is divided into two parts, namely the detection and monitoring center module and the network area early warning center module. Timely defense measures can be used to detect the network during hacker attacks, collect relevant data to analyze attack trends, and activate warning functions. After the warning is completed, corresponding defense measures can be taken [20]. The network monitoring center will process data that enters the network at all times. The network warning center can obtain corresponding data and detect abnormal information in the data. The detection methods involve misuse detection and anomaly detection.

The primary task of the regional warning center is to accurately receive the warning data from the monitoring center. Then, the received warning information is compared with local information to discover the correlation between the two. Thus, the timely alarm function of the early warning center can be achieved. The specific function of the warning center module is to upload warning information. Based on IP address mapping, the detailed information of the region is accurately searched. The data communication between different regions is independent of each other. When a certain warning center malfunctions, other warning centers need to provide warning prompts.

The alarm information contains a large amount of data, and the data content is complex. Therefore, it is necessary to identify alarm information so that all types of alarm information can be uniquely identified. The system will mark all alarm information with a unified Alert ID to achieve the display function of unique identification. The Alert ID tag mainly consists of information intrusion environment and sending information. The structure of alarm information is shown in Figure 4.

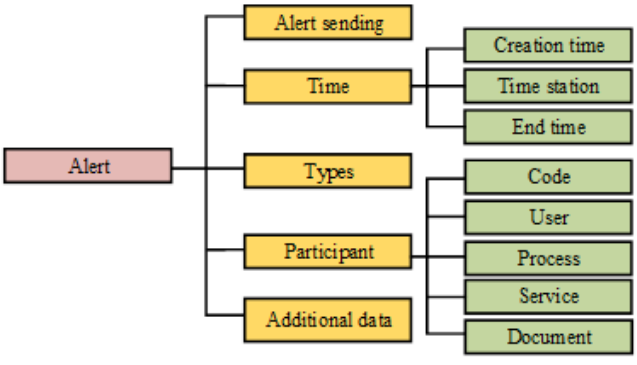


Figure 4: AISM types

The sender of alarm information is unique. It has only one Agent ID. The time and date attributes of alarm information are described through time instances. Time is composed of three sub types, namely creation time, time station and end time. The end time represents the time

when the system experienced an abnormal situation [21]. A timestamp can record in detail the time when an exception occurred. Therefore, it can effectively avoid information confusion between alarm messages. Type is a description for the type of virus attack, which includes two specific attributes, CVE and EventType. With a default value of 0, CVE can be widely accepted. CVE can improve the standardization and sharing of data. At the same time, increasing the operability of data and consistency of reports in IDS provides convenience for IDS comparison [11]. In the alarm information, it is necessary to strictly prevent the abuse of activity information. Defining attack sources and targets separately may result in overlapping alarm information. Therefore, the source and target need to be uniformly defined. The unified type is used to describe these two roles. According to different attributes, these two types are strictly distinguished. This approach not only covers more data content, but also simplifies the alarm information model.

The attack source and target content are different. Therefore, there are significant differences in the expressions of IDS included, such as different network addresses, user names, and host names. Different attack sources provide different expressions, including attack sources based on host names. Data nodes, specific usage accounts, ID addresses, and corresponding service types that can provide timely descriptions of alarm entities. The service type describes the specific information of the attack source and target service [7]. In a sense, services refer to the resources available to users through the network. The attribute types of a service are service address and service port. In addition, services can also be learned as the requesting and receiving services. The specific expression of ports is shown in Table 1.

Table 1: Port definition

Code	Related explanations
la	Represents data ports located between 1-1024
“500	Represents data ports located between 500 and 1024
ha	Represents data ports less than 6000
”!6000	Represents the non functional usage of port 6010
any	Represents various forms of data port usage

Finally, the simplification of alarm information involves classifying and summarizing various relationships between information, eliminating and merging redundant and duplicate information, and reducing the number of alarms [15]. When there is too much alarm information, it is correlated in different ways. Firstly, filter the alarm information that does not meet the conditions, as shown in Formula (4).

$$[M, p(M)D] \Longrightarrow \infty \tag{4}$$

$[p(M)]$ indicates that warning M does not meet the alarm condition D. Secondly, merge duplicate alarm information, as shown in Formula (5).

$$[S, S, \cdots, S] \Longrightarrow S \tag{5}$$

Based on the above operations, alarm information can be effectively streamlined.

4 Analysis of Apriori Algorithm Optimization and System Verification Results

To verify the performance of the improved Apriori algorithm proposed in the research, the spatial complexity of the improved Apriori algorithm is first analyzed. The temporary storage space required for the improved Apriori algorithm is shown in Figure 5. The small number of intermediate itemsets indicates that the Apriori algorithm has high computational efficiency. At the same time, it consumes less storage space. This indicates that the improved Apriori algorithm consumes less space.

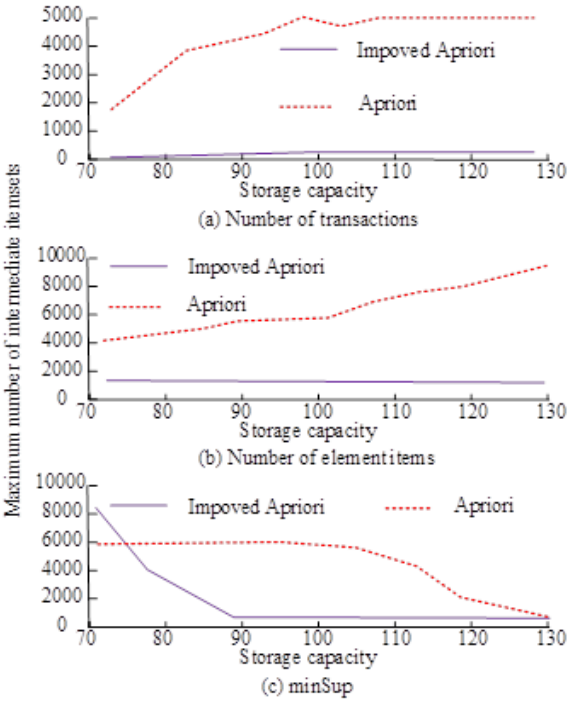


Figure 5: Comparison of storage space after improved Apriori algorithm

The performance of data mining algorithms greatly affects the efficiency of detection centers. Therefore, the traditional Apriori algorithm is compared with the optimized Apriori algorithm. The WinPcap network data capture tool intercepted 5000 links. For the improved Apriori algorithm, the support level is 0.4, the confidence level is 0.6, and the interest level values are different.

These two algorithms are used to classify 5000 link situations, with 3500 data used for algorithm training and the remaining data used for algorithm validation. Several data attributes are selected for mining work. The experimental results are shown in Figure 6. In Figure 6, as the interest value increases, the amount of mining data for optimizing the Apriori algorithm also increases. Although the values of support and credibility are fixed at this time, the rule data still decreases. Therefore, effective data mining rules can only be obtained by setting interest threshold data reasonably. The optimized Apriori algorithm can effectively enhance the rule data usage value.

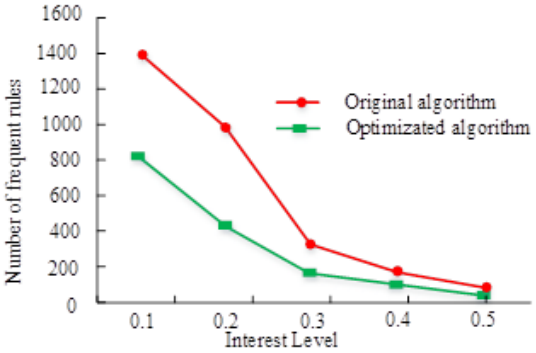


Figure 6: Comparison of data mining between traditional Apriori and optimized Apriori

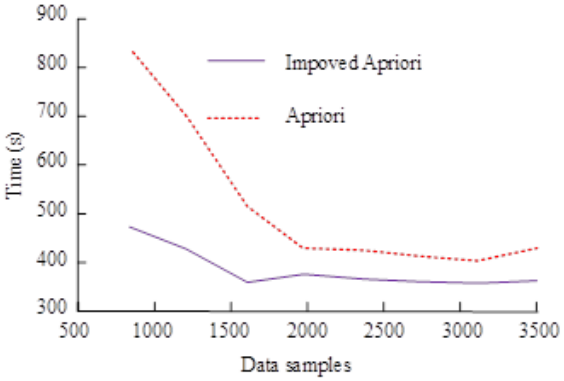


Figure 7: Comparison of improved Apriori algorithm runtime

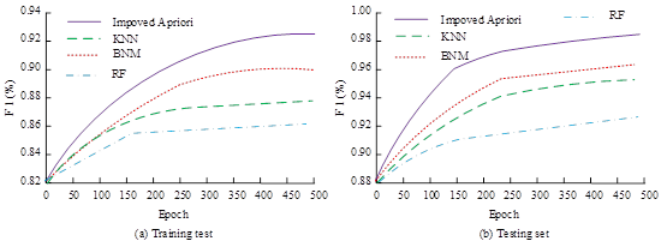


Figure 8: F1 value of four models

The given dataset is used to test the running time of the improved Apriori algorithm. The results obtained are shown in Figure 7. In Figure 7, under different data sample conditions, the running time of the improved algorithm is lower than that of traditional methods. Specifically, when the sample data is small, the difference in runtime between the two is more significant. When the sample data is 1800, the difference in running time is significantly reduced. When the sample data is 3500, the running time of the Apriori algorithm is 470s. The running time of the improved Apriori algorithm is 360s. The improved method performs significantly better than traditional methods.

Common data mining algorithms include K-Nearest Neighbor (KNN), Naive Bayesian Model (NBM) and Random forest (RF). The performance of the above data mining algorithm is verified. The results obtained are shown in Figure 8. The F1 values of the four models on the training and testing sets are shown in Figure 8. In Figure 8(a), on the training set, the F1 value of the optimized Apriori algorithm reaches 0.931, which is 0.061, 0.071, and 0.068 higher than the KNN, NBM, and RF models, respectively. In Figure 8(b), on the test set, the F1 value of the optimized April algorithm reaches 0.985, significantly higher than other methods. It indicates that the proposed method has better performance.

The UNSW-NB15 is selected to validate the performance of the proposed method in the study. This dataset has nine types of attacks, namely, DoS, Fuzzers, Analysis, Backdoors, Exploits, Generic, Reconnaissance, Shellcode and Worms. The number of records in the training set is 175,341 records and the testing set is 82,332 records from the different types, attack and normal. In this dataset, 7000 and 3000 pieces of data were randomly selected in a 7:3 ratio for experiments. The experimental datasets used include DoS, Fuzzers, Analysis, Backdoors, Exploits and Worms. The application effectiveness of this method is evaluated based on the detection rate (DR), false detection rate (FDR), and miss detection rate (MDR) in different scenarios. The DR, FDR, and MDR are shown in Figure 9. In different scenarios, the proposed method has higher DR. Taking the results of the Dos dataset as an example, the DR, FDR, and MDR of this model are 82.03%, 11.12%, and 18.75%, respectively. This indicates that the proposed method has better adaptability and good performance in different scenarios. This method can achieve significant results in practical applications and meet the warning needs of different scenarios.

In addition, to verify the application effectiveness of the early warning system, the experiment takes the internal network of Company A as the research object. Two important detection points are set up in the network. The layout of experimental detection points is shown in Figure 10.

This test sets up two detection domains. The data obtained from the company's testing in November 2018 is used as an example for research. In the test, the alarm

Table 2: Early warning conditions at each stage

Time	Initial alarm	Processed alarm	Number of notifications	Number of alerts	Number of alarms	Emergencies
2018-11-20 01:00	21	16	20	1	2	0
2018-11-21 04:00	58	51	25	1	1	0
2018-11-22 06:00	74	72	67	1	1	0
2018-11-23 11:00	37	36	31	1	1	0
2018-11-24 17:00	1540	21	1501	1	1	0
2018-11-25 18:00	1701	18	1692	1	1	0
2018-11-26 21:00	26	14	21	1	1	0
2018-11-27 22:00	180	13	174	1	1	0
2018-11-28 23:00	871	10	866	1	1	0
2018-11-29 02:00	84	4	81	1	1	0
2018-11-30 14:00	160	5	150	1	1	0

Figure 9: Performance comparison for this model in different scenarios

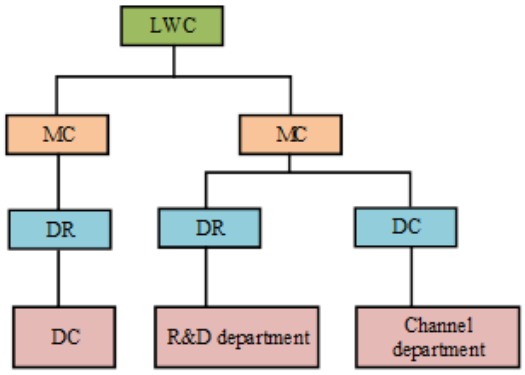


Figure 10: Test environment structure

data generated in each time period, the original data of the monitoring center and the early warning data generated by the early warning center need to be counted in detail. The statistical period is 1 hour. The obtained statistical results are shown in Table 2. From Table 2, there is a large amount of alarm data before preprocessing. However, after being processed by the system, the amount of alarm data significantly decreases. This indicates that the early warning system filters the data information for the alarm process. In addition, the maximum number of alarms in this early warning system can reach 12000, which is consistent with the actual situation. In summary, the early warning system has high practicality.

5 Conclusion

In the research of distributed network virus warning and defense, the data mining algorithms is extremely important. On the basis of the Apriori algorithm, the shortcomings of the traditional Apriori algorithm are optimized. A distributed network warning and defense system based on the Apriori algorithm is established to effectively reduce the network virus infection rate and improve data security. In the research, the data mining performance of traditional Apriori algorithm and optimized Apriori algorithm are compared. The internal network of Company A is used as the research object. The performance measurement results of the early warning system indicate that there is more alarm data before preprocessing. However, after using the early warning and defense system, the amount of alarm data significantly decreases. This indicates that the early warning and defense system can effectively block network viruses, thereby reducing the amount of alarm data. In addition, the maximum number of alarms in this early warning system can reach 12000, which is in line with the actual situation. From the above, the proposed method can deeply mine the potential rules of relevant data and achieve network security

warning. The effectiveness of this early warning system is high and can meet the needs of different virus attack scenarios. Although this research can provide a method reference for the early warning and defense of distributed network viruses, there are still issues with limited sample data and insufficient experimental accuracy in the research. In addition, there are various forms and types of network viruses, which are constantly being updated with the development of technology. Therefore, in future research, more intelligent methods can be attempted to optimize warning systems, achieve more accurate warnings, and optimize network security management.

References

[1] A. Ali, I. S. Muhammad, F. Umbreen, A. Nauman, I. Zafar, R. Ali, R. Muhammad, R. M. Azizur, "Optimal existence of fractional order computer virus epidemic model and numerical simulations," *Mathematical Methods in the Applied Sciences*, vol. 44, no. 13, 10673-10685, 2021.

[2] C. Aníbal, H. Fernando, P. Manuel, "Sufficient conditions for the existence of positive periodic solutions of a generalized nonresident computer virus model," *Quaestiones Mathematicae*, vol. 44, no. 2, pp. 259-279, 2021.

[3] F. Aur, S. S. Deniz, K. Yazici, "Visual preferences assessment of landscape character types using data mining methods (Apriori algorithm): the case of Altnsa and Inkoy (Van/Turkey)," *Journal of Agricultural Science and Technology*, vol. 22, no. 1, pp. 247-260, 2020.

[4] S. Barnabás, R. Zoltán, "Cyber Security analysis of smart buildings from a cyber security architecture point of view," *Interdisciplinary Description of Complex Systems*, vol. 21, no. 2, pp. 141-147, 2023.

[5] M. Cheng, S. Li, Y. Wang, G. H. Zhou, P. Han, Y Zhao, "A new model for Network security situation assessment of the industrial Internet," *Computers, Materials and Continua*, vol. 75, no.2, pp. 2527-2555, 2023.

[6] L. Federico, F. Alberto, "From DevOps to DevSecOps is not enough. CyberDevOps: an extreme shifting-left architecture to bring cybersecurity within software security lifecycle pipeline," *Software Quality Journal*, vol. 31, no. 2, pp. 619-654, 2023.

[7] N. J. Gonçalves, S. H. Rodrigues, T. T. M. Monteiro, "Preventing computer virus prevalence using epidemiological modeling and optimal control," *Discontinuity, Nonlinearity, and Complexity*, vol. 9, no. 2, pp. 187-197, 2020.

[8] J. Li, "Research on an optimised encryption algorithm for network information security communication," *International Journal of Communication Networks and Distributed Systems*, vol. 29, no. 1, pp. 31-46, 2023.

[9] N. Liao, Y. Song, S. Su, X. S. Huang, H. L. Ma, "Detection of probe flow anomalies using information entropy and random forest method," *Journal of Intelligent and Fuzzy Systems: Applications in Engineering and Technology*, vol. 39, no. 4, pp. 433-447, 2020.

[10] M. B. Luther, "Behavioural analytics in cyber security for digital forensics application," *International Journal of Computer Science and Information Technology*, vol. 15, no. 1, pp. 83-90, 2023.

[11] V. MadhuSudanan, R. Geetha, "Dynamics of epidemic computer virus spreading model with delays," *Wireless Personal Communications*, vol. 115, no. 14, pp. 2047-2061, 2020.

[12] D. Natalia, L. Salvatore, "Apriori algorithm for association rules mining in aircraft runway excursions," *Civil Engineering and Architecture*, vol. 8, no. 3, pp. 206-217, 2020.

[13] T. Niksa-Rynkiewicz, M. Landowski, P. Szalewski, "Application of Apriori algorithm in the lamination process in Yacht production," *Polish Maritime Research*, vol. 27, no. 3, pp. 59-70, 2020.

[14] J. Niu, "Design of automatic monitoring system for network information security in cloud computing environment," *International Journal of Information and Computer Security*, vol. 21, no. 1-2, pp. 19-34, 2023.

[15] N. Özdemir, S. Uçar, İ. B. B. Eroğlu, "Dynamical analysis of fractional order model for computer virus propagation with kill signals," *International Journal of Nonlinear Sciences and Numerical Simulation*, vol. 21, no. 3-4, pp. 239-247, 2020.

[16] K. B. Sunil. G, Sunita. S, Tanuja, "Dynamics of virus-patch model with latent effect," *International Journal of Computer Mathematics*, 99, no.9, pp. 1754-1769, 2022.

[17] F. Y. Wang, "Analysis and research on China's industrial Internet security development," *International Journal of Plant Engineering and Management*, vol. 27, no. 4, pp. 244-252, 2022.

[18] S. Wang, G. Long, "Research on campus network security protection system framework based on cloud data and intrusion detection algorithm," *Soft Computing*, vol. 27, no. 10, pp. 6835-6844, 2023.

[19] X. Wang, D. Huang, X. Zhao, "Design of the sports training decision support system based on improved association rule, the Apriori algorithm," *Intelligent Automation and Soft Computing*, vol. 26, no. 4, pp. 755-763, 2020.

[20] H. Xie, "Research and case analysis of Apriori algorithm based on mining frequent item-sets," *Open Journal of Social Sciences*, vol. 9, no.4, pp. 458-468, 2021.

[21] F. F. Yang, Z. Z. Zhang, "Dynamics of a nonlinear SIQRS computer virus spreading model with two delays," *Aims Mathematics*, vol. 6, no. 4, pp. 4083-4104, 2021.

[22] Y. Zhou, "Design and implementation of book recommendation management system based on improved Apriori algorithm," *Intelligent Information Management*, vol. 12, no. 3, pp. 75-87, 2020.
[23] L. Zhu, "Implementation of web log mining device under Apriori algorithm improvement and confidence formula optimization," *International Journal of Information Technology and Web Engineering*, vol. 15, no. 4, pp. 53-71, 2020.

Biography

Wang Lin obtained a Bachelor's degree in Computer Science and Technology from Henan University of Technology in 2005. In 2008, she obtained a Master's degree in Computer Technology from Wuhan University of Technology. At present, the main research direction is computer software and networking.

Wang Chuang obtained a Bachelor's degree in Computer Technology from Wuhan University of Chemical Technology in 2004 and a Master's degree in Computer Technology from Wuhan University of Technology in 2009. At present, the main research direction is computer software and networking.

Network Anomaly Attack Detection System Based on Incremental Learning Combined with SCV and SVM Algorithms

Lijie Li

(Corresponding author: Lijie Li)

School of Information and Intelligent Engineering, Ningbo City College of Vocational Technology
Ningbo 315100, China
Email: lilijieei@163.com

(Received Aug. 10, 2023; Revised and Accepted June 2, 2024; First Online Aug. 17, 2024)

Abstract

Network anomaly attack detection refers to analyzing network behavior and extracting features to predict and judge it. Since network anomaly attacks are a dynamic and uncertain behavior, their behavior is fuzzy and complex, so it is difficult for traditional data flow analysis methods to detect network anomaly attacks effectively. To improve detection accuracy, a network anomaly attack detection system based on incremental learning combined with a candidate support vector algorithm is proposed to optimize the support vector machine. Based on the historical data, the supporting candidate vector is proposed. This study adopts the retention strategy to classify the old samples. Combined with the Kuhntak condition, the new samples are screened, and the number of training samples in the incremental process is reduced to improve the efficiency of the detection model. The updated algorithm has a higher anomaly detection rate than the classical support vector algorithm. The normal false alarm rate of the model is 91.7%, and the abnormal detection rate is 3.5%. The classification effect of the final detection model is better than that of the classical support vector algorithm. This method has good accuracy and real-time performance in predicting and judging network anomaly attacks.

Keywords: Candidate Support Vector; Incremental Learning; Network Anomaly Detection; Support Vector Machine

1 Introduction

Currently, network security has faced severe challenges, which is difficult for traditional detection technologies to find problems [1]. Machine learning algorithms can automatically learn and adjust models according to different data types, so useful features can be extracted from network data [2]. Support Vector Machine (SVM) divides the input space into a high-dimensional characteristic space. Then a linear model is built in the high-dimensional feature space, and the nonlinear mapping of the training sample is obtained based on this linear model [5]. Support Candidate Vectors (SCVs) are widely used in machine learning because they do not require any assumptions, the distance between data points and feature Spaces and the neighborhood relationships between data points [8]. Incremental learing differs from traditional machine learning in that it does not require a model to be specified in advance, but simply makes predictions about the input data at each iteration. Therefore, it has a wide range of applications in bioinformatics and biomedical fields [10].

Aiming at the problems of the network traffic detection model, such as weak adaptive ability, long update cycle and poor continuous learning ability, this paper proposes the research of network anomaly attack detection system based on SCV optimization SVM algorithm based on incremental learning. The aim of the research is to promote the healthy development of the network environment and improve the resilience of the network attack defense technology. The research content is split into five parts. The first part is the introduction, which describes the impact of malicious traffic brought by network attacks under the contemporary background of the development of network information technology. The second part is the related work, the application of network anomaly detection technology and SVM algorithm in various fields, and the study status of many scholars on the two technologies. In the third part, according to the characteristics of large scale, fast change, variety and high dimension of network traffic data, the network anomaly detection model based on SCV optimization SVM in incremental learning is studied.

In the first section, incremental learning, KKT condition and SVM algorithm are studied and analyzed, and incremental learning is introduced into SVM algorithm to make SVM classification algorithm more suitable for real network environment. In the second section, two clas-

sical incremental support vector machine algorithms are introduced into SVM algorithm, and the network anomaly attack detection model of SCV-KKT-ISVM algorithm is constructed. In the fourth part, the accuracy of network detection between the model and the classical SVM is analyzed by comparison experiment, as well as the effectiveness of the detection model and the accuracy of classification. The fifth part is the summary and prospect of the research methods and results.

2 Related Work

Currently, the scale of the network continues to expand, whose exposed vulnerabilities are gradually increasing. With the increasing number and types of network attacks, many scholars have carried out research on the detection technology of network anomaly attacks. Tang et al. [13] proposed a framework consisting of a two-stage detection module and a mitigation module. It combined port traffic characteristics with a classifier that counts traffic based on a flow table to accurately detect low-rate denial-of-service (LDoS) attacks. The framework distinguished between a victim port and a benign port by calculating an exception score for each port. The framework was capable of identifying LDoS attacks. Han et al. [14] suggested a network combining sparse autoencoder and kernel. Compared with other feature extraction methods, the recognition rate of the proposed method could reach 98.68%, and the average dimensionality reduction time was 5.59 s, which possessed superior operation efficiency. Subasini et al. [15] proposed a network anomaly detection technique to explore the fractal characteristics of network traffic. Then, based on fractal theory and wavelet analysis, the network traffic model was proposed. The autocorrelation function of the proposed model could reach the mean square error of 4.762×10-4, which verified that it can alleviate the influence of non-stationarity of network traffic on modeling accuracy, which performed well.

Traditional network defense technology has been unable to meet people's requirements and status quo of network security. Therefore, using machine learning algorithms with self-learning ability to detect network anomaly attacks has become the research direction of many scholars in this field. Ding et al. [17] proposed an incremental learning strategy that adds normal new data to the policy. To solve the problem of slow calculation velocity, a dynamic downsampling method was proposed. This method could decrease calculation time by at least 80%. Incremental learning could maintain high estimation accuracy and low false positive rate over the long run. By using current machine learning techniques and analysis, Wang et al. [18] highlighted the relationship between Autism Spectrum Disorder (ASD) and a family history of jaundice or autism in children. The four subgroups of data were split into training data and test data by hierarchical cross-validation method, and provided to sequence-minimum-optimized SVM mixed classifier. The success rate of this method was 100%, and the detection rate of ASD was 95.52%. Based on jaundice and family history of autism, this model made a very meaningful explanation for the more effective detection of ASD in children.

To sum up, machine learning and network anomaly detection technology have achieved a lot. However, as the scale of traffic in the network environment gradually expands, the data information gradually surges with the increase of traffic. Therefore, there is a continuous learning requirement for detection models. To enhance the detection efficiency and effectiveness of the network anomaly detection method, this study proposes a network anomaly attack detection system based on incremental learning SCV optimization SVM algorithm.

3 Optimize the Network Anomaly Detection Based on Incremental Learning Combined with SCV SVM Algorithm

Network anomaly detection refers to the amount of data transmitted between two or more communication nodes in a real network environment. However, with the expansion of data scale, high latitude and rapid change of data, intelligent network anomaly detection technology has gradually become a development trend [4,19,20]. The standard SVM algorithm uses Batch Learing to simply merge old samples and new samples, and then trains the combined data set to get a new detection model. SCV refers to non-SV samples that are not retained in the training data set and affect subsequent models. Based on this, this study proposes a network anomaly detection method based on incremental learning and SCV optimization SVM algorithm.

3.1 Research on Incremental Learning-Based Support Vector Machine Algorithm

SVM is only suitable for small sample training. As the sample size increases, the time required to train the model will also increase, so incremental learning is introduced into SVM algorithm. However, in the real network environment, the feature similarity of data samples is extremely high, so it is tough to find a suitable hyperplane in the sample space to make the samples linearly separable [21, 22]. Therefore, the correlation between KKT (Karush Kuhn Tucker) condition and sample distribution in SVM algorithm is studied. When maximizing the interval, most samples are classified normally, but a small number of samples are allowed to be misclassified. The optimized expression for this process is shown in Equa-

tion (1).

$$\begin{cases} \min_{\bar{w},b,\zeta_i} \frac{1}{2}||\bar{w}||^2 + C\sum_{i=1}^{m}\zeta_i s.t.\zeta_i \geq 0 \\ y_i(w^T \cdot x_i + b) \geq 1 - \zeta_i \end{cases} \quad (1)$$

In Equation (1), C represents the penalty factor. Parameters that need to be set in advance, ζ_i represents the relaxation factor; $\zeta_i, b, \bar{w}$ are the parameters. The four types of sample distribution under KKT conditions are shown in Equation (2).

$$\begin{cases} \lambda_i = 0 \rightarrow y_i f(x_i) \geq 1 \\ 0 < \lambda_i < C \rightarrow y_i f(x_i) = 1 \\ \lambda_i = C \rightarrow y_i f(x_i) \leq 1 \end{cases} \quad (2)$$

In Equation (2), $y_i f(x_i) \geq 1$ satisfies the necessary and sufficient conditions of KKT condition. As long as the distribution is outside and above the classification interval $y_i f(x_i) = 1$, and the sample is correctly classified, the sample is a KKT condition. The SVM algorithm determines that the model is only related to the number of SV samples, which is not related to the spatial dimension of SV. Therefore, KKT algorithm is suitable for incremental learning. Network traffic is characterized by large scale, high dimension and indivisibility, resulting in non-linear, divisible and highly overlapping traffic data [23]. SVM algorithm has the advantage of solving the problem of high latitude and linear indivisible data, and the principle of this algorithm is shown in Figure 1.

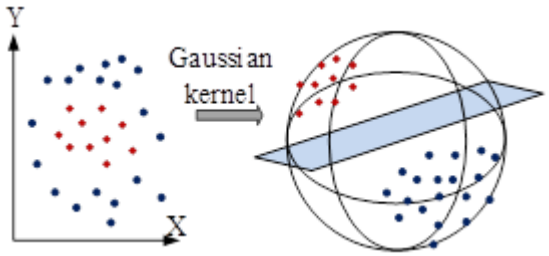


Figure 1: Principle and structure of SVM algorithm

The computational complexity of SVM algorithm is only related to the number of samples. After the over-primitive problem is transformed into a dual problem, the computational complexity is only related to the number of samples, and has nothing to do with the spatial dimension, so as to solve the high-dimensional problem of data [6, 24]. The distance expression from point (x, y) to the decision boundary line is shown in Equation (3).

$$d = \frac{y \cdot (w \cdot w + b)}{||w||_2} \quad (3)$$

In Equation (3), if both w and b are enlarged meanwhile, the numerator and denominator will also be enlarged meanwhile, but the objective function is not affected. Thus, the molecule is enlarged or compressed to equal 1, and its maximization equation is shown in Equation (4).

$$\max \frac{1}{||w||_2} s.t y^i(w \cdot x^i + b) \geq 1(i = 1, 2, \cdots, m) \quad (4)$$

In Equation (4), when $w \cdot x + b = 0$, that is the hyperplane is separated, and the equation is optimized, as shown in Equation (5).

$$\min(\frac{1}{2}||w||_2^2) s.t y^i(w \cdot x^i + b) \geq 1(i = 1, 2, \cdots, m) \quad (5)$$

The low dimensional space of the nonlinear separable samples is converted into the high dimensional space, and then the optimal classification hyperplane of the algorithm is found in the high dimensional space [3, 7, 11]. Introducing incremental learning into the algorithm can make machine learning more intelligent and more in line with the actual network traffic data changes [16]. The classification and process of incremental learning are shown in Figure 2.

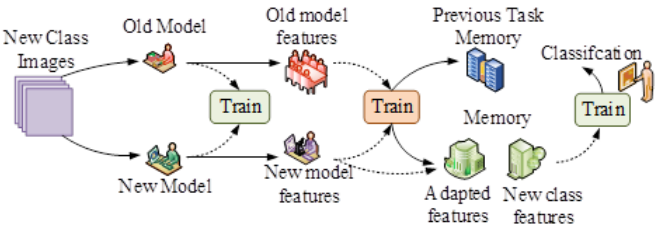


Figure 2: Classification and process of incremental learning

Incremental learning refers to acquiring new knowledge from new data, consolidating old knowledge, and discarding useless knowledge when new data is added [9, 12, 25]. In fact, the change of model parameters is limited by adding a penalty term to the loss function, so as to hinder the change of important parameters on the old task. The mathematical expression is shown in Equation (6).

$$\theta_s^*, \theta_o^*, \theta_n^* \leftarrow \arg\min_{\hat{\theta_s},\hat{\theta_o},\hat{\theta_n}} (\lambda_o L_{old}(Y_o, \hat{Y_o}) + L_{new}(Y_n, \hat{Y_n}) + R(\hat{\theta_s}, \hat{\theta_o}, \hat{\theta_n})) \quad (6)$$

The expression defining the loss function by Equation (6) is shown in Equation (7).

$$Loss = \lambda_o L_{old}(Y_o, \hat{Y_o}) + L_{new}(Y_n, \hat{Y_n}) + R(\hat{\theta_s}, \hat{\theta_o}, \hat{\theta_n}) \quad (7)$$

For the new task, the mathematical expression of L_{new} is shown in Equation (8).

$$L_{new}(Y_n, \hat{Y_n}) = -Y_n \log \hat{Y_n} \quad (8)$$

Equation (8) represents prediction under new tasks, and incremental learning is the new objective function obtained after learning each additional subset of data, as shown in Equation (9).

$$M_i = f(D_i, M_{i-1}) \quad (9)$$

In Equation (9), D_i represents the data set; M_{i-1} denotes the model; M_i represents the objective function obtained by training the model. Incremental learning is introduced into SVM algorithm, called Incremental Support Vector Machines (ISVM) algorithm, which can determine the SV set of the old model classification hyperplane and the new data to participate in the training of the new model. In the incremental process, the ISVM algorithm for screening new samples based on KKTT conditions is called KKT-ISVM algorithm, and the specific process of this algorithm is shown in Figure 3.

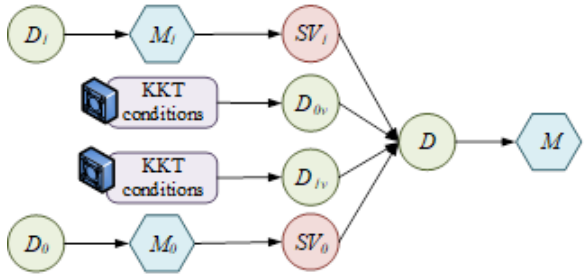


Figure 3: The process of incremental learning for KKT-ISVM algorithm

Incremental learning does not retain the historical data after training, which can reduce the pressure of historical data on storage space in the real network environment. For the training model, incremental learning also reduces the computational complexity and time. The KKT conditions for model D_o, model M_o and model SV_o are obtained by training the data set M_o. Through repeated training, finally training $\{SV_0 \cup D_{0v} \cup SV_1 \cup D_{1v}\}$ to get the final model M after the first incremental training. The algorithm has the best generalization ability for unknown samples and the optimal classification hyperplane. The new sample data set is filtered by KKT condition, and the whole new data set is reduced, so as to reduce the redundancy of new data and relieve the pressure of data storage in memory.

3.2 Network Traffic Anomaly Detection Model Based on Incremental Learning SCV Optimization SVM Algorithm

ISVM introduce incremental learning into support vector machine algorithms, but do not retain the historical training data set, which will affect the non-SV samples that are subsequently updated by the model [17]. Therefore, to retain most of the samples with potential value, SCV algorithm is proposed to improve the updating speed of the model. To retain ISVM introduce incremental learning into support vector machine algorithms, but do not retain the historical training data set, which will affect the non-SV samples that are subsequently updated by the model [17]. To mine potential valuable samples from historical data sets, a retention model based on SCV algorithm is proposed. The specific process of this model is shown in Figure 4.

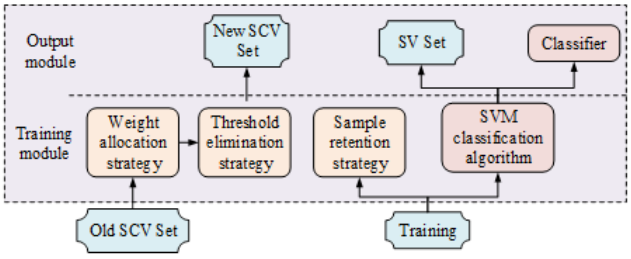


Figure 4: The sample retention model based on SCV algorithm

The sample retention model based on SCV algorithm mainly consists of two parts: training module and output module. The training module includes SCV algorithm and SVM classification algorithm. The SCV algorithm consists of three parts: retention strategy, threshold elimination strategy and weight allocation strategy [22]. After sample training, the classifier, SCV dataset, and SV dataset in the output module can be obtained. Based on the SV dataset, it does not appear at a type of center in the sample space, so it is required to calculate the distance between the center point in the sample space and the sample. First, the mathematical expression of the center point m_i in different sample Spaces is shown in Equation (10).

$$m_i = \frac{1}{n}\sum_{s=1}^{n_i} x_{sk} \tag{10}$$

In Equation (10), i represents the number of categories, whose value range is $i = \{1, 2\}$; n_i denotes the total number of samples in different categories; x_{sk} represents the k-th attribute of sample s. The expression of the Euclidean metric d_{ij} for two sample points x_1 and x_2 is shown in Equation (11).

$$d(x_1, x_2) = \sqrt{\sum_{k=1}^{m} |x_{1k} - x_{2k}|^2} \tag{11}$$

The size of radius R_1 and R_2 in the sample space determines whether valuable samples can be effectively retained. R_1 is the Euclidean distance between the center and the support vector. The mathematical equation of R_2 is shown in Equation (12).

$$R_2 = \frac{w^T \cdot x_i + b}{||w||} \tag{12}$$

In Equation (12), R_2 represents the distance between the center and the hyperplane. For non-SV samples, the conditions in Equation (12) must be satisfied if they are considered as candidate support vectors, and the definition

is shown in Equation (13).

$$d \geq R_1 AND r \leq R_2 \tag{13}$$

In Equation (13), d represents the Euclidean distance between the sample point x and the center point C of the class to which the sample belongs; r denotes the Euclidean distance between the sample point x and the SVM hyperplane. After the detection model is updated, the valuable samples in the historical data set are retained through the retention strategy above SCV algorithm. The probability of becoming SV is determined by the distance from the hyperplane, so the weight of a template in SCV is related to the distance between the hyperplanes, and the definition of its judgment criteria is shown in Equation (14).

$$W_i = \frac{d}{d + R_1} + \frac{r}{r + L} \tag{14}$$

In Equation (14), L represents the distance between SV and the hyperplane; W_i represents the value of the weight of i sample; The greater the value of d, the smaller the value of r, the more likely it is that the sample will become SV in incremental training and the greater the weight assigned. When SV's weight W is 1, its threshold T is defined as shown in Equation (15).

$$T_i = W_i - 1 = \frac{d}{d + R_I} + \frac{r}{r + L} - 1 \tag{15}$$

To allow more data to enter the subsequent incremental learning, the threshold value is appropriately relaxed, and the samples are retained and stored in the SCV aggregation. After the classifier, SV and SCV data sets are obtained through the sample retention model based on SCV algorithm, the algorithm ends. The specific process of incremental learning based on SCV algorithm is shown in Figure 5.

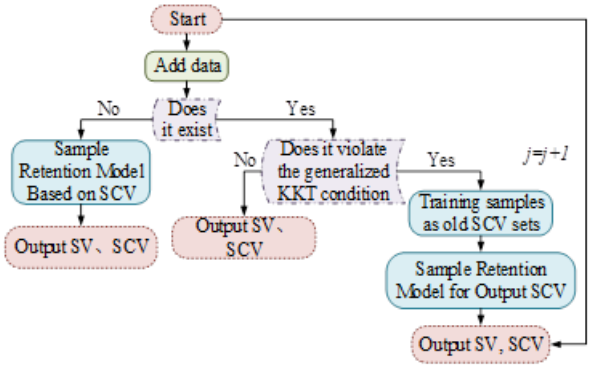


Figure 5: Based on SCV-KKT-ISVM algorithm process

The algorithm determines whether the sample violates the KKT condition by KKT condition, and filters the sample under the generalized KKT condition to keep the sample violating the KKT condition and make it become SV sample. The new samples obtained by screening and the SV dataset obtained by training are taken as the new training set and put into the sample retention model based on SCV algorithm together with the old SCV set. This results in a new classifier, SV dataset, and SCV dataset. SCV algorithm is the non-SV sample that will affect the model in the historical data set, KKT-ISVM algorithm is the new sample that will affect the updated historical data set, representing the SV sample obtained from the classification hyperplane of the old model. Based on this, this study proposes a network anomaly attack detection model based on incremental learning SCV-optimized SVM algorithm, namely SCV-KKT-ISVM algorithm detection model. The model is shown in Figure 6.

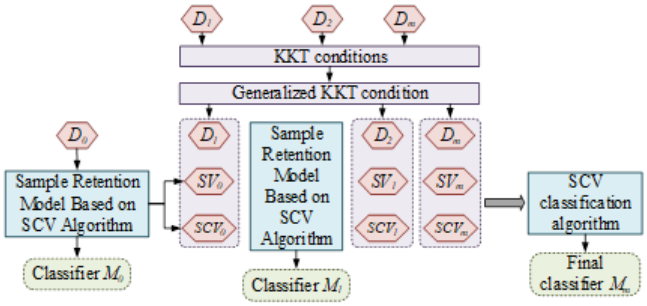


Figure 6: A network anomaly attack detection model based on SCV-KKT-ISVM algorithm

In this model, it is required to first determine whether there are samples in the new data that violate the KKT condition, and there are two cases. First, if it does not exist, it means that the old model has strong generalization ability in the face of new data and can be used as the result of training [13]. Secondly, if it exists, the generalized KKT conditions in the old model are used to screen the new data, and the violated samples are added to the new model for training. Then, non SV samples in the newly added data are filtered through the retention policy in the SCV algorithm. Then the qualified samples are added to the core SCV set by weight allocation strategy and threshold elimination strategy in SCV algorithm.

4 Test and Analysis of Network Traffic Anomaly Detection Model Based on Incremental Learning SCV Optimization SVM Algorithm

To verify the effectiveness and applicability of the proposed model, the SCV-KKT-ISVM algorithm and Batch-SVM algorithm are tested on the KDDup99 dataset and included in the next Department of Defense intrusion detection evaluation plan, which is closer to traffic data in real network environments. The KDDup99 dataset, which is included in the next DOD Intrusion Detection Evaluation Program, more closely resembles traffic data in

real network environments. To ensure the effectiveness of the comparison of the four algorithms, the comparison experiment is conducted in the same experimental condition. The hardware configuration of Intel(R)Core(TM)i7-10710U@1.61GHz and 32.0 GB running memory are used in the experiment. Using Windows x64 operating system, Python development voice, PyCharm development tools, Python development environment. The main evaluation indexes in the experiment are abnormal detection rate, normal false alarm rate and accuracy rate. Auxiliary indexes include training time, actual number of support vectors and test time. The first is to learn after 5 increments, each incrementing 2500 data. Model classification accuracy, training time and actual number of support vectors for SCV-KKT-ISVM algorithm and Batch-SVM algorithm. The experimental results of classification accuracy and time consumed are shown in Figure 7.

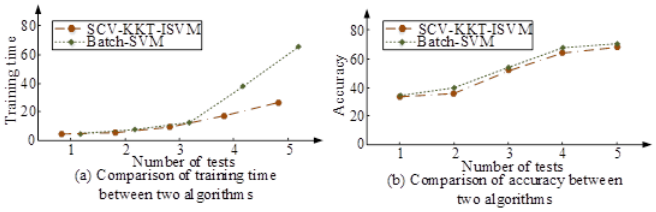


Figure 7: Comparison of training time and accuracy between SCV-KKT-ISVM algorithm and Batch-SVM algorithm

In Figure 7, when the number of incremental learningrises, the time consumed and precision of the detection model under the SCV-KKT-ISVM algorithm and Batch-SVM algorithm show a trend of gradual increase. In Figure 7, the Batch-SVM algorithm takes more time than the SCV-KKT-ISVM algorithm. When the training times are the first 3 times, the difference between the two algorithms is small. However, after three training sessions, the Batch-SVM algorithm shows a surge trend. Because the Batch-SVM algorithm does not pre-process the data and the retrograde, the number of increments increases, so the training time is longer. In Figure 7(b), the accuracy of SCV-KKT-ISVM is roughly the same as that of Batch-SVM trained together. It can be seen that the SCV algorithm retains most of the SV of the model classification hyperplane and effectively retains the potentially valuable samples.

In Figure 8, as the number of training increases, the recall rates of the detection models under both the SCV-KKT-ISVM algorithm and the Batch-SVM algorithm show a decreasing trend. In Figure 8(a), under high-dimensional data, the recall rate of the Batch-SVM algorithm is higher than that of the SCV-KKT-ISVM algorithm, with a value of 0.32, while the recall rate of the SCV-KKT-ISVM algorithm is 0.28, verifying that the accuracy of this algorithm is higher than that of existing advanced algorithms. In Figure 8(b), under low dimensional data, there is a significant difference in the recall rate between the SCV-KKT-ISVM algorithm and the Batch-SVM algorithm. Among them, the recall rate of the Batch-SVM algorithm is 0.56, while the recall rate of the SCV-KKT-ISVM algorithm is 0.43. This algorithm has higher accuracy when facing low dimensional data. For high-dimensional network data, the SCV-KKT-ISVM algorithm can more effectively improve the recall rate, which verifies the effectiveness of the algorithm. To better understand the comparison between the two algorithms, the results are shown in Table 1.

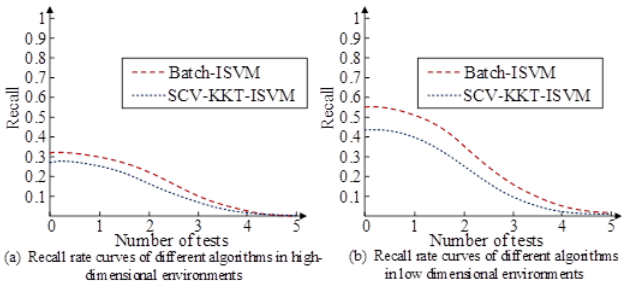


Figure 8: Comparison of Training Time and Recall Rate between SCV-KKT-ISVM Algorithm and Batch SVM Algorithm

In Table 1, with the increase of training times, the training time of SCV-KKT-ISVM algorithm is reduced by 3 times compared with Batch-SVM algorithm. The number of support vectors required is also reduced by about 100; But the accuracy is about the same. In the process of incremental learning, SCV-KKT-ISVM algorithm not only enhances the classification accuracy, but also greatly improves the speed of detecting model update and enhances the ability of model continuous learning. Therefore, this study compares the anomaly detection rate and normal false alarm rate of SCV-KKT-ISVM algorithm, KKT-ISVM algorithm, and ISVM algorithm. The results are shown in Figure 9.

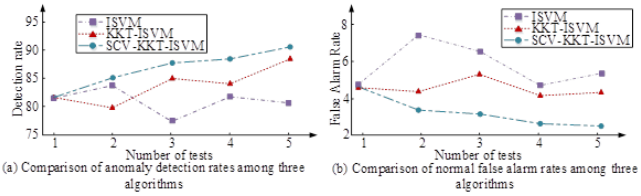


Figure 9: Comparison of anomaly detection rate and normal false alarm rate of three algorithms

In Figure 9, the classification performance of the initial detection model obtained by the three algorithms after the first training is consistent. In the last four incremental training sessions, the detection models obtained by the three algorithms are different. In Figure 9, the detection rates of the KKT-ISVM algorithm and SCV-KKT-ISVM algorithm both show an upward trend with

Table 1: Comparison of algorithm training effects

Training frequency	Batch SVM algorithm			SCV-KKT-ISVM algorithm		
	Number of support vectors	Accuracy	Training time	Number of support vectors	Accuracy	Training time
1	94	82.9%	3.96s	90	83.1%	3.93s
2	150	85.1%	6.52s	108	83.7%	4.81s
3	393	90.6%	11.67s	296	89.9%	9.54s
4	670	95.8%	37.54s	589	94.5%	16.17
5	1074	96.5%	65.87s	908	96.1%	25.91s

increasing training times. The detection rate of the SCV-KKT-ISVM algorithm is as high as 91.38, while the detection rate of the ISVM algorithm is still lower than the initial one, only 79.96. In Figure 9(b), the false alarm rates of both KKT-ISVM algorithm and ISVM algorithm increase with the increase of training times. Among them, the false positive rate of the ISVM algorithm is as high as 5.68; The false alarm rate of the SCV-KKT-ISVM algorithm gradually decreases to 2.07. It can be seen that the SCV-KKT-ISVM algorithm has stronger generalization ability and better classification performance on test data compared to the KKT-ISVM algorithm and ISVM algorithm. Table 2 indicates the specific values of abnormal detection rate and normal false alarm rate obtained after testing the detection model.

From Table 2, compared with the KKT-ISVM algorithm and ISVM algorithm, the updated SCV-KKT-ISVM algorithm has a higher abnormal detection rate and a lower normal false alarm rate, with an average abnormal detection rate of 87.76% and an average normal false alarm rate of 4.2%. The classification performance of the final detection model is superior to the KKT-ISVM algorithm and ISVM algorithm. This algorithm not only preserves valuable samples, but also has more obvious classification effect after the model is detected. The training time comparison results of the three algorithms are shown in Figure 10.

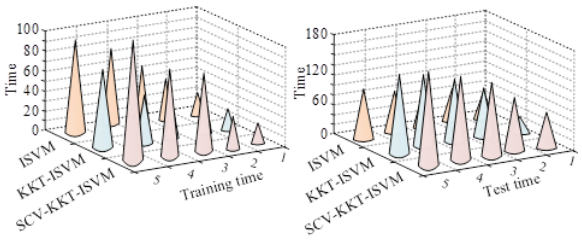


Figure 10: Comparison of training and testing time for three algorithms

In Figure 10, KKT-ISVM performs better than ISNM, but it takes a long time. Because in the process of increment, the algorithm needs to carry out cross judgment and training classification. The generalized KKT conditions in SCV-KKT-ISVM algorithm can filter new data and reduce historical data to reduce redundant data. Therefore, the algorithm does not lead to an increase in time as the number of increments increases. The introduction of SCV algorithm can eliminate the value strategy according to the change of sample space and reduce the number of training samples, thus improving the speed of model updating. Compared with ISVM algorithm and KKT-ISVM algorithm, SCV-KKT-ISVM algorithm has greater advantages, which not only enhances the classification performance, but also enhances the adaptive ability of network traffic anomaly detection model.

5 Conclusion

Currently, the number of network attacks has soared. The defense technology for network attacks can resist most of the known attack types, but the traditional defense technology update period is too long, and can not better solve the new attack types and attack traffic. In this study, a network anomaly attack detection model based on SCV optimization SVM algorithm based on incremental learning is proposed for the characteristics of network traffic anomaly attacks. The experimental results shows that compared with the SCV-KKT-ISVM algorithm and the Batch-SVM algorithm, it is found that the recall rate shows a decreasing trend with increasing training times. Under high-dimensional data, the recall rate of Batch-SVM is 0.32, higher than the 0.28 of SCV-KKT-ISVM; On the contrary, under low dimensional data, the recall rate of Batch-SVM is 0.56, lower than the 0.43 of SCV-KKT-ISVM. The initial detection model classification performance of the three algorithms is consistent, but there are differences after incremental training. The detection rates of KKT-ISVM and SCV-KKT-ISVM show an increasing trend with the number of training sessions. The detection rate of SCV-KKT-ISVM is as high as 91.38, and the detection rate of ISVM is 79.96. The false alarm rate of ISVM is as high as 5.68, while the false alarm rate of SCV-KKT-ISVM drops to 2.07. The results show that SCV-KKT-ISVM has stronger test data generalization ability, with better classification performance than KKT-ISVM and ISVM. The anomaly detection rate is 87.76%, and the false alarm rate is 4.2%. The classification effect has

Table 2: Comparison of DR and FAR of Three Algorithms

Training frequency	ISVM		KKT-ISVM		SCV-ISVM	
	DR (%)	FAR (%)	DR (%)	FAR (%)	DR (%)	FAR (%)
1	82.7	5.5	82.7	5.5	82.7	5.5
2	84.9	8.3	80.9	5.3	86.2	4.3
3	78.5	7.4	86.1	6.2	88.8	4.1
4	82.7	5.6	85.3	5.2	89.4	3.6
5	81.7	6.2	89.4	5.2	91.7	3.5

been optimized. Compared with the anomaly detection method in reference [10], the SCV-KKT-ISVM algorithm not only improves classification accuracy, but also greatly enhaces the speed of model updating and enhances the continuous learning ability of the model. However, the SCV algorithm selected in this study is only the Gaussian kernel function with the best effect in general, and whether other kernel functions are suitable for anomaly detection in network environment needs further research and discussion.

References

[1] F. Bazikar, S. Ketabchi, and H. Moosaei, "Smooth augmented Lagrangian method for twin bounded support vector machine," *Numerical Algebra, Control and Optimization*, vol. 12, no. 4, pp. 659-678, 2022.

[2] H. F. Bhat, and M. A. Wani, "Novel PSSM-based approaches for gene identification using support vector machine," *Journal of Information Technology Research*, vol. 14, no. 2, pp. 152-173, 2021.

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

[4] S.-F. Chiou, E. F. Cahyadi, C.-Y. Yang, and M.-S. Hwang, "An improved Chang-Lee's smart card-based authentication scheme," *Journal of Physics: Conference Series*, vol. 1237, pp. 042044, 2019.

[5] C. Ding, Y. Chen, Z. Liu, A. M. Alshehri, and T. Liu, "Fractal characteristics of network traffic and its correlation with network security," *Fractals: An Interdisciplinary Journal on the Complex Geometry of Nature*, vol. 30, no. 2, pp. 2-12, 2022.

[6] W. R. Ghanem, M. Shokir, and M. Dessoky, "Defense against selfish PUEA in cognitive radio networks based on hash message authentication code," *International Journal of Network Security*, vol. 4, no. 1, pp. 12-21, 2016.

[7] Negin Hamian, Majid Bayat, Mahdi R.Alaghband, Zahra Hatefi, and Seyed Morteza Pournaghi, "Blockchain-based user re-enrollment for biometric authentication systems," *International Journal of Network Security*, vol. 14, no. 1, 2022, pp. 18-38, 2022.

[8] F. X. Han, S. Y. Liu, T. Z. Zhang, X. Lu, and Y. Li, "Sparse auto-encoder combined with kernel for network attack detection," *Computer Communications*, vol. 173, no. 1, pp. 14-20, 2021.

[9] Y. Hao, J. Li, N. Wang, X. Wang and X. Gao, "Spatiotemporal consistency-enhanced network for video anomaly detection," *Pattern Recognition*, vol. 121, no. 11, pp. 108-232, 2022.

[10] I. Hidayat, M. Z. Ali, and A. Arshad, "Machine learning-based intrusion detection system: An experimental comparison," *Journal of Computational and Cognitive Engineering*, vol. 2, no. 2, pp. 88-97, 2022.

[11] M.-S. Hwang, E. F. Cahyadi, S.-F. Chiou, and C.-Y. Yang, "Reviews and analyses the privacy-protection system for multi-server," *Journal of Physics: Conference Series*, vol. 1237, pp. 022091, 2019.

[12] K. A. Kumar, and A. Anjum, "A chaos maps based method using encryption scheme for securing DICOM images: A comparative analysis," *International Journal of Network Security*, vol. 12, no. 3, pp. 128-135, 2020.

[13] D. Lakhmiri, R. Alimo, and S. Le Digabel, "Anomaly detection for data accountability of Mars telemetry data," *Expert Systems with Applications*, vol. 189, no. 5, pp. 2-7, 2022.

[14] M. Lyu, H. H. Gharakheili, C. Russell, and V. Sivaraman, "Hierarchical anomaly-based detection of distributed DNS attacks on enterprise networks," *IEEE Transactions on Network and Service Management*, vol. 18, no. 1, pp. 1031-1048, 2021.

[15] M. Marvi, A. Arfeen, and R. Uddin, "An augmented K-means clustering approach for the detection of distributed denial-of-service attacks," *International Journal of Network Management*, vol. 31, no. 6, pp. 2-23, 2021.

[16] M. M. Nabi, and F. Nabi, "Cybersecurity mechanism and user authentication security methods," *International Journal of Network Security*, vol. 14, no. 1, pp. 1-9, 2022.

[17] E. Nsugbe, "Toward a self-supervised architecture for semen quality prediction using environmental and lifestyle factors," *Artificial Intelligence and Applications*, vol. 1, no. 1, pp. 35-42, 2023.

[18] A. H. Poursaeed, and F. Namdari, "Real-time voltage stability monitoring using weighted least square support vector machine considering overcurrent protection," *International Journal of Electrical Power & Energy Systems*, vol. 136, no. 5, pp. 2-18, 2022.

[19] K. Shaheed, A. Mao, I. Qureshi, Q. Abbas, M. Kumar, and X. Zhang, "Finger-vein presentation attack detection using depthwise separable convolution neural network," *Expert Systems with Application*, vol. 198, no. 6, pp. 2-16, 2022.

[20] G. Singh, and N. Khare, "GSFI_SMOTE: A hybrid multiclass classifier for minority attack detection in internet of things network," *International Journal of Ad Hoc and Ubiquitous Computing*, vol. 38, no. 3, pp. 45-61, 2021.

[21] D. Tang, X. Wang, Y. Yan, D. Zhang, and H. Zhao, "ADMS: An online attack detection and mitigation system for LDoS attacks via SDN," *Computer Communications*, vol. 181, no. 1, pp. 454-471, 2022.

[22] A. M. Usman, and M. K. Abdullah, "An assessment of building energy consumption characteristics using analytical energy and carbon footprint assessment model," *Green and Low-Carbon Economy*, vol. 1, no. 1, pp. 28-40, 2023.

[23] X. Wang, Z. Liang, A. S. V. Koe, Q. Wu, X. Zhang, H. Li, and Q. Yang, "Secure and efficient parameters aggregation protocol for federated incremental learning and its applications," *International Journal of Intelligent Systems*, vol. 37, no. 8, pp. 4471-4487, 2021.

[24] Z. Wang, C. Liu, and F. Yan, "Condition monitoring of wind turbine based on incremental learning and multivariate state estimation technique," *Renewable Energy*, vol. 184, no. 1, pp. 343-360, 2022.

[25] M. Y. Zhu, Z. Chen, K. F. Chen, and Y. Zhang, "Attention-based federated incremental learning for traffic classification in the Internet of Things," *Computer Communications*, vol. 185, no. 5, pp. 168-175, 2022.

Biography

Lijie Li, obtained Master Degree in School of Computer Engineering and Science from Shanghai University, Shanghai, China, in 2007. Presently, he is working as a Vice Professor and the Dean in the Department of Computer Application, Ningbo City College of Vocational Technology, China. He has published more than 30 articles in journals and conferences proceedings. His areas of interest include machine learning, image processing, pattern recognition and dynamic optimization.

SDN-based Privacy Protection Model for IoT Node Awareness

Fengqing Tian[1], Haili Xue[2], Guangchun Fu[1], and Guohui Liu[3]
(Corresponding author: Fengqing Tian)
School of Information Engineering, Henan Institute of Science and Technology[1]
School of Computing, Xinxiang Vocational and Technical College[2]
Xinxiang 453000, China
Zhongheyi Testing Technology Co., Ltd, Zhengzhou 450000, China[3]
Email: tfengqing@163.com
(Received Sept. 8, 2023; Revised and Accepted June 26, 2024; First Online Aug. 17 & 22, 2024)

Abstract

The issue of privacy leakage during data transmission in software-defined networks has not been resolved. For this issue, the current study presents two software-defined network privacy protection models. The first model optimizes multi-paths and incorporates information masquerading schemes. The second model focuses on information masquerading alone. The experimental data showed that the research model was effective in optimizing the information transmission path; the average recall rate of the deep neural network classifier was 97%, the average false alarm rate was 8%, and the average missing alarm rate was 5%. Therefore, this designed model can effectively prevent attackers from accurately stealing information. The software-defined network privacy protection model for information camouflage can effectively classify data, which is conducive to the camouflage of private information.

Keywords: Information Camouflage; Information Leakage; Multi-Path Optimization; Privacy Protection; Software Defined Network

1 Introduction

Software defined networks (SDN) achieve programmable network management and flexibility by separating the network control plane and data plane. In SDN, network administrators can centrally manage the entire network through a central controller, thereby simplifying network management work [13]. During the use of SDN, there is a risk of information leakage due to insufficient authentication strength of the controller, security issues with the Open Flow protocol, vulnerabilities in network devices used by users, and hacker attacks during data analysis [24].

The leakage of personal privacy information makes it easier for hackers to attack hosts, causing network crashes. Information leakage by enterprises can lead to violations of network regulations and affect their normal operations. Therefore, it is very necessary to maintain the privacy in the process of SDN transmission. To address the issue of SDN privacy information leakage, this study constructs a Multi-path SDN Privacy Protection model (MP-SDN-PP) and an Information Camouflage SDN Privacy Protection model (IC-SDN-PP) based on multi-path optimization and information camouflage schemes.

The main contributions of this research are as follows: It solves the problem of sensitive information leakage during data transmission in SDN. Traditional SDN may disclose sensitive information, including routing policies and paths. However, this study employs randomization to assign transmission paths for sensitive information, consequently providing effective protection of sensitive information during data transmission. Deep learning network is used to disguise traffic, and multi-path-based data transmission privacy protection method is used to cut off the calculation of sensitive traffic and its transmission path by collecting traffic and other behaviors. Therefore, a reliable way is provided to maintain the transmission security of SDN, the active defense of sensitive traffic is realized, and a new way is provided to protect node privacy.

This study is divided into four sections. The first part explores the implementation of SDN by various scholars, along with suggested approaches to mitigate SDN information leakage. The second part is the construction of two models, MP-SDN-PP and IC-SDN-PP. The third part is an explanation of the performance testing of the two models constructed. The fourth part summarizes the paper and identifies the shortcomings of this study.

2 Related Works

SDN is a new type of network data architecture, which can help administrators configure network hardware directly through the controller to achieve global visibility of the network, with high flexibility and controllability. Some experts and scholars have conducted relevant research on the application of SDN. E Ahvar *et al.* found that the development of the internet has led to a sharp increase in the number of objects connected to the internet. The increase in connected devices was transforming the current internet into the future large-scale Internet of Things (IoT). 5G networks with high communication and computing capabilities would be applied in devices for data sharing and processing large-scale IoTs. SDN with emerging cloud related technologies could support the above application aspects and would be widely applied to support the development of large-scale IoT and new applications [1]. S Rawas found that with the increasing demand for cloud services, the traffic inside cloud data centers has significantly increased. Therefore, to provide high-quality services to customers, an optimal resource allocation was proposed and machine virtual model was integrated for modern large-scale cloud data centers. This model optimized the control function of the model through SDN, improved the performance of the model, and reduced the consumption of resources and communication costs [18]. I Alam *et al.* found that IoT has a wide range of applications, but the physical infrastructure of heterogeneous network systems has become more complex, requiring efficient and dynamic scheduling solutions. Based on this issue, an IoT network with functional virtualization and software definition based on SDN and network functionality was designed. This network model could be used to address the complexity issues in network architecture and provide solutions for infrastructure management, configuration, and scheduling issues in IoT [2]. Lin *et al.* found that existing cloud computing embedded vehicle networks cannot guarantee timely data processing or service access. Moreover, the existing network architecture did not support scalable network management, resulting in the inability to implement intelligent data computing strategies. On the basis of this issue, to improve the flexibility and controllability of the network, a fog-based base station was constructed and an architecture that supports SDN was proposed. This structure could solve the problem of delayed data processing and improve the scalability of network architecture [20].

While SDN provides convenience for administrators, it may also lead to the leakage of privacy information. Based on it, some scholars have proposed some privacy protection mechanisms. R Xie *et al.* proposed a Cross Path attack that can disrupt SDN control channels by utilizing shared links in control and data flow paths. Due to data traffic not entering the control channel, the attack was very covert. To address this issue, an adversarial path reconnaissance model was developed that could identify attack links. The model's ability to identify target paths was both feasible and effective, achieving an accuracy rate of 98% while also controlling costs [21]. X Wang *et al.* found that SDN not only brings convenience to people, but also increases the potential attack surface for cybercriminals. SDN had the characteristics of resource constraints and heterogeneity, and traditional network security solutions were difficult to achieve ideal results. Therefore, an ID based SDN security architecture model had been proposed, which provided endogenous trusted services for IoT on the network side by embedding unforgeable terminal identities in data streams. This service provided greater scalability and manageability for network security monitoring [19]. Xu *et al.* found that there are increasing privacy breaches in the uninstallation of Internet of Vehicles services, but there is insufficient regulation of SDN. Therefore, a secure service uninstallation method was designed that supported SDN to improve the service and edge utility of Internet of Vehicles. This method could solve the inherent uncertainty of SDN controllers on edge networks and was very effective in practical applications [22]. C Ke *et al.* found that the fog nodes of intelligent healthcare lack effective security mechanisms, and users' privacy data may be stolen by malicious users. In addition, fog computing was also subject to resource limitations and internal attacks by IoT. To address it, an intelligent medical fog node security authentication scheme based on SDN had been proposed. This scheme deployed an authentication algorithm in SDN to verify the credibility of fog nodes, while IoT only needed to send information to the SDN gateway, reducing the computational complexity of IoT. Through experiments, it had been proven that this scheme was practical in applications [9].

In summary, in the context of the big data era, SDN can provide people with convenience in their daily lives, but it also causes privacy breaches for users. Therefore, research on improving the security of SDN is of great significance.

3 Building a Privacy Protection Model Based on SDN

SDN is a new type of network data transmission system that has been widely used in the field of software development [12]. To address the issue of privacy leakage during data transmission, this chapter is divided into two sections to construct a privacy protection model. Section 1 combines random selection algorithm to construct the MP-SDN-PP model. Section 2 classifies traffic through intra domain controllers and constructs an IC-SDN-PP model.

3.1 Construction of MP-SDN-PP-based Model

The SDN intra domain transmission model mainly consists of four parts, namely the SDN controller, host con-

nected to different switches, switches that support the Open Flow protocol, and attackers [11]. The SDN controller can obtain a global network view through network links, calculate the information transmission path between the sender and receiver, and send the path information to a switch that supports the Open Flow protocol. There are various sensitive information in the transmission path within the SDN domain, including user personal information, data exchanged between users, etc. [10]. Assuming that both the SDN controller and Open Flow switch are operating normally, privacy information will not be leaked. Moreover, both the source and target nodes are operating normally and do not disclose private information to attackers. To ensure a broader range of available transmission paths during the communication process, the concept of deep traversal is introduced to explore all possible routes between the sender and receiver. Equation (1) is the set formula for paths.

$$path_i = (v_1, v_2, \cdots, v_N) \quad i = 1 \leq i \leq N \tag{1}$$

In Equation (1), N represents the node numbers, and any path does not intersect. A node matrix is defined to represent the set of paths, as shown in Equation (2) [8].

$$V[i][j] = \begin{cases} w(v_i, v_j) & \text{if } (v_i, v_j) \in E \\ 0 & \text{else if } i = j \\ \infty & \text{else} \end{cases} \tag{2}$$

In Equation (2), E represents the set of links. V is the set of switch nodes. v_i represents the node from which the sender transmits information to the receiver, and v_j is the node from which the receiver feedback information to the sender. The source node is set to v_1, the target node is v_9, and the corresponding node matrix is $V[10][10]$. Searching for paths from v_1 until all paths are found before ending. Figure 1 shows the process of path lookup.

Utilizing the techniques described above for routing may lead to the crossing of nodes, causing an overabundance of traffic at particular nodes and resulting in congestion of information transmission. Therefore, to avoid selecting paths with too many intersecting nodes, a path correlation formula is introduced for improvement. Equation (3) is a path correlation expression.

$$D_{path_i} = \sum_{i=1, j \leq K, i \neq l}^{N} \frac{1}{2}\left(\frac{J_{ij}}{N_i} + \frac{J_{ij}}{N_j}\right) \tag{3}$$

In Equation (3), J_{ij} represents the number of intersecting nodes between paths. To reduce the consumption of network resources, improve the utilization rate of network resources, and reduce path costs, the path cost formula is introduced as Equation (4).

$$c(path_i) = \sum_{i=1}^{N} c_i \tag{4}$$

In Equation (4), c_i represents the cost required to transmit a single data information. Therefore, the average link cost formula is Equation (5).

$$c_{average} = \frac{\sum_{i=1}^{N} c(path_i)}{n} \tag{5}$$

In Equation (5), n represents the number of links, and $\sum_{i=1}^{N} c(path_i)$ is the sum of all individual costs. In the node matrix, the weight of a path is Equation (6).

$$w(path) = \sum_{i=2}^{N} w(v_i, v_{i+1}) \tag{6}$$

In Equation (6), (v_i, v_{i+1}) represents the path weight and w is the link delay. The link delay formula is Equation (7).

$$L(v_i, v_j) = T_{total} - \frac{T_{v_i}}{2} - \frac{T_{v_j}}{2} \tag{7}$$

In Equation (7), T_{v_i} represents the time it takes for the controller to respond and send a message to the controller. T_{v_j} is the time it takes for the second switch to respond, and T_{total} represents the total time consumed. In selecting a path, it is first necessary to ensure path correlation. Each path's correlation must be such that the intersection of paths is below the path intersection threshold, which is $d(path_i) \leq \max D$. Secondly, all paths are sorted based on the link delay value. Finally, the path cost is sorted from small to large and the path with the lowest path cost is selected. After selecting the path that meets the constraint conditions, a trusted path can be obtained [16]. Implementing a random selection algorithm on the SDN controller can prevent attackers from accurately attacking the transmission path. The random selection algorithm can select a path by generating a random number, and the max value formula of the random number is Equation (8).

$$W = \sum_{k=1}^{K} w(path_j) \tag{8}$$

When the transmission path is attacked, the attack probability is related to the path intersection threshold $\max D$. Assuming there are K trusted paths, the probability of a random path being attacked is Equation (9).

$$p(path_i) = 1 - (1 - p_b)^S \left(\frac{1 - p_a}{1 - p_b}\right)^Q \tag{9}$$

In Equation (9), p_a is the probability of intersecting nodes being attacked, and p_b represents the probability of non intersecting nodes being attacked. S is the number of intermediate nodes, and Q represents the number of intersecting nodes. Therefore, the probability of a path being attacked is influenced by the number of intersecting nodes, and the average probability of a path being attacked is calculated as Equation (10).

$$P_{average} = \frac{\sum_{k=1}^{K} w(path_k)}{K} \tag{10}$$

In Equation (10), K is the number of trusted paths and $\sum_{k=1}^{K} w(path_k)$ is the probability of any trusted path being attacked. From this, the smaller the average probability of a path being attacked, the safer the path for

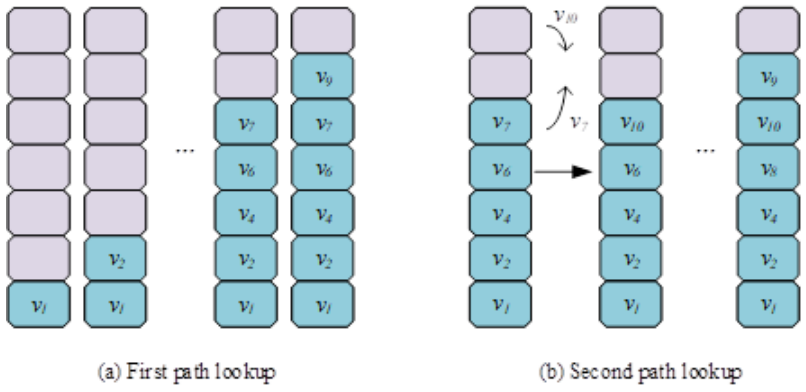


Figure 1: Path lookup process diagram

transmitting information, and vice versa, the more dangerous it is. The larger the threshold max D of path intersection and the higher the value K of the number of trusted paths, the more intersecting nodes there will be. The smaller the threshold max D of path intersection, the smaller the value K of the number of trusted paths, and the lower the path correlation, resulting in a decrease in model performance. To solve this problem, an max D and $P_{average}$ trade-off model is constructed, and the model flowchart is shown in Figure 2.

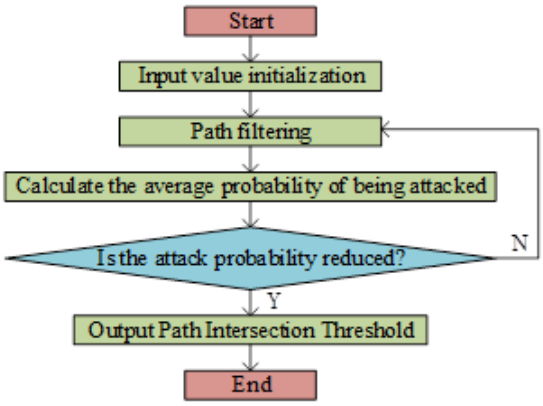


Figure 2: Model flowchart

In Figure 2, the first is to input the number of paths, the probability p_a of intersecting nodes being attacked, and the probability p_b of non intersecting nodes being attacked, and initialize the path intersection threshold max D to 0. Then path filtering is performed based on the max D value and the $P_{average}$ value is calculated. To determine whether the $P_{average}$ value has decreased. If it has decreased, increase max D; if it has not decreased, output the max D value. Figure 3 is the structural diagram of MP-SDN-PP.

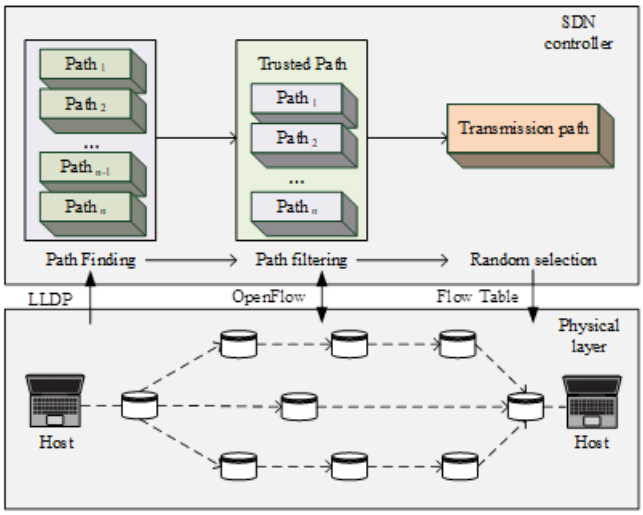


Figure 3: Structure diagram of multi-path privacy protection model

3.2 Construction of IC-SDN-PP

During the process of information transmission between SDN domains, attackers can capture and intercept data packets in the transmission path. Privacy information can be disguised to become ordinary information to prevent attackers from intercepting it [7]. Firstly, the information transmitted in the transmission path is classified, and deep learning algorithms are used to classify private information and determine its category. Then the Variational Auto-encoder (VAE) model is introduced. This model is an unsupervised complex probability distribution function learning model that can make the features of private information follow the probability distribution of ordinary information features. To ensure that the model follows a Gaussian distribution, an end-to-end deep learning model is constructed. The idea of supervised learning is introduced into the VAE model to build a Supervised Variational Auto-encoder (Supervised-VAE) model. The

schematic diagram of the model is expressed in Figure 4.

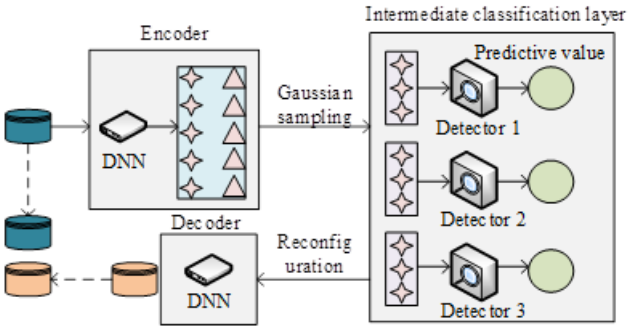


Figure 4: Structure diagram of the Supervised-VAE model

In Figure 4, the Supervised-VAE model consists of three parts: encoder, decoder, and intermediate classification layer. The encoder can compress low-dimensional information into high-dimensional space through the DNN structure, making the information follow a Gaussian distribution [15]. The experiment uses reparameterization techniques to sample from the Gaussian distribution and obtain intermediate classification layer features [14]. The calculation of the feature formula for the intermediate classification layer is Equation (11).

$$z = \mu + \sigma \times \epsilon \tag{11}$$

In Equation (11), μ and σ represent the mean and variance of the original data compressed into a Gaussian space, while ϵ is the values that follow a Gaussian distribution. When the number of information categories is determined, the number of binary classifiers in Supervised-VAE is also determined. To avoid errors in the feature representation of the middle layer and ensure that each middle layer feature has information expression ability, an attention layer structure is added to the model [23]. Before the feature values of the middle layer are input into the detector, a fully connected layer will also pass through. The information category classifier is displayed in Figure 5.

In Figure 5, the activation function of the full connection layer uses the Softmax function. When the eigenvalue is activated through the activation function, the weighted intermediate eigenvalue is obtained. The calculation formula is Equation (12).

$$\begin{cases} a_i = soft\max(z) \\ z_i' = a_i \times z \end{cases} \tag{12}$$

In Equation (12), a_i represents the weight vector that is the same as the feature value z of the middle layer, and the sum is 1. z_i' represents a specific feature of Class i information category. The decoder in the model can reconstruct input features, thus disguising private information as ordinary information. To accelerate the entire training

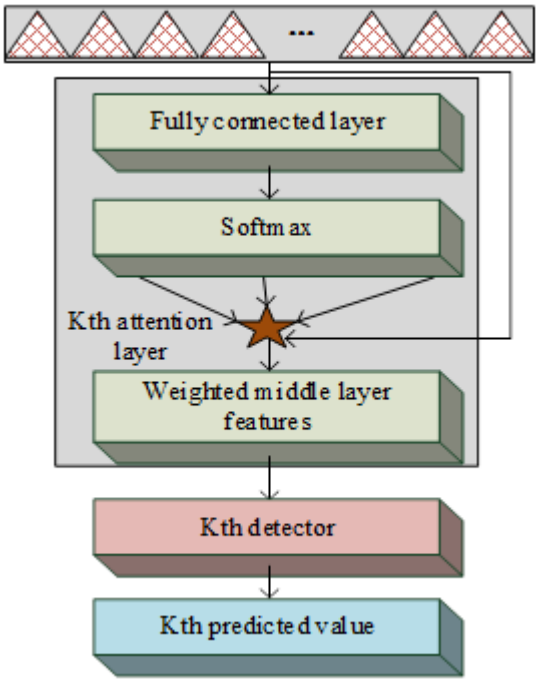


Figure 5: Structure diagram of information category classifier

process, the decoder is structured to be completely symmetrical with the encoder. This implies that the number of hidden layer units in the encoder and decoder are identical and correspond to each other [17]. The divergence calculation formula of the loss function in the unchanged VAE model is Equation (13).

$$\begin{cases} L_{KL} = -KL(q(z,X)||p(z,X)) \\ L_{recons} = MSE(X||q(z,X)) \\ L_{class} = \sum_{i=1}^{N} y_i \log \hat{y}_i + (1-y_i)(1-\log \hat{y}_i) \end{cases} \tag{13}$$

In Equation (13), $p(z,X)$ is the standard normal distribution. MSE represents the mean square error. y_i is the true value of the information category classifier. $\hat{y}_i$ is the predictive value of the information category classifier, and N' represents the number of samples in the training set. The final loss function calculation formula of the Supervised-VAE model is obtained by integrating Equation (12), as listed in Equation (14).

$$L = L_{KL} + L_{recons} + L_{class} \tag{14}$$

In Equation (14), L_{recons} is the reconstruction error value of the loss function, L_{class} represents the training value of the loss function, and L_{KL} is the divergence of the loss function. After the above improvements to the VAE model, the entire model can be trained to randomly sample information in Gaussian space. Privacy information only needs to meet the standard normal distribution of ordinary information to achieve the purpose of camouflaging private information into ordinary information [3, 6]. The calculation formula for the accuracy, recall rate, false

alarm rate, and missed alarm rate of disguised information classification through a classifier is Equation (15).

$$\begin{cases} Accuracy = \frac{TP}{TP+FN} \\ Recall = \frac{TP}{TP+FP} \\ Falsealarm = \frac{FP}{TN+FP} \\ Mis\sin galarm = \frac{FN}{TP+FN} \end{cases} \quad (15)$$

In Equation (15), TP is the number of correctly classified private information, FN represents the number of misclassified private information. FP is the number of misclassified other information, and TN represents the number of correctly classified other information. During the transmission of privacy information between SDN domains, privacy information can be transmitted through multiple paths. These paths can not only transmit private information, but also ordinary information, creating a good environment for hiding private information.

4 Simulation Experiments and Performance Analysis

In response to the privacy leakage issue in SDN data transmission, multi-path and information camouflage privacy protection models have been constructed. This chapter will conduct performance tests on both models.

4.1 Performance Testing of Multi-path Privacy Protection Model

This experiment uses RYU as the controller of SDN, with Mininet as the operating environment, TCP as the controller protocol, OpenvSwitch as the virtual switch, 64 bit Ubuntu 14.04 as the virtual machine parameter, and 4GB of running memory. The multi-path privacy protection model with the traditional MPA model is tested to analyze their performance in path lookup. The test results are exhibited in Figure 6.

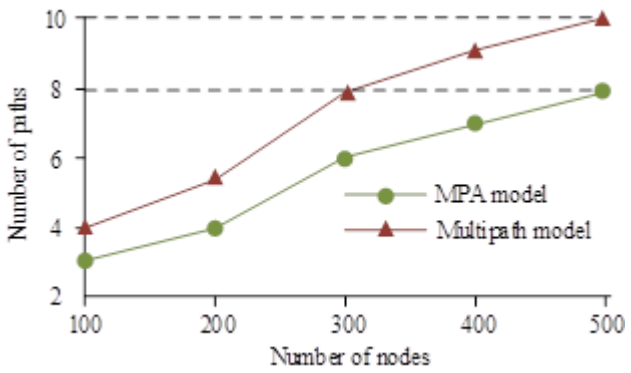


Figure 6: Comparison Curve of Path Search Results

From Figure 6, the MPA and multi-path model find 2, 4, 6, 7, 8 paths and 4, 5, 8, 9, 10 paths when the nodes are 100, 200, 300, 400, and 500, respectively. So when the number of nodes is 500, the two models have the strongest path finding ability, and the multi-path model can find more paths. The number of nodes is set to 500, selecting 4, 5, and 6 paths, and testing the network latency of the two models. The test results are appeared in Figure 7.

Figure 7 shows the time consumption curves of MPA and multi-path models under different path tests. As the traffic load increases, the average network latency under different paths shows an increasing trend. In Figure 7, when K=6 and the traffic load is 15Mbit/s, the time consumed is the least, and the average network delay is the least: MPA is 2s, and the multi-path model is 1.7s. From this, it can be concluded that K=6 is the optimal path transmission state, and the threshold at this time is also the best. Moreover, the testing time of the multi-path model is shorter than that of MPA, resulting in a faster testing speed. A single path scenario is selected to compare with a multi-path scenario to analyze the similarity of single path and multi-path transmission traffic. The test result curve is listed in Figure 8.

From Figure 8, the similarity of information transmitted in a single path scenario is significantly higher than that in a multi-path scenario. The multi-path model employs varied random paths for transmitting information, thereby reducing the similarity of information and impeding attackers from accessing private information through information interception. This enforces some disruption to the attacker's attack behavior. Therefore, the multi-path model can improve the security of information transmission and prevent the leakage of private information. The rise in the number of links will result in higher transmission expenses. Hence, this study opts for attribute encryption models that rely on ABE and click on protocol. It also compares the operational expenses of the multi-path model with these two conventional privacy protection models. Figure 9 shows the cost comparison curve.

From Figure 9, the ABE has the highest cost consumption, with an average consumption of 225. The average consumption of the click protocol model is 187. The multi-path model has the lowest cost consumption, with an average of 67. The multi-path model can achieve the minimum cost consumption while ensuring privacy protection.

4.2 Performance Testing of Information Camouflage Privacy Protection Model

This experiment uses Ryu controller, operating environment is Mininet, inter domain transmission protocol is BGP, development language is Python 3.6.1, Openflow 1.3 protocol and Tensorflow are selected as deep learning frameworks. Different classifiers are selected to simulate the attacker's attack behavior in the experiment, and DNN is compared with C4.5, SVM, and NB classifiers for the experiment. Figure 10 shows the test results.

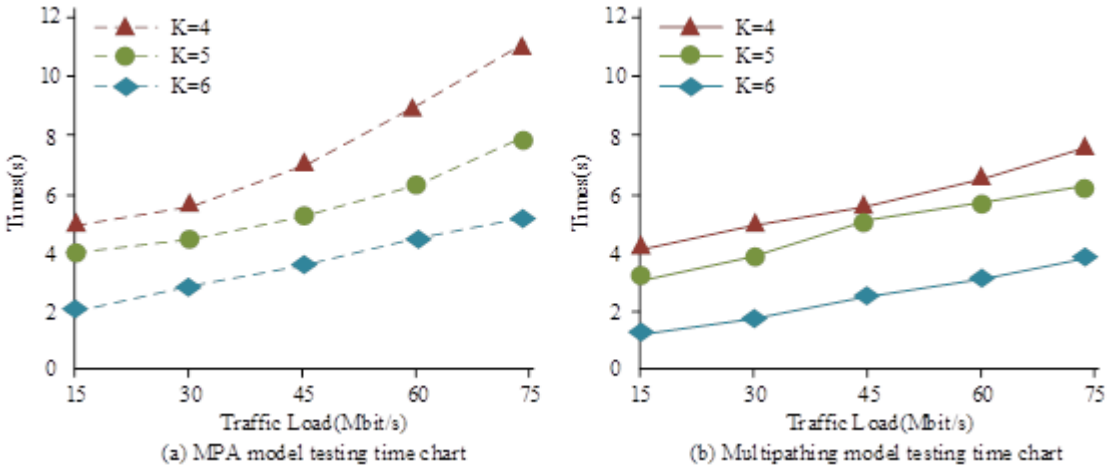


Figure 7: Comparison of Test Time under Different Paths

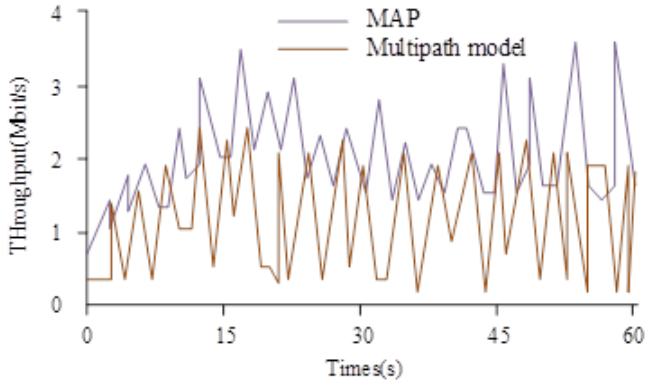


Figure 8: Similarity comparison curve

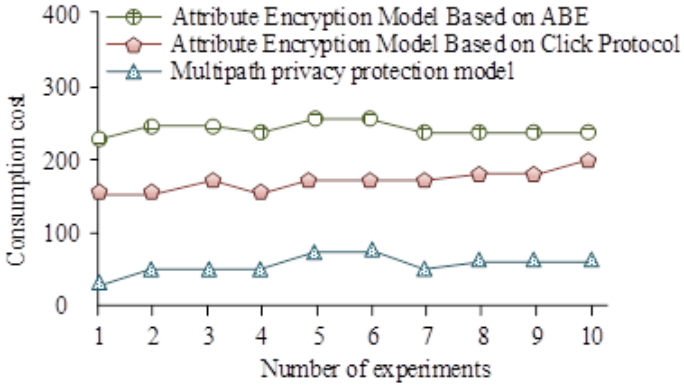


Figure 9: Calculation Cost Comparison Curve

Figure 10(a) shows the accuracy testing of four classifiers. The average accuracy values of DNN, C4.5, SVM, and NB are 95%, 90%, 92%, and 87%, respectively. The DNN classifier used in the information camouflage model has the highest accuracy in testing and has good performance in classifying ordinary and private information. Figure 10(b) shows the recall rate test. The average recall rates for DNN, C4.5, SVM, and NB are 97%, 95%, 94%, and 92%, respectively. DNN has the highest recall rate, indicating a high accuracy in predicting privacy information. Figure 10(c) shows the false alarm rate test. The DNN classifier has the lowest false alarm rate, with an average of 8%, indicating that it has a high accuracy in classifying ordinary information. Figure 10(d) shows the missed alarm rate test. The average value of DNN testing is 5%, indicating its high accuracy in classifying private information. There are two schemes for processing transmitted information: packet filling and information deformation. The information camouflage model is compared and tested with these two schemes. The test results are displayed in Table 1.

From Table 1, under the DNN classifier, the testing accuracy of packet filling, traffic deformation scheme, and information camouflage model are 80.15%, 75.88%, and 59.40%, respectively. The test results are all greater than those tested under the other three classifiers. This indicates that different information processing methods have an impact on the performance of classifiers, but DNN has significantly better testing accuracy compared to other schemes. Under multi-path conditions, the Moore dataset is selected to test the Supervised-VAE model. Figure 11 is the distribution map of information classification features.

From Figure 11, in the early stages of training, the data in the dataset presents a chaotic distribution feature. As the training time increases, data of different colors are gathered together. The Supervised-VAE model can effectively classify data and completely distinguish between ordinary information and private information,

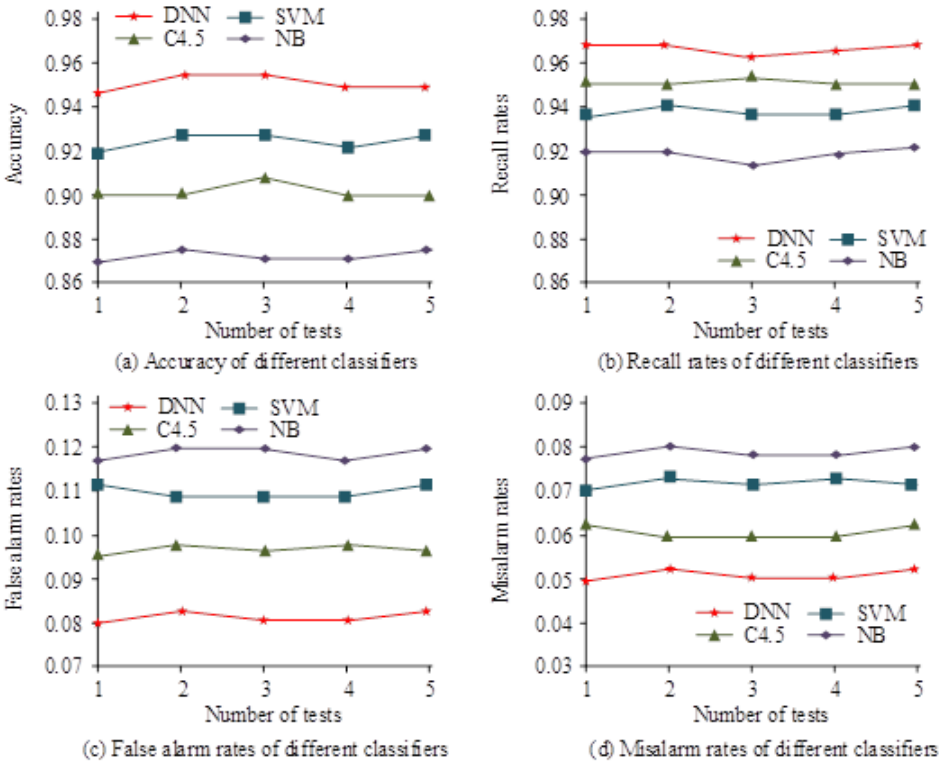


Figure 10: Performance testing of different classifiers

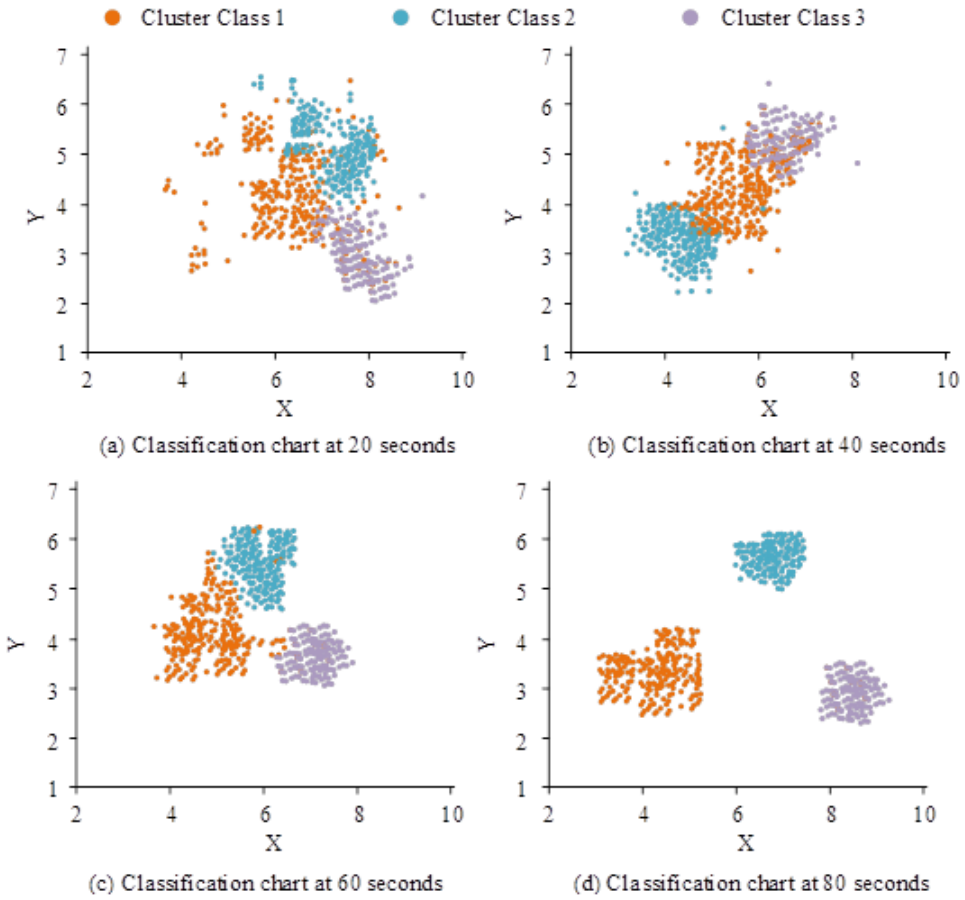


Figure 11: Classification Test Diagram

Table 1: Results of Different Information Processing Methods

Classifier	DNN	C4.5	SVM	NB
Original information	92.75%	90.67%	87.55%	85.33%
Packet padding	80.15%	75.05%	73.69%	70.25%
Flow deformation	75.88%	70.69%	67.82%	65.32%
Steganography	59.40%	57.36%	55.99%	50.87%

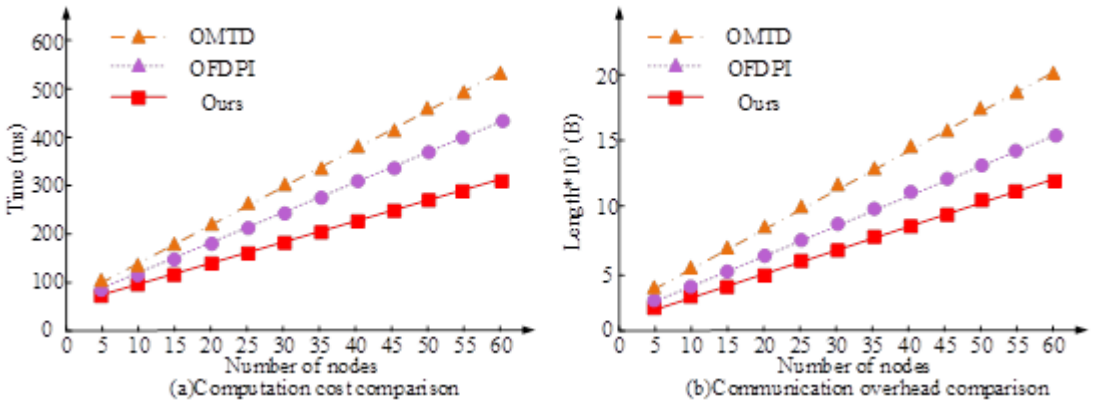


Figure 12: Performance comparison of each model

which is conducive to the disguise of private information. This study further introduces the Optimized Moving Target Defense (OMTD) model proposed by S Chiba *et al.* to compare with the Openflow Deep Packet Inspection (OFDPI) model proposed by Q Cheng *et al.* [4, 5]. The experimental results are shown in Figure 12.

Figure 12(a) shows the comparison of the computing costs of each model. When the number of nodes is small, the gap between the computing costs of the models is small, but with the increase of the number of nodes, the computing costs of the models are increased to a certain extent. When the total number of nodes is 60, the calculation time of the proposed model is 242ms, which is 230ms and 267ms lower than that of the OMTD model and OFDPI model, respectively, with an average decrease of 53.88%. Figure 12(b) shows the comparison of the communication cost of each model. When the total number of nodes is 60, the communication cost of the proposed model is 10,000B, which is 4,329B and 9,899B lower than that of the OMTD model and OFDPI model, respectively. In conclusion, the privacy protection model proposed in this study has the best performance.

5 Conclusion

SDN resolves traffic congestion issues in current networks, offering convenient user experiences in day-to-day life and improved network optimization modes for developers. This study focused on the privacy leakage problem encountered by SDN during information transmission, and constructed MP-SDN-PP and IC-SDN-PP based on multi-path optimization and information camouflage schemes. The data showed that the multi-path model could find up to 10 paths when the number of nodes is 500. When K=6 and the flow load was 15Mbit/s, the time consumed was the least, at 1.7s. And this model had the lowest cost consumption, with an average consumption of 67. The multi-path model had good path optimization effect in information transmission and could effectively prevent attackers from accurately stealing information. The DNN classifier in the information camouflage model had an average accuracy of 95%, a recall rate of 97%, a false alarm rate of 8%, and a missed alarm rate of 5%. When using different information processing schemes, the accuracy of packet filling and traffic deformation schemes testing was 80.15% and 75.88%, while the accuracy of information camouflage model testing was 59.40%. Under the training of the Moore dataset, IC-SDN-PP could effectively classify data, which was conducive to the disguise of privacy information. However, the hardware devices used in the paper may be inadequate to handle the storage flow in real-world application scenarios due to the significant data demand in the intra domain transmission mode and the high configuration of the controller. As a result, there is a risk of message storms occurring. Therefore, future research should prioritize optimizing the quantity of routed data and enhancing the link's maximum load and congestion. In addition, the attack probability of nodes in real scenarios is indefinite, so it is necessary to further classify nodes by machine algo-

rithm to obtain the optimal transmission path.

References

[1] E. Ahvar, S. Ahvar, S. M. Raza, J. M. S. Vilchez, M. G. Lee, "Next generation of SDN in cloud-fog for 5G and beyond-enabled applications: Opportunities and challenges," *Network*, vol. 1, no. 1, pp. 28-49, 2021.

[2] I. Alam, K. Sharif, F. Li, Z. Latif, M. M. Karim, S. Biswas, B. Nour, Y. Wang, "A survey of network virtualization techniques for internet of things using SDN and NFV," *ACM Computing Surveys*, vol. 53, no. 2, pp. 1-40, 2020.

[3] Z. Chen, "Research on internet security situation awareness prediction technology based on improved RBF neural network algorithm," *Journal of Computational and Cognitive Engineering*, vol. 1, no. 3, pp. 103-108, 2022.

[4] Q. Cheng, C. Wu, H. Zhou, D. Kong, R. Wei, "Machine learning based malicious payload identification in software-defined networking," *Journal of Network and Computer Applications*, vol. 192, no. 10, pp. 103186-103198, 2021.

[5] S. Chiba, L. Guillen, S. Izumi, T. Abe, T. Suganuma, "An SDN-based moving target defense as a countermeasure to prevent network scans," *IEICE Transactions on Communications*, vol. 105, no. 7, pp. 1400-1407, 2022.

[6] A. Cohen, H. Esfahanizadeh, B. Sousa, "Bringing network coding into SDN: Architectural study for meshed heterogeneous communications," *IEEE Communications Magazine*, vol. 59, no. 4, pp. 37-43, 2021.

[7] T. Duan, V. Dinavahi, "Fast path recovery for single link failure in sdn-enabled wide area measurement system," *IEEE Transactions on Smart Grid*, vol. 13, no. 2, pp. 1645-1653, 2021.

[8] M. S. Hwang, E. F. Cahyadi, C. Y. Yang, S. F. Chiou, "An improvement of the remote authentication scheme for anonymous users using an elliptic curve cryptosystem," in *IEEE 4th International Conference on Computer and Communications (ICCC'18)*, pp. 1872-1877, 2018.

[9] C. Ke, Z. Zhu, F. Xiao, Z. Huang, Y. Meng, "SDN-based privacy and functional authentication scheme for fog nodes of smart healthcare," *IEEE Internet of Things Journal*, vol. 9, no. 18, pp. 17989-18001, 2022.

[10] O. Lemeshko, O. Yeremenko, B. Sleiman, "Fast reroute model with realization of path and bandwidth protection scheme in SDN," *Advances in Electrical and Electronic Engineering*, vol. 18, no. 1, pp. 23-30, 2020.

[11] P. Li, A. A. Laghari, M. Rashid, J. Gao, T. R. Gadekallu, A. R. Javed, S. Yin, "A deep multimodal adversarial cycle-consistent network for smart enterprise system," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 1, pp. 693-702, 2022.

[12] D. Liu, L. Shan, L. Wang, S. Yin, H. Wang, C. Wang, "P3oi-melsh: Privacy protection target point of interest recommendation algorithm based on multi-exploring locality sensitive hashing," *Frontiers in Neurorobotics*, vol. 15, pp. 660304-660304, 2021.

[13] L. Lu, "Multi-path allocation scheduling optimization algorithm for network data traffic based on SDN architecture," *IMA Journal of Mathematical Control and Information*, vol. 37, no. 4, pp. 1237-1247, 2020.

[14] Z. Ma, B. Li, "A DDoS attack detection method based on SVM and K-nearest neighbour in SDN environment," *International Journal of Computational Science and Engineering*, vol. 23, no. 3, pp. 224-234, 2020.

[15] R. Muliono, "Web-based library information system design at SDN 056004 Basilam," *Journal of Research Computer Science*, vol. 1, no. 1, pp. 14-26, 2021.

[16] M. M. Nabi, F. Nabi, "Cybersecurity mechanism and user authentication security methods," *International Journal of Electronics and Information Engineering*, vol. 14, no. 1, pp. 1-9, 2022.

[17] R. Purnomo, W. Priatna, A. Y. P. Yusuf, "Optimization of the use of information technology in learning administration at SDN Hurip Jaya 03," *Jurnal Abdimas Umtas*, vol. 4, no. 2, pp. 925-930, 2022.

[18] S. Rawas, "Energy, network, and application-aware virtual machine placement model in SDN-enabled large scale cloud data centers," *Multimedia Tools and Applications*, vol. 80, no. 10, pp. 15541-15562, 2021.

[19] X. Wang, K. Xu, W. Chen, Q. Li, M. Shen, B. Wu, "ID-based SDN for the internet of things," *IEEE Network*, vol. 34, no. 4, pp. 76-83, 2020.

[20] X. Wang, S. Yin, H. Li, J. Wang, L. Teng, "A network intrusion detection method based on deep multi-scale convolutional neural network," *International Journal of Wireless Information Networks*, vol. 27, no. 1, pp. 503-517, 2020.

[21] R. Xie, J. Cao, Q. Li, K. Sun, G. Gu, "Disrupting the SDN control channel via shared links: Attacks and countermeasures," *IEEE/ACM Transactions on Networking*, vol. 30, no. 5, pp. 2158-2172, 2022.

[22] X. Xu, Q. Huang, H. Zhu, S. Sharma, X. Zhang, L. Qi, A. Z. M. Bhuiyan, "Secure service offloading for internet of vehicles in SDN-enabled mobile edge computing," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 6, pp. 3720-3729, 2020.

[23] Y. Yang, X. Song, "Research on face intelligent perception technology integrating deep learning under different illumination intensitiesm," *Journal of Computational and Cognitive Engineering*, vol. 1, no. 1, pp. 32-36, 2022.

[24] J. Zhang, M. Ye, Z. Guo, C. Y. Yen, H. J. Chao, "CFR-RL: Traffic engineering with reinforcement learning in SDN," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 10, pp. 2249-2259, 2020.

Biography

Fengqing Tian obtained his PhD in Mechanical Engineering (2021) from University of Shanghai for Science and Technology. Presently, he is working as associate professor in School of Information Engineering, Henan Institute of Science and Technology. His areas of interest include electronic information, smart agriculture, edge computing, and internet of things.

Haili Xue obtained her M.S. degree in Mechanical Engineering (2013) from Nanjing University of Science & Technology. Presently, she is working in School of Computing, Xinxiang Vocational and Technical College. Her areas of interests include computer science, image processing, and smart agriculture.

Guangchun Fu obtained his M.S. degree in Communication and Information System (2005) from Zhengzhou University. He is an associate professor in School of Information Engineering, Henan Institute of Science and Technology. His areas of interest include embedded system, intelligent control, and internet of things.

Guohui Liu obtained his bachelor's degree in Computer Engineering (2005) from Henan Institute of Science and Technology. Presently, he is working as a Chief Engineer in Zhongheyi Testing Technology Co., Ltd. His areas of interest include Computer science, smart city, smart agriculture, intelligent detection.

E-commerce Scheme Based on Proxy t-out-of-n Oblivious Signature

Jingyu Chen[1,2], Linming Gong[1,2], Xiangxiang Ma[1,2], and Daoshun Wang[3]
(Corresponding author: Jingyu Chen)
The Shaanxi Key Laboratory of Clothing Intelligence, School of Computer Science, Xi'an Polytechinic University[1]
Xi'an 710048, China
State and Local Joint Engineering Research Center for Advanced Networking and Intelligent Information Services[2]
School of Computer Science, Xi'an Polytechnic University, Xi'an, Shaanxi 710048, China
Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China[3]
No. 58 Shan-gu Road, Lintong, Xi 'an 710600, China
Email:18864615619@163.com
(Received July 13, 2023; Revised and Accepted Apr. 20, 2024; First Online Aug. 17, 2024)

Abstract

With the development of the internet and the demand for online shopping, intelligent online shopping systems have received widespread attention in recent years. However, existing online shopping systems' security and privacy issues have not been well addressed. This article proposes a shopping scheme based on proxy t-out-of-n oblivious signatures, which uses t-out-of-n oblivious signature technology to ensure that users select multiple messages from a predetermined set of messages, ensuring the legality and privacy of the selection while using proxy signature technology to provide the attribute of signature unforgeability. We have analyzed the correctness of the proposed scheme and demonstrated its security in terms of integrity, unforgeability, confidentiality of selected messages, selection restriction, non-repetitive, verifiability, distinguishability, non-repudiation, and prevention of abuse. Our scheme has lower computational and communication costs than other oblivious signature schemes. In addition, the privacy-preserving multi-choice shopping scheme we propose enables users to shop more securely and conveniently.

Keywords: Digital Signature; Electronic Shopping; Oblivious Transfer; Privacy Protection

1 Introduction

In recent years, online transaction applications such as online shopping and online banking have started to be heavily used by people. However, the privacy leakage problem that accompanies the participation in these online activities is quite serious. To cope with this problem, digital signature [14] and oblivious transfer [22] techniques play a key role. By using digital signature techniques, a signer can sign a message using a private key. Then, anyone can verify the correctness of this signature by using the signer's public key. Unlike traditional signatures, digital signatures cannot be forged, nor can the signer deny any signature he or she has made. By using oblivious transfer techniques, privacy such as the votes of Internet users in online elections or the user's choice of items in online shopping can be protected.

Various schemes based on oblivious transfer for digital commodity transactions have been proposed currently.The scheme proposed by Aiello, Bill *et al.* [1] can address issues such as blacklisted user exclusion, purchase restrictions on user wallet amounts, and privacy protection.It also prevents problems such as fraud and double payments. Borges *et al.* [3] proposed a conversion of the 1-out-of-n OT protocol to a POT (Priced Oblivious Transfer) protocol construct that enables payment functionality through an electronic coin system to handle virtual transactions. This scheme is unlinkable because the sender cannot determine whether two executions of the protocol are the same receiver. This scheme relies on POT generation, which is not efficient enough in scenarios where a large number of items are available. Biesmans *et al.* [2] proposed a pay-per-use and pay-per-channel CAS that protects user privacy, using a POT scheme combined with BABE(broadcast attribute-based encryption) together, allowing subscribers to purchase TV programs without disclosing to the service provider which programs are purchased, but the broadcast center is given the subscriber's private key and there is a potential risk of privacy leakage. Aditya *et al.* [12] use the UC(universal composability) framework [5] to describe an ideal POT functionality for aggregating statistics and dynamic pricing, but with a high communication complexity.

Digital signature techniques can also be well applied in online purchasing schemes, such as the OSBE(Oblivious

signature-based envelope) [7, 15], which can be used in mobile agent applications to guarantee that only the recipient holding a trusted third-party signature can open the envelope and compute the shared key. However, this scheme relies on a trusted third party and cannot determine the legitimacy of the recipient at the time of signature transfer. The scheme [24] achieves unconditional anonymity using heterogeneous ring signatures, but its security may be affected by the identity-based signature technology.Some researchers have further proposed signature schemes with joint blind and proxy signatures [4, 17, 21]. However, applying blind signatures in an electronic shopping system makes it impossible for the signer to know whether the signed message is selected from a legitimate candidate message.

To address the above issues, the present paper proposes a proxy oblivious signature scheme in which the signer cannot know which messages are selected by the signature recipient at the time of signing, but can be sure that the message selected by the signature recipient is the intended message, otherwise the signature will not be accepted by the verifier. Thus, restricted signatures prevent potentially illegal or malicious users from obtaining legitimate signatures for illegal or criminal activities. The main contributions of the present paper are as follows:

1) Our scheme combines the advantages of proxy signatures and t-out-of-n oblivious signatures, satisfying the following properties: integrity, unforgeability, confidentiality of selected messages, selective restriction, non-repeatability, verifiability, distinguishability, non-repudiation and prevention of abuse.

2) Correctness analysis and security analysis of the scheme have been performed and performance comparison shows that our scheme is more efficient.

3) Based on the signature scheme, we designed a privacy-protected multi-choice electronic online shopping system to solve the privacy problem in the electronic shopping system, guarantee the privacy of the selected message products, and provide a safer and more convenient shopping experience for users.

2 Preliminaries

In this section, we briefly introduce some technical backgrounds involved in our scheme.

2.1 Proxy Signature

To solve the problem that the signer is unable or inconvenient to sign, the concept of proxy signature was introduced by Mambo *et al.* [16]. Three entities are included in the proxy signature scheme: the original signer, the proxy signer, and the recipient signer. The original signer can authorize one or more proxy signers to sign the message on its behalf. Proxy signature schemes can be divided into two types: full proxy signatures and partial proxy signatures.

Full proxy signing means that the original signer sends its private key to the proxy signer over a secure channel so that the proxy signer can generate the same signature as the original signer. The original signer takes full responsibility for the signatures generated by the proxy signer. However, full proxy signatures do not prevent the misuse of proxy signatures and are not identifiable and non-repudiation, so they are not suitable for commercial applications.

Partial proxy signatures are further divided into two types: unprotected proxy signatures and protected proxy signatures.

For unprotected proxy signatures, both the original signer and the proxy signer can provide a valid proxy signature, but an unauthorized third party cannot generate a valid proxy signature. In this case, the proxy signature key is the delegation key.

For protected proxy signatures, only the designated proxy signer can provide a valid proxy signature. Neither the original signer nor the third party can generate a valid proxy signature. In this case, the proxy signature key consists of two parts: the delegation key and the private key of the proxy signer.

Compared with full proxy signatures, partial proxy signatures are more secure and reliable because they can limit the abuse of proxy signatures and have identifiability and non-repudiation for commercial applications.

2.2 Oblivious Signatures

In 1981, oblivious transfer [18] was introduced as a protocol where a sender and receiver can exchange messages without the sender knowing which message the receiver has selected.

In 1994, oblivious signatures were introduced by Chen [6], which allow a receiver to select a message to be signed without revealing the message to the signer. There are two types of oblivious signatures, one where the signer holds a key and another where the receiver holds the key. The goal of oblivious signatures is to ensure that the signed message is one of the pre-selected messages.

In 2008, Tso *et al.* [20] formalized the concepts of oblivious signatures and proposed 1-out-of-n oblivious signature protocol based on Schnorr signatures. They also improved the performance of the protocol. In 2019, Tso et al [19] combined two unintentional signatures into one scheme to achieve a two-in-one unintentional signature. The scheme ensures that nobody knows who the signer is and which message was signed.

In 2017, Chiou et al [10] proposed an oblivious signature combined with a proxy signature that meets seven security requirements, including integrity, unforgeability, unlinkability, non-repudiation, verifiability, distinguishability, and ambiguity. In 2018, Chiou and Chen [9]proposed a t-out-of-n oblivious signature that satisfies choice

restrictiveness and non-repeatability requirements, making it suitable for multi-choice e-voting applications.

Overall, oblivious signatures are useful in scenarios where privacy protection and fraud prevention are required. They have promising applications in areas such as e-voting and e-cash.

3 The Proposed Proxy t-out-of-n Oblivious Signature Protocol

This section completely describes the operation procedure of the proposed proxy t-out-of-n oblivious signature protocol. The protocol is based on the security requirements in **Definition 1**.

3.1 Attacker Model

The signature scheme proposed in this paper involves four entities: the original signer **A**, the proxy signer **B**, the receiver **R**, and the verifier **V**. In this signature scheme, **A** and **B** interact with each other over a secure channel to ensure the security of the message and the signature. However, any other participant (i.e., **R** or **V**) can only communicate with **B** through an insecure public channel, which provides opportunities for adversaries to intercept and attack. Thus, the adversary model [8, 11, 13, 23] assumes the following:

1) An adversary can intercept and tamper with all messages transmitted over an insecure channel, including signature requests, signature responses, and signature verification requests.
2) An adversary can forge a signature request and send it to **B** or **R** in an attempt to obtain a legitimate signature.
3) An adversary can forge a signature response and send it to **R** in an attempt to trick **R** into accepting an illegitimate signature.
4) An adversary can forge a signature verification request and send it to **V** in an attempt to trick **V** into accepting an illegitimate signature.
5) An adversary can use a replay attack to replay an already transmitted message back to the channel in an attempt to trick **B** or **R** into accepting a duplicate message.

3.2 Security Requirements

The system requirements for the proposed signature scheme are described as **Definition 1**.

Definition 1. *(System requirements for proxy t-out-of-n oblivious signature protocol). Assume that the original signer, the proxy signer, the receiver and the verifier interact in the proxy t-out-of-n oblivious signature protocols. A protocol is secure if it can meet the following conditions:*

1) *Integrity: As long as the receiver and signer can execute the protocol honestly, once the protocol is completed, the receiver has access to the signed message and can verify the integrity of the message.*
2) *Unforgeability: Even though the protocol is public, it is still difficult for an adversary to forge the signature within an acceptable time frame.*
3) *Privacy of selected messages: The signer cannot determine which messages the recipient has selected, protecting the recipient's privacy.*
4) *Choice restriction: the receiver cannot obtain a valid signature for any message other than the bar message.*
5) *Non-repetitive: During the signing process, the receiver cannot obtain multiple signatures on the same message.*
6) *Verifiability: Once a signed message is received, anyone can test the validity of the signature.*
7) *Distinguishability: From a signed message, anyone can distinguish whether the signature is a proxy signature or not.*
8) *Non-repudtionia: Once the proxy signer has signed the proxy authorization specification, the proxy authorization specification becomes valid, the original signer cannot deny the proxy signer's authorization, and the proxy signer cannot deny that he/she signed the document.*
9) *Prevention of abuse: Once the proxy signer has obtained the proxy authorization from the original signer, it is limited to using the proxy authority within the specified protocol, and should be clearly provable if the proxy authority is abused.*

3.3 Proposed Proxy t-out-of-n Oblivious Signature Protocol

The proposed proxy t-out-of-n oblivious signature protocol is based on the RSA signature scheme. The protocol includes four roles (original signer O, proxy signer P, signature receiver R and signature verifier V) and is divided into four phases (initialization, proxy authorization, signing and verification of the signature).

Initialization phase. This phase first defines the parameters, O and P to generate the public keys (e_O, N_O) and (e_P, N_P) and private keys (d_O, N_O) and (d_P, N_P) of the RSA cryptosystem , then generates the delegated authority m_{wt} to prove the signature authority of the proxy signer and let the proxy signer select the appropriate hash function $H(\cdot)$.

Proxy phase. O delegates signature authority to P, as shown in Figure 1.

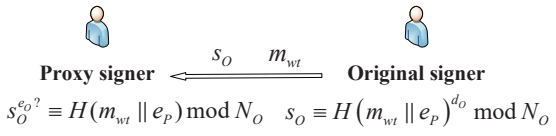


Figure 1: Proxy phase

Step 1: O computes $s_O \equiv H(m_{wt}||e_P)^{d_O} \bmod N_O$ and then sends s_O and m_{wt} to P;

Step 2: P verifies if Equation (1) holds,

$$s_O^{e_O} \stackrel{?}{\equiv} H(m_{wt} \parallel e_P) \bmod N_O. \quad (1)$$

and if it does, then P obtains signature authority from O.

Signing phase. P sends n plaintext messages $m_i (i = 1, 2, ..., n)$ and lets the receiver choose t of them, and finally the receiver receives a valid signature from P, as shown in Figure 2.

Step 1: P chooses i random variables $r_i \in {}_R Z_{N_P}^*, (i = 1, 2, ..., n)$ and computes Equation (2),

$$s_i \equiv H(m_{wt}||m_i||r_i)^{d_P} \bmod N_p. \quad (2)$$

then sends $(s_i, r_i, m_i), i = 1, 2, ..., n$, s_O and m_{wt} to the receiver R.

Step 2: R verifies whether Equation (1) and Equation (3) hold,

$$s_i^{e_P} \stackrel{?}{\equiv} H(m_{wt}||m_i||r_i) \bmod N_P. \quad (3)$$

and if they do, then R selects t from n plaintext messages and corresponds to t variables $r_{a_j}, (j = 1, 2, ..., t, t < n, 1 \le a_1, a_2, ..., a_t \le n)$. R selects a random number $b_j \in {}_R Z_{N_P}^*$, computes $c_j = {b_j}^{e_P} \cdot r_{a_j} \bmod N_P$, and sends $c_j, (j = 1, 2, ..., t)$ to P.

Step 3: P receives c_j, calculates $\beta_j = {c_j}^{d_P} \bmod N_P$, and sends $\beta_j, (j = 1, 2, ..., t)$ to R.

Step 4: R receives the latter β_j and uses the inverse of b_j to compute $v_j = b_j^{-1} \beta_j \bmod N_P$ and obtain the complete signature $\sigma(m_{a_j}) = (s_{a_j}, v_j)$, where $\sigma(\cdot)$ denotes the signature value.

Verification phase. This phase verifies that the signature received by the receiver is correct. The detailed process is shown in Figure 3.

Step 1: Receiver R sends $(\sigma(m_{a_j}), m_{a_j})$ to V.

Step 2: Verifier V verifies whether Equation (1) and Equation (4) hold,

$$s_{a_j}^{e_P} \stackrel{?}{\equiv} H(m_{wt}||m_{a_j}||{v_j}^{e_P}) \bmod N_P. \quad (4)$$

and if they do, then the signature is a valid signature.

4 The Proposed Multi-choice Online Shopping Program with Privacy Protection

4.1 System Requirements for Online Shopping Scheme

Online shopping system must have the following security features:

Verifiability: After the transaction is completed, the buyer can verify the content of the products announced by the seller to protect their rights and interests. At the same time, anyone can verify the validity of the signature.

Privacy: The buyer's identity information and the information of the selected products will not be known by the seller to ensure the buyer's privacy.

Non-repudiation: Once the proxy signer signs the explicit authorization specification, the original signer will not be able to deny the authorization to the proxy signer, and the proxy signer will not be able to deny that he or she has signed the message.

Unforgeability: The signature of the selected products can only be generated by a valid agreement and cannot be forged.

The online shopping system must meet the following system requirements:

1) The seller cannot deny a legitimate signature.
2) To protect the content of the products selected by the user.
3) Have multiple choices for multiple products.

4.2 Online Shopping Scheme with Privacy Protection

There are three entities in the solution, including Seller S, Bank B and User U. The definition of each entity is as follows:

1) Seller: Sells a variety of products, each set to $m_i (i = 1, 2, ..., n)$ and corresponding to a price of $p_i (i = 1, 2, ..., n)$.
2) Bnak: Interacts with the user as an agent of the seller and collects the user's purchase fee accordingly.
3) The user: Purchases a digital product, pays the bank, and interacts with the seller to get the selected products with the signature given by the bank.

When the user purchases the product, the seller is unaware of the user's choice and cannot deny it. The user can and will only get the products of his choice after paying for it. The system flow of the whole scheme is shown in Figure 4.

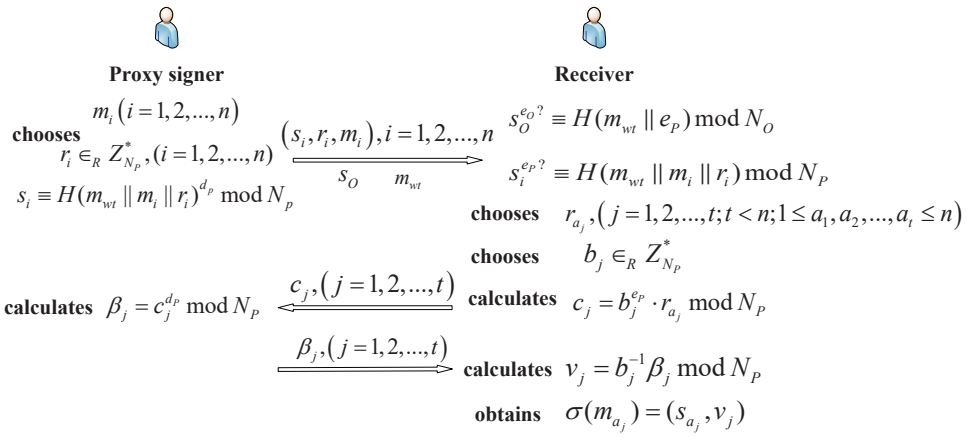


Figure 2: Signing Phase

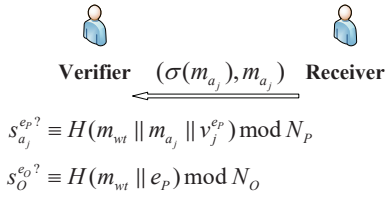


Figure 3: Verification Phase

Initialization phase. This phase first defines the parameters and the system relies on the RSA cryptosystem to generate the public and private keys of each entity , then generates the delegated authority to prove the signature authority of the proxy signer m_{wt}. And allows the proxy signer to select the appropriate hash function $H(\cdot)$.

Step 1: B computes $s_B \equiv H(m_{wt}||e_S)^{d_B} \bmod N_B$ and then sends s_B and m_{wt} to S;

Step 2: S verifies whether Equation (5) holds,

$$s_B^{e_B} \stackrel{?}{\equiv} H(m_{wt}||e_S) \bmod N_B. \quad (5)$$

and if it does, then S obtains signature authorization from B.

Selection request phase. U sends a selection request to S. S receives the user's request and requests the products index from B, then sends the products list to U.

Step 1: B calculates $h_i \equiv H(m_i||e_S)$, and then sends it to S.

Step 2: S receives it and calculates $\{p_i\} = \{H(h_i||t_S)\}$, t_S is the corresponding time. Then chooses random variable $r_i \in {}_RZ^*_{N_P}, (i = 1,2,...,n)$, calculates $s_i \equiv H(m_{wt}||m_i||p_i||r_i)^{d_S} \bmod N_S$ and sends each product introduction m_i and (m_{wt}, s_i, r_i) to user U.

Selection of products phase. Users shop for products.

Step 1: U verifies whether Equation (5) and Equation (6)

$$s_i^{e_B} \stackrel{?}{\equiv} H(m_{wt}||m_i||p_i||r_i) \bmod N_B. \quad (6)$$

are correct.

Step 2: U selects the desired item and corresponds to variable $r_{a_j}, (j = 1,2,...,t, t < n, 1 \le a_1, a_2, ..., a_t \le n)$.

Step 3: U selects a random number $b_j \in {}_RZ^*_{N_S}$, calculates $c_j = {b_j}^{e_S} \cdot r_{a_j} \bmod N_S$, and sends $c_j, (j = 1,2,...,t)$ to S.

Step 4: S calculates $\beta_j = {c_j}^{d_S} \bmod N_S$ and sends $\beta_j, (j = 1,2,...,t)$ to U.

Step 5: U receives β_j, calculates $v_j = b_j^{-1}\beta_j \bmod N_S$, and gets the seller's signature $\sigma(m_{a_j}) = (s_{a_j}, v_j)$.

Reconciliation phase. The bank verifies the correctness of the signature and performs the products count.

Step 1: U sends $(\sigma(m_{a_j}), m_{a_j})$ to B.

Step 2: B verifies whether Equation (5) and Equation (7) hold.

$$s_{a_j}^{e_S} \stackrel{?}{\equiv} H(m_{wt}||m_{a_j}||p_{a_j}||{v_j}^{e_S}) \bmod N_S. \quad (7)$$

Step 3: B finds and the corresponding item and counts the quantity.Then sends to S for pickup, and sends the products to the buyer with the signature.

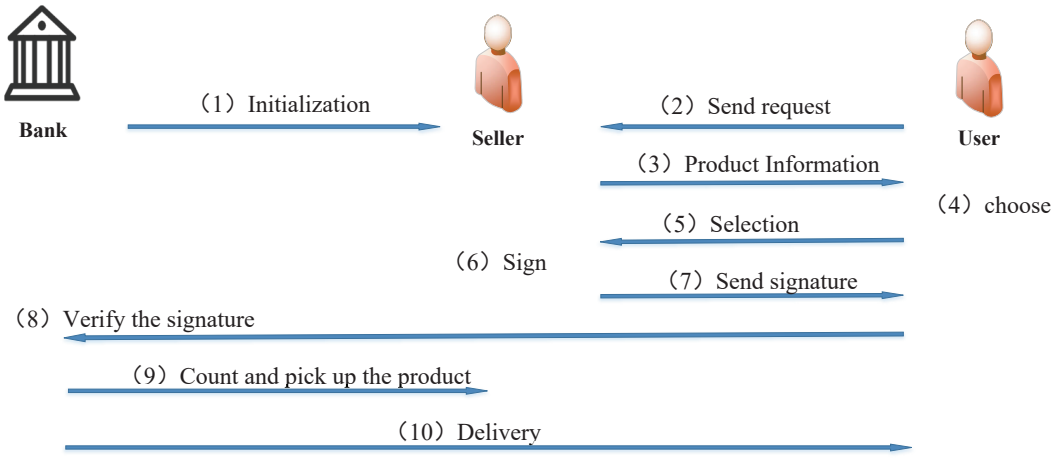


Figure 4: Online shopping system flow chart

5 Correctness Analysis and Safety Analysis

5.1 Correctness Analysis

The correctness of the proxy authorization is first proved for the proxy signature phase.The proof of Equation (1) is shown in Equation (8).

$$\begin{aligned} s_O &\equiv H(m_{wt}||e_P)^{d_O} \bmod N_O \\ {s_O}^{e_O} &\equiv H(m_{wt}||e_P)^{d_O \cdot e_O} \bmod N_O \\ &\equiv H(m_{wt}||e_P)^{1 \bmod \varphi(N_O)} \bmod N_O \\ &\equiv H(m_{wt}||e_P)^{k\varphi(N_O)+1} \bmod N_O \end{aligned} \tag{8}$$

If $H(m_{wt}||e_P)$ and N_O are coprime, the Equation (8) can be obtained from Euler's theorem, as shown in Equation (9).

$$\begin{aligned} H(m_{wt}||e_P)^{\varphi(N_O)} &\equiv 1 \bmod N_O \\ H(m_{wt}||e_P)^{k\varphi(N_O)} &\equiv 1 \bmod N_O \\ H(m_{wt}||e_P)^{k\varphi(N_O)+1} &\equiv H(m_{wt}||e_P) \bmod N_O \end{aligned} \tag{9}$$

So Equation (1) is proved.

If $H(m_{wt}||e_P)$ and N_O are not coprime, due to $N_O = p_O \cdot q_O$, so $H(m_{wt}||e_P)$ is a multiple of p_O or q_O, we may want to set Equation (10).

$$H(m_{wt}||e_P) = t \cdot p_O, (t \in Z^+). \tag{10}$$

At this time, $H(m_{wt}||e_P)$ and q_O must be coprime, otherwise $H(m_{wt}||e_P)$ is also a multiple of q_O, and thus is also a multiple of $p_O \cdot q_O$, which contradicts Equation (11).

$$H(m_{wt}||e_P) < N_O = p_O \cdot q_O. \tag{11}$$

From the fact that $H(m_{wt}||e_P)$ and q_O are coprime and from Euler's theorem, we know that Equation (12).

$$H(m_{wt}||e_P)^{\varphi(q_O)} \equiv 1 \bmod q_O. \tag{12}$$

So we can know the Equation (13),

$$\begin{aligned} H(m_{wt}||e_P)^{k\varphi(q_O)} &\equiv 1 \bmod q_O \\ [H(m_{wt}||e_P)^{k\varphi(q_O)}]^{\varphi(p_O)} &\equiv 1 \bmod q_O \\ H(m_{wt}||e_P)^{k\varphi(N_O)} &\equiv 1 \bmod q_O \end{aligned} \tag{13}$$

which satisies Equation (14),

$$H(m_{wt}||e_P)^{k\varphi(N_O)} = 1 + r \cdot q_O. \tag{14}$$

both sides of the equation (14) are multiplied by equation (10) at the same time to obtain the equation (15).

$$\begin{aligned} H(m_{wt}||e_P)^{k\varphi(N_O)+1} &= H(m_{wt}||e_P) + r \cdot t \cdot p_O \cdot q_O \\ &= H(m_{wt}||e_P) + r \cdot t \cdot \varphi(N_O) \end{aligned} \tag{15}$$

We can know the Equation (16).

$$H(m_{wt}||e_P)^{k\varphi(N_O)+1} \equiv H(m_{wt}||e_P) \bmod N_O. \tag{16}$$

So Equation (1) is proved.The same reason can prove Equation (3).

For the correctness of the signature needs to be verified Equation (4).The proof is show in the Equation (17),

$$\begin{aligned} H(m_{wt}||m_{a_j}||{v_j}^{e_P}) &= H(m_{wt}||m_{a_j}||(b_j^{-1}\beta_j)^{e_P}) \\ &= H(m_{wt}||m_{a_j}||(b_j^{-1}{c_j}^{d_P})^{e_P}) \\ &= H(m_{wt}||m_{a_j}||(b_j^{-1}(b_j^{e_P}r_{a_j})^{d_P})^{e_P}) \\ &= H(m_{wt}||m_{a_j}||(b_j^{-1}b_j r_{a_j}^{d_P})^{e_P}) \\ &= H(m_{wt}||m_{a_j}||r_{a_j}) \\ &= s_{a_j}^{e_P} \bmod N_P \end{aligned} \tag{17}$$

So the correctness of the signature protocol is proven.

Similarly the online shopping scheme based on this signature protocol is also correct.

5.2 Security Analysis of the Proposed Oblivious Signature Protocol

1) Integrity: By verifying the Equation (17), the integrity of the oblivious signature (s_{a_j}, v_j) has been proven.

2) Unforgeability: The receiver cannot forge the signature of the signer by computing β_j from the given c_j, e, and N. This can be reduced to the RSA problem, when (e', N') is the public key of RSA,there is $c' = m'^{e'} \bmod N'$. And the constraint condition is $m', c' \in Z_N^*$.Then ,it is not possible to calculate m' when given e', N' and c'.

3) Privacy of selected messages: In the proposed signature protocol, the receiver randomly selects the blinding factor b_j and computes the blinded message c_j, which is sent to the proxy signer. An attacker cannot determine b_j or r_{a_j} from the intercepted message. The signer can compute and decrypt b_j to obtain n potential b_j corresponding to n random numbers r_{a_j}, but does not know which b_j the receiver selected, so the probability of r_{a_j} being chosen by the receiver is $\frac{1}{n}$. Therefore, the proposed signature protocol can perfectly protect the privacy of the receiver.

4) Choice restriction: During the signing phase, the receiver selects t random numbers r_{a_j} and generates c_j to choose the message. If the receiver chooses a number r' that does not belong to $r_i, i = 1, 2, ..., n$, the receiver will receive $\beta_j = \left(b_j^e r'\right)^d$ from the signer and extract $v_j = \left(b_j^{-1}\beta_j\right) = (r')^d \bmod N$, but will not be able to pass the verification Equation (18).

$$s_{a_j}^e \stackrel{?}{\equiv} H(m_{a_j}||{v_j}^e). \tag{18}$$

Additionally, if the receiver attempts to obtain a signature on an illegal message , the receiver cannot choose any message except for the predetermined message, because the final signature is $\sigma(m_{a_j}) = (s_{a_j}, v_j)$, where $s_{a_j} \equiv H(m_{a_j}||r_{a_j})^d \bmod N$ is generated by the signer and bound to the message m_{a_j}.

5) Non-repetitive: If the receiver attempts to obtain multiple signatures on the same message, the receiver can randomly choose b_1 and b_2, then compute $c_1 = {b_1}^e r_{a_j} \bmod N$ and $c_2 = {b_2}^e r_{a_j} \bmod N$ during the signing phase. However, after extracting the signature parameters v_1 and v_2 and computing the Equation (19),

$$v_1 = b_1^{-1}\beta_1 = v_2 = b_2^{-1}\beta_2 = r_{a_j}^d \bmod N. \tag{19}$$

The receiver will obtain the same signature, thus preventing the receiver from obtaining multiple signatures on the same message.

6) Verifiability: In the verification phase, the verifier checks if the Equation (4) holds. Since the Equation (20) holds,

$$\begin{aligned} H(m_{wt}||m_{a_j}||v_j^{e_P}) &= H(m_{wt}||m_{a_j}||r_{a_j}) \\ &= s_{a_j}^{e_P} \bmod N_P. \end{aligned} \tag{20}$$

the signature can be verified by all verifiers.

7) Distinguishability: By using different congruences to verify the validity of the original signature and the proxy signature, anyone can easily distinguish the proxy signature from a normal signature.

8) Non-repudtionia: During signing, P and O's private keys use a hash function. No one else has access to these private keys and they cannot create a legitimate signature. Similarly, neither the proxy signer nor the original signer can deny a verifiable signature.

9) Prevention of abuse: The signature is verified and authenticated using an authorization code to record the proxy signer's ability, time, and usage conditions. The authorization code cannot be forged, so the proxy signer should not use the signature for unauthorized purposes to prevent the protocol from being abused.

5.3 Security Analysis of the Proposed Online Shopping Scheme

1) In the reconciliation phase, the bank uses the verification equation to check the contents of the products. When redeeming the products, anyone can substitute the public keys of S and B into the verification the Equation (5) and the Equation (7) to check the validity of the purchase order.

2) Each purchase step does not require the use of user identification information, so it will not reveal the user's identity. During the purchase process, the user's selection is blinded using a random number b_j, so the user's purchasing privacy is protected because the seller does not know the user's selection.

3) When the bank collects the products based on the signature, the seller cannot change the contents of the products, substitute them with inferior products, or deny their own signature. The user's selection is verified using the public keys of S and B, so the seller cannot deny the validity of the purchase signature.

4) The products selected by the user must be legally signed using the seller's private key. After the signing phase, it will be impossible to forge another valid signature for the products.

Table 1: Computation cost comparison

Schemes	Original signer	(Proxy) signer	Receiver	Verifier
[6]	–	$3nMe$	$(2n+10)Me$	$8Me$
[20]	–	$2nMe$	$(2n+2)Me$	$2Me$
[10]	$2Me$	$(n+2)Me$	$(2n+2)Me$	$2Me$
[9]	–	$(n+t)Me$	$2nMe$	$2tMe$
this work	Me	$(n+t)Me$	$(n+t)Me$	tMe

Table 2: Communication cost comparison

Schemes	$OS \to PS$	$PS \to R$	$R \to PS$	$R \to V$
[6]	–	$3nlp + nlq$	lq	$7lp + lq + lH$
[20]	–	$n(lq + lH)$	lp	$lq + lH$
[10]	$lp + lq$	$n(lq + lH)$	lp	$lq + lH$
[9]	–	$(2n+t)lN + nlm$	tlN	$t(lN + lm + lH)$
this work	$lN + lm$	$(n+t+1)lN + (n+1)lm$	tlN	$4tlN + 2tlm + tlq$

6 Comparison

The oblivious signature scheme proposed in this paper is more universal and supports selecting multiple messages based on t-out-of-n oblivious transfer scheme compared to schemes of Chen *et al.* and Tso *et al.* based on the 1-out-of-n oblivious transfer scheme. In addition, the signature scheme proposed in this paper satisfies the additional proxy signature property, which is more efficient, reliable, and convenient compared to Chiou and Chen's signature scheme based on t-out-of-n oblivious transfer. We compare the computational and communication costs of the proposed signature scheme with other related oblivious signature schemes and show the comparison results in tables. In Table 1, we compare the most important computational operation of each scheme, the modular exponentiation are denoted by the symbol Me, and the results show that the proposed scheme outperforms other protocols in terms of computational cost. In Table 2, we compare the performance of each scheme in terms of communication cost, where $q|p-1$ and (lN, lp, lq, lm, lH) represent the lengths of N, p, q, message, and hash function. The communication cost is also not higher than that of other oblivious signature schemes.

Table 3 shows that the proposed scheme in this paper provides more attributes and features than other protocols, while satisfying obliviousness, Ambiguity, multi-selectivity, and proxy functionality, indicating that our protocol has stronger functionality.

7 Conclusions

This article proposes a proxy t-out-of-n oblivious signature protocal, and based on this protocal, a new privacy-enhancing online shopping scheme is proposed, aiming to provide more comprehensive privacy protection and tighter security. Compared with other schemes, our scheme provides secure properties such as protecting buyer's choice privacy and supporting multiple product selection. In addition, our scheme is more convenient for users compared to traditional online shopping scheme. Next, we will continue to improve the protocol to enhance efficiency and security.

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] Bill Aiello, Yuval Ishai, and Omer Reingold, "Priced oblivious transfer: How to sell digital goods," in *Advances in Cryptology¡ª EUROCRYPT 2001: International Conference on the Theory and Application of Cryptographic Techniques Innsbruck, Austria, May 6–10, 2001 Proceedings 20*, pp. 119–135. Springer, 2001.

[2] Wouter Biesmans, Josep Balasch, Alfredo Rial, Bart Preneel, and Ingrid Verbauwhede, "Private mobile pay-tv from priced oblivious transfer," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 2, pp. 280–291, 2017.

Table 3: Ability comparison

Schemes	obliviousness	Ambiguity	multi-selectivity	Proxy functionality
[6]	√	√	–	–
[20]	√	√	–	–
[10]	√	√	–	√
[9]	√	√	√	–
this work	√	√	√	√

[3] Ricard Borges and Francesc Sebé, "An e-coin based construction for unlinkable priced oblivious transfer," *The Computer Journal*, p. bxad031, 2023.

[4] Xavier Bultel, Pascal Lafourcade, Charles Olivier-Anclin, and Léo Robert, "Generic construction for identity-based proxy blind signature," in *International Symposium on Foundations and Practice of Security*, pp. 34–52. Springer, 2021.

[5] Ran Canetti, "Universally composable security: A new paradigm for cryptographic protocols," in *Proceedings 42nd IEEE Symposium on Foundations of Computer Science*, pp. 136–145. IEEE, 2001.

[6] Lidong Chen, "Oblivious signatures," in *ESORICS*, pp. 161–172, 1994.

[7] Rongmao Chen, Yi Mu, Willy Susilo, Guomin Yang, Fuchun Guo, and Mingwu Zhang, "One-round strong oblivious signature-based envelope," in *Information Security and Privacy: 21st Australasian Conference, ACISP 2016, Melbourne, VIC, Australia, July 4-6, 2016, Proceedings, Part II*, pp. 3–20. Springer, 2016.

[8] Shin-Yan Chiou, "Common friends discovery for multiple parties with friendship ownership and replay-attack resistance in mobile social networks," *Wireless Networks*, vol. 24, no. 4, pp. 1055–1069, 2018.

[9] Shin-Yan Chiou and Jiun-Ming Chen, "Design and implementation of a multiple-choice e-voting scheme on mobile system using novel t-out-of-n oblivious signature.," *Journal of Information Science & Engineering*, vol. 34, no. 1, 2018.

[10] Shin-Yan Chiou, Tsung-Ju Wang, and Jiun-Ming Chen, "Design and implementation of a mobile voting system using a novel oblivious and proxy signature," *Security and Communication Networks*, vol. 2017, 2017.

[11] Shin-Yan Chiou, Zhaoqin Ying, and Junqiang Liu, "Improvement of a privacy authentication scheme based on cloud for medical environment," *Journal of medical systems*, vol. 40, pp. 1–15, 2016.

[12] Aditya Damodaran, Maria Dubovitskaya, and Alfredo Rial, "Uc priced oblivious transfer with purchase statistics and dynamic pricing," in *Progress in Cryptology–INDOCRYPT 2019: 20th International Conference on Cryptology in India, Hyderabad, India, December 15–18, 2019, Proceedings 20*, pp. 273–296. Springer, 2019.

[13] Rajesh Gupta, Sudeep Tanwar, Sudhanshu Tyagi, and Neeraj Kumar, "Machine learning models for secure data analytics: A taxonomy and threat model," *Computer Communications*, vol. 153, pp. 406–440, 2020.

[14] Ravneet Kaur and Amandeep Kaur, "Digital signature," in *2012 International Conference on Computing Sciences*, pp. 295–301. IEEE, 2012.

[15] Ninghui Li, Wenliang Du, and Dan Boneh, "Oblivious signature-based envelope," in *Proceedings of the twenty-second annual symposium on Principles of distributed computing*, pp. 182–189, 2003.

[16] Masahiro Mambo, Keisuke Usuda, and Eiji Okamoto, "Proxy signatures: Delegation of the power to sign messages," *IEICE transactions on fundamentals of electronics, communications and computer sciences*, vol. 79, no. 9, pp. 1338–1354, 1996.

[17] Xu-Feng Niu, Wen-Ping Ma, Bu-Qing Chen, Ge Liu, and Qi-Zheng Wang, "A quantum proxy blind signature scheme based on superdense coding," *International Journal of Theoretical Physics*, vol. 59, pp. 1121–1128, 2020.

[18] Michael O Rabin, "How to exchange secrets with oblivious transfer," *Cryptology ePrint Archive*, 2005.

[19] Raylin Tso, "Two-in-one oblivious signatures," *Future Generation Computer Systems*, vol. 101, pp. 467–475, 2019.

[20] Raylin Tso, Takeshi Okamoto, and Eiji Okamoto, "1-out-of-n oblivious signatures," *Lecture Notes in Computer Science*, vol. 4991, pp. 45–55, 2008.

[21] Shaobin Wang, Hong Fan, and Guohua Cui, "A proxy blind signature schemes based dlp and applying in e-voting," in *Proceedings of the 7th international conference on Electronic commerce*, pp. 641–645, 2005.

[22] Vijay Kumar Yadav, Nitish Andola, Shekhar Verma, and S Venkatesan, "A survey of oblivious transfer protocol," *ACM Computing Surveys (CSUR)*, vol. 54, no. 10s, pp. 1–37, 2022.

[23] Huijie Yang, Jian Shen, Junqing Lu, Tianqi Zhou, Xueya Xia, and Sai Ji, "A privacy-preserving data transmission scheme based on oblivious transfer and blockchain technology in the smart healthcare," *Security and Communication Networks*, vol. 2021, pp. 1–12, 2021.

[24] Caixue Zhou, "An efficient heterogeneous ring signature scheme," *International Journal of Network Security*, vol. 24, no. 5, pp. 904–912, 2022.

Biography

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.

Linming Gong. 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

Xiangxiang Ma. 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.

Daoshun Wang. 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 Cryptographic algorithms, video intelligent behavior analysis, Information hiding and digital watermarking technology.

A Blockchain Technology: Analysis of a Secure Payment Mode for Enterprise E-Commerce Import and Export Trade

Xianfeng Dong and Jing Li

(Corresponding author: Xianfeng Dong)

Henan Mechanical & Electrical Vocational College, Zhengzhou
Henan 451151, China
Email: dongxf_xf@hotmail.com

(Received July 14, 2023; Revised and Accepted July 28, 2024; First Online Aug. 17, 2024)

Abstract

This paper briefly introduces the payment mode of e-commerce import and export trade based on blockchain. In this mode, the blockchain is utilized to store transaction information, while the Bresson, Catalano, and Pointcheval (BCP) encryption algorithm is employed to encrypt the transaction amount. Additionally, the additive homomorphism of the algorithm is utilized to aid in storing and validating the blockchain. Subsequently, simulation experiments were conducted in a laboratory. The results demonstrated that this payment mode enabled the successful uploading of normal transaction information onto the blockchain while effectively safeguarding the privacy of the transaction amount. It was observed that the receiver exhibited the highest efficiency in processing transaction information, whereas the initiator exhibited the lowest efficiency. Moreover, this mode effectively resisted violent decryption attempts by third parties.

Keywords: BCP Encryption Algorithm; Blockchain; Import and Export Trade; Secure Payment

1 Introduction

With the rapid development of the Internet and economic globalization, e-commerce, particularly import and export trade, has become an essential driver for enterprise economic growth [2]. However, the development of the e-commerce import and export trade model has also brought challenges. For instance, import and export trade payment is mostly centralized within the banking system. To ensure security, banks require verification and clearing from multiple intermediary institutions [4]. Additionally, both parties involved in the transaction must possess cross-border transaction qualifications; otherwise, they must involve a qualified agent, significantly reducing transaction efficiency. Furthermore, numerous entities involved in the payment process impose fees, increasing payment costs.

The involvement of multiple payment links also leads to prolonged settlement cycles [11], during which fluctuations in national currency exchange rates further escalate payment costs. The traditional import and export trade payment mode is centralized, rendering the entire payment process vulnerable to risks if any link encounters information security issues. Blockchain, a distributed database technology, allows secure and traceable transactions among network participants without needing a centralized trust institution. Blockchain technology's decentralized, transparent, tamper-proof nature [6] makes it suitable for import and export trade payments. The identity and transaction information of the parties involved in import and export trade can be stored and verified using blockchain technology. Smart contracts within the blockchain enable quick and automated transaction processing.

Zhao *et al.* [15] proposed a novel architecture called secure publish-subscribe (SPS) without middleware, i.e., fair payment with a reputation based on blockchain. The effectiveness of SPS was verified through implementing innovative contract protocols on Ethereum. Zhang *et al.* [13] presented a blockchain-based decentralized supply chain system. The system had secure information sharing to ensure the safety of product source records without relying on any intermediary agent. Ma *et al.* [7] put forward a blockchain infrastructure service-based DRM platform with high credit and security and confirmed the reliability and security of the scheme through evaluation experiments. This paper briefly introduces the payment mode of e-commerce import and export trade based on blockchain. It utilizes blockchain to store transaction information, and the Bresson, Catalano, and Pointcheval (BCP) encryption algorithm is employed to encrypt the transaction amount. The additive homomorphism of the cryptographic algorithm aids in storing and verifying the blockchain.

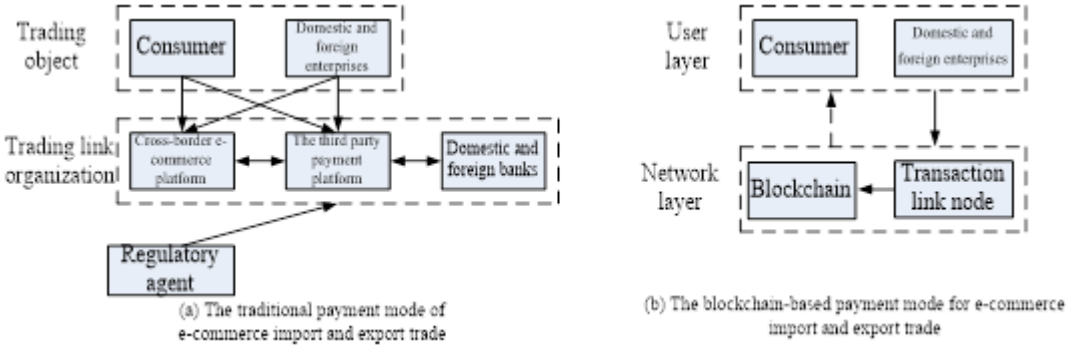


Figure 1: Basic structure of traditional and blockchain-based payment modes for e-commerce import and export trade

2 Payment Mode for E-Commerce Import and Export Trade

Figure 1 depicts the traditional and blockchain-based payment modes for e-commerce import and export trade. In the traditional e-commerce import and export trade payment mode, three central components can be identified: the transaction entities, the transaction intermediary institutions, and the regulatory bodies. The transaction entities include consumers, domestic enterprises, and foreign enterprises [14]. Enterprises can also act as consumers. The transaction intermediary institutions comprise cross-border e-commerce platforms, third-party payment platforms, and domestic and foreign banks. The regulatory bodies are responsible for overseeing the transaction intermediary institutions. In the traditional payment mode of e-commerce import and export trade, consumers and domestic and foreign enterprises first reach a transaction order on the e-commerce platform. They then utilize a third-party payment platform to complete the checkout process. During the checkout, the third-party payment platform is required to transfer different national currencies through domestic and foreign cooperative banks. Throughout the transaction process, regulatory agencies oversee the activities of the transaction intermediary institutions [1].

However, in the traditional payment mode of e-commerce import and export trade, each organization involved in the transaction operates independently and possesses complete information about consumers and enterprises. If any organization's information is compromised, it jeopardizes the overall information security of users throughout the transaction process. To address this issue, blockchain, with its decentralized and tamper-proof characteristics, can be applied to the secure payment of e-commerce import and export trade [10]. The import and export trade payment mode incorporating blockchain can be divided into the user and network layers. The user layer comprises consumers and domestic and foreign enterprises, while the network layer comprises the blockchain and the transaction link nodes. The consumers and domestic and foreign enterprises in the user layer remain consistent with those in the traditional payment mode. However, each transaction link organization in the traditional payment mode transforms into a transaction link node within the network layer, forming nodes within the blockchain. Using the blockchain-based payment mode, consumers and domestic and foreign enterprises engage in transaction-related operations within various nodes. The transaction information generated from these operations is organized into blocks and stored chronologically in the blockchain. Authorized consumers and domestic and foreign enterprises can access and query transaction information within the blockchain [5].

3 Process of the Blockchain-Based Payment Mode

Blockchain plays a crucial role in the payment mode of import and export trade by securely storing transaction information generated throughout the transaction process in an immutable and traceable manner [3]. As a result, blockchain can effectively cover the entire transaction process, from "placing an order" to "clearing the money and goods." The information generated at each stage of the transaction process can be recorded and stored in the blockchain chronologically.

There are many links in the transaction process, but the storage process of the information generated by the transaction link is generally consistent. This paper focuses on the transfer payment link in the transaction process, and its steps are as follows.

1) First is the initialization phase of the payment mode, in which the central bank (a single bank with a cross-border transfer function or a collection of multiple domestic and foreign cooperative banks) involved in the cross-border transaction generates the public parameters based on the given security parameters [8] and generates the respective public and private keys

for each node user in the blockchain through the public parameters. The formulas for generation are:

$$\begin{cases} MK = (p, q) \\ N = (2p+1)(2q+1) \\ g^{pq} \bmod N^2 = 1 + \lambda N \\ PP = (N, \lambda, g) \\ pk_i = g^{a_i} \\ sk_i = a_i \end{cases} \tag{1}$$

where MK is the master key, which consists of two numbers, p and q, λ is a security parameter, N is a number with a bit length of λ, g is a random integer that satisfies the condition within the formula, PP is a public parameter, which consists of N, λ, and g, a_i is a random integer for blockchain node user i, pk_i and sk_i are the public and private keys of blockchain node user i, respectively.

2) Then, it is the preparation stage. When the users participating in the blockchain payment mode carry out the transfer transaction, they first use their public key pk_i to encrypt the transaction amount. The encryption algorithm used in this paper is the BCP encryption algorithm [12], and its encryption formula is:

$$\begin{cases} s_i = E_{pk_i}(x_i) = (A_i, B_i) \\ A_i = g^{r_i} \bmod \ N^2 \\ B_i = g^{a_i r_i}(1 + x_i N) \bmod \ N^2 \end{cases} \tag{2}$$

where x_i is the amount within the transaction message sent by user i, s_i is the ciphertext after encrypting x_i, represented as (A_i, B_i), $E_{pk_i}(\cdot)$ is an encryption function, and r_i is a random integer. After obtaining ciphertext s_i, it is put into the block body together with the block's header hash, parent hash, timestamp, random number, and other parameters, which serve as the block's information m. Then, the SHA-512 algorithm in the hash algorithm is used to compute the summary information of m, which is also stored in the block.

3) The transaction initiator node broadcasts the block to the blockchain network, and every node in the network verifies the information of the broadcasted block. This paper uses the practical byzantine fault tolerance (PBFT) consensus algorithm [9] for consensus verification. When over two-thirds of the nodes complete the verification, consensus is achieved, and the block is appended to the main chain. Consequently, the blockchain ledger of each node is updated.

When the node verifies the block information, it first uses the SHA-512 algorithm to compute the information in the block, i.e., m, to verify whether the summary information is consistent; if it is consistent, the verification passes and continues to the next verification step. Otherwise, the transaction is terminated. Then, the BCP encryption algorithm's additive homomorphism is utilized to validate the transaction information, comparing the sum of the initiator + receiver's account balances before and after the transfer and the BCP encrypted ciphertexts. The calculation formulas are:

$$\begin{cases} E_{pk_i}(x_1 + y_1) = E_{pk_i}(x_1) \otimes E_{pk_i}(y_1) \\ E_{pk_i}(x_2 + y_2) = E_{pk_i}(x_2) \otimes E_{pk_i}(y_2) \end{cases} \tag{3}$$

where x_1 and x_2 are the account balances of the initiator before and after the transfer, y_1 and y_2 are the receiver's account balances before and after the transfer, and $\otimes$ denotes the multiplication operation in the cipher space under the same public key. The ciphertext $E_{pk_i}(y_1)$ of the account balance of the transaction receiver before the transfer is stored in the blockchain, while the ciphertext $E_{pk_i}(y_2)$ of the account balance after the transfer is encrypted by the receiver with the same public key after it receives the block information broadcast from the initiator, which is also broadcasted throughout the network. When $E_{pk_i}(x_1 + y_1) = E_{pk_i}(x_2 + y_2)$, the authentication passes.

4 Simulation Experiments

4.1 Experimental Environment

The simulation experiments were conducted in a laboratory setting. Various servers in the laboratory were utilized as user nodes responsible for transferring payments. Moreover, the blockchain network was established using virtual machines on Ethernet. These virtual machines served as some blockchain nodes. The virtual machine parameters were uniformly configured with a single-core i5 CPU, operating at a frequency of 2.5 GHz, and equipped with a memory of 4 GB. The virtual machine nodes did not perform the transfer transactions as users. Instead, they assisted in the distributed storage of transaction information. A total of six virtual machine nodes were deployed for this purpose. In the laboratory setup, the servers acted as users for the transfer transactions. Specifically, server 1 served as the central bank responsible for transferring the transaction amount, server 2 acted as the initiator of the transaction transfer, and server 3 acted as the receiver of the transaction transfer.

4.2 Experimental Setup

Server 2 was set up to have a 100 yuan balance in server 1's central bank account, and server 3 was set up to have a 0 yuan balance in server 1's central bank account.

1) Payment information uploading test
First, server 2 initiated the operation of "transferring 30 yuan" to server 3, and then two scenarios were set up:

Table 1: Transaction information query results for server 1 and virtual nodes under two payment scenarios

Scenario setting	The search result of server 1	The query result of the virtual node
Scenario 1	xx:xx:xx, server 2 transfers yuan 30 to server 3, 70 yuan balance	xx:xx:xx, server 2 transfers *** yuan to server 3, *** yuan balance
Scenario 2	No record	No record

a. The central bank transferred 30 yuan from server 2 to server 3, and then the transaction information is queried from server 1 and the virtual node respectively.

b. Server 1 was interfered so that the central bank fails to transfer 30 yuan from server 2 to server 3 but transfers it to another virtual node, and then the transaction information was queried from server 1 and the virtual node, respectively.

2) Operational efficiency of the blockchain-based payment model
50, 100, 150, 200, and 250 transactions were set, respectively. The encryption time consumed by the transaction initiator and the decryption time consumed by the transaction receiver and the central bank in the payment mode under different transaction volumes were recorded.

3) Security testing of the blockchain-based payment mode
In order to test the security of the payment mode, an additional server, server 4, was set up as a third-party attacker. The attacker comes to launch an attack to brute force the encrypted transaction information, and the time of brute force was set to 10, 20, 30, 40, and 50 min, respectively.

4.3 Experimental Results

To verify the uploading function of the proposed blockchain-based payment mode, two scenarios, namely Scenario 1 and Scenario 2, were set up. The query results of the payment operation by server 1 and virtual nodes under these scenarios are presented in Table 1. In Scenario 1, which represents a normal payment scenario, server 1 and the virtual nodes could retrieve the corresponding transaction information records. The difference was that server 1 could access complete transaction amounts while the virtual node could not. On the other hand, in Scenario 2, an abnormal payment scenario occurred after the central bank was disrupted. In this scenario, neither the server nor the virtual nodes could locate the relevant transaction records. The interference with the central bank prevented normal money transfers, leading to the block containing transaction information failing to pass verification during the uploading process. Consequently, the transaction was terminated, resulting in the absence of transaction information in the blockchain.

The operational efficiency of the blockchain-based payment mode was tested, and the results are presented in Table 2. The table demonstrated that both encryption and decryption time increased as the transaction volume increased. The average encryption efficiency of the initiator was 29 sums/s. The receiver's average decryption efficiency was 71 sums/s. The central bank's average decryption efficiency was 35 sums/s. These results showed that the receiver exhibited the highest efficiency in processing transaction information, while the initiator's processing efficiency was the lowest. This discrepancy can be attributed to the fact that during the decryption process, the receiver did not actually convert the ciphertext into plaintext. Instead, it verified the correctness of the transaction information by using the BCP encryption algorithm's additive homomorphism to perform homomorphic multiplication operations. This approach significantly reduced the time consumption involved in the process.

The extent of brute force cracking of transaction information using the third-party server is illustrated in Figure 2. The figure depicts that as the cracking time increased, the completeness of the third-party server's violent decryption of the transaction information also increased. However, it is observed that the rate of increase in decryption completeness gradually slowed down. Even after 50 minutes of violent cracking, the decryption completeness reached only 5.4%. Furthermore, with the increase of the cracking time, the decryption completeness stabilized, indicating that the security of the ciphertext can be guaranteed.

5 Conclusion

This paper briefly overviews the payment mode for e-commerce import and export trade based on blockchain technology. The approach utilizes blockchain to store transaction information and employs the BCP encryption algorithm to encrypt transaction amounts. Additionally, the additive homomorphism property of the encryption algorithm assists in storing and validating the blockchain. Subsequently, laboratory simulation experiments were conducted. In the normal payment scenario, the central bank's server and other virtual nodes queried

Table 2: Operational efficiency of the blockchain-based payment mode

Trading volume	50 sums	100 sums	150 sums	200 sums	250 sums	Average efficiency (sum/s)
The encryption time of the initiator/s	1.73	3.45	5.17	6.90	8.62	29
The decryption time of the receiver/s	0.71	1.41	2.11	2.82	3.52	71
The decryption time consumption of the central bank/s	1.43	2.86	4.28	5.71	7.14	35

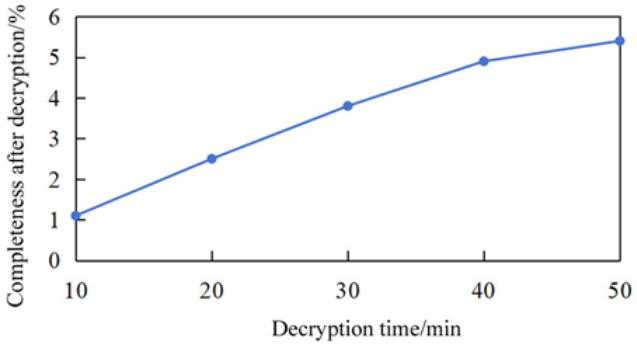


Figure 2: The cracking degree of the encrypted information using brute force cracking by the third-party server

transaction information. However, only server 1 accessed the complete transaction amount, while the virtual nodes did not obtain specific transaction amounts. In the abnormal payment scenario, the server acting as the central bank and other virtual nodes could not query transaction information. The time consumption for encryption and decryption increased with the transaction volume. The average encryption efficiency of the initiator was 29 sums/s. The average decryption efficiency of the receiver was 71 sums/s. The central bank's average decryption efficiency was 35 sums/s. The completeness of transaction information after decryption using brute force by a third-party server increased with the cracking time. Moreover, there was a gradual stabilization trend in the completeness of decryption as the cracking time increased.

References

[1] T. A. Alghamdi, I. Ali, N. Javaid, M. Shafiq, "Secure Service Provisioning Scheme for Lightweight IoT Devices With a Fair Payment System and an Incentive Mechanism Based on Blockchain," *IEEE Access*, vol. 8, pp. 1048-1061, 2019.

[2] M. Cherepniov, "Decentralized scheme for secure database creation and storage," *International Journal of Open Information Technologies*, vol. 8, pp. 109-115, 2020.

[3] P. Dangayach, "Pharmaceutical supply chain tracking system based on blockchain technology and radio frequency identification tags," *International Journal of Business Research*, vol. 19, no. 4, pp. 37-44, 2019.

[4] R. Jabbar, N. Fetais, M. Kharbeche, M. Krichen, K. Barkaoui, M. Shinoy, "Blockchain for The Internet of Vehicles: How to use Blockchain to secure Vehicle-to-Everything (V2X) Communication and Payment?," *IEEE Sensors Journal*, vol. 21, no. 14, pp. 15807-15823, 2021.

[5] M. Li, S. Shao, Q. Ye, G. Xu, G. Huang, "Blockchain-enabled logistics finance execution platform for capital-constrained E-commerce retail," *Robotics and Computer Integrated Manufacturing: An International Journal of Manufacturing and Product and Process Development*, vol. 65, pp. 1-14, 2020.

[6] C. Lin, D. He, X. Huang, M. K. Khan, K. K. R. Choo, "DCAP: A Secure and Efficient Decentralized Conditional Anonymous Payment System Based on Blockchain," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 2440-2452, 2020.

[7] Z. Ma, W. Huang, H. Gao, "Secure DRM Scheme Based on Blockchain with High Credibility," *Chinese Journal of Electronics*, vol. 27, no. 05, pp. 1025-1036, 2018.

[8] W. Shen, J. Qin, J. Yu, R. Hao, J. Hu, J. Ma, "Data Integrity Auditing without Private Key Storage for Secure Cloud Storage," *IEEE Transactions on Cloud Computing*, vol. 9, no. 4, pp. 1408-1421, 2021.

[9] F. Tian, "An agri-food supply chain traceability system for China based on RFID & blockchain technology," in *13th International Conference on Service Systems and Service Management (ICSSSM'16)*, pp. 1-6, 2016.

[10] H. Wu, N. Su, C. Ma, P. Liao, D. Li, "A privacy protection solution based on NLPCA for blockchain supply chain financial system," *International Journal of Financial Engineering*, vol. 07, no. 3, pp. 2050019, 2020.

[11] Y. Ye, Z. Ren, X. Luo, J. Zhang, W. Wu, "Garou: An Efficient and Secure Off-Blockchain Multi-Party Payment Hub," *IEEE Transactions on Network and*

Service Management, vol. 18, no. 4, pp. 4450-4461, 2021.

[12] B. Zaabar, O. Cheikhrouhou, F. Jamil, M. Ammi, M. Abid, "HealthBlock: A secure blockchain-based healthcare data management system," *Computer Networks*, vol. 200, pp. 1-16, 2021.

[13] G. Zhang, Z. Yang, W. Liu, "Blockchain-based decentralized supply chain system with secure information sharing," *Computers & Industrial Engineering*, vol. 182, pp. 1-9, 2023.

[14] M. Zhang, J. Liu, K. Feng, F. Beltran, Z. Zhang, "SmartAuction: A blockchain-based secure implementation of private data queries," *Future Generations Computer Systems*, vol. 138, pp. 198-211, 2023.

[15] Y. Zhao, Y. Li, Q. Mu, B. Yang, Y. Yu, "Secure Pub-Sub: Blockchain-Based Fair Payment With Reputation for Reliable Cyber Physical Systems," *IEEE Access*, vol. 6, pp. 12295 - 12303, 2018.

Biography

Xianfeng Dong, born in February 1974, graduated from Zhengzhou University with a bachelor's degree in December 2006. He holds a professional title of senior economist and is working at Henan Mechanical & Electrical Vocational College. He is interested in management and vocational education.

Dr. Jing Li is currently an associate professor, as well as a Dean, at School of Information Engineering, Henan Mechanical and Electrical Vocational College, Zhengzhou, China. . Her research interests include artificial intelligence and pattern recognition.

Color Image Encryption Based on Chaotic Systems and Dynamic Transformation Matrices

Chunming Xu and Yong Zhang
(Corresponding author: Chunming Xu)

School of Mathematics and Statistics, Yancheng Teachers University
No.50, Kaifang Avenue, Yancheng 224002, China
Email: ycxcm@126.com

(Received Dec. 17, 2023; Revised and Accepted June 21, 2024; First Online Aug. 17, 2024)

Abstract

Matrix transformation has been successfully used in image encryption. However, conventional image encryption algorithms often employ fixed transformation matrices, which limits the randomness and versatility of image encryption. This paper proposes an image encryption algorithm based on dynamic transformation matrices to address this issue. Random variables generated from chaotic sequences correlated with the plaintext are used to select the transformation matrices. As a result, even slight changes in the plaintext or chaotic sequences lead to completely different encryption results. Furthermore, the proposed algorithm combines dynamic transformation matrices with Zigzag transformation to further shuffle the color image's R, G, and B components, thereby enhancing the encryption effectiveness. Experimental simulations and result analyses demonstrate the algorithm's strong security and high efficiency.

Keywords: Chaotic System; Color Image Encryption; Transformation Matrix; ZigZag Transform

1 Introduction

With the development of the Internet and the explosive growth of digital data, digital images have been widely used in various fields, including communication, storage, medical imaging, military intelligence, etc. Therefore, the demand for protecting images has become increasingly urgent. These factors have prompted researchers to conduct in-depth research on image encryption to ensure the security and privacy of images and meet the security requirements of image data in the digital age [5, 12, 13, 21, 25].

Image encryption is a technique that transforms image data into an unreadable form to ensure the security and privacy of images during transmission and storage. It applies cryptography and algorithms to transform images, making it difficult for unauthorized individuals to understand or restore the original image content [2, 4, 20, 22, 23].

Traditional encryption methods such as Data Encryption Standard [3], Advanced Encryption Standard [17], Rivest Shamir Adleman [14], and Elliptic Curve Cryptography [16] are mainly suitable for encrypting text data. Image data is two-dimensional, large scale, and not free from redundancy. It features visual perception, contextual relevance, and diversity. All these make traditional encryption algorithms unsuitable for image encryption.

Chaos is a complex, unpredictable, and highly sensitive dynamic behavior. Chaos exhibits characteristics such as unpredictability, sensitivity to initial conditions, aperiodicity, fractal structure, and specific statistical properties. These characteristics make chaos an interesting and complex phenomenon and have wide research and application values in fields such as science, mathematics, and engineering. Chaos has also been widely applied in image encryption [10, 18, 26]. By combining chaotic sequences with image data, highly random encrypted images can be generated. This method exhibits strong resistance against statistical analysis and cryptographic attacks but requires high parameters and initial conditions for the decryption process.

There is a close relationship between images and matrices. In computers, images are typically represented and processed as matrices or multidimensional arrays. In recent years, researchers have proposed image encryption algorithms based on matrix transformations [1,7,9]. However, the transformation matrices used are often fixed and unchanged, which limits the randomness and usability of image encryption. To that end, this paper proposes an image encryption algorithm with dynamically changing parameters to enhance the complexity and security of the encryption process.

The rest of the paper is organized as follows. In Section 2, we review some fundamental knowledge. Section 3 introduces the the proposed image encryption scheme. Section 4 presents the experimental results and the security of the presented method. Finally, we conclude this paper in Section 5.

2 Fundamental Knowledge

2.1 Transformation Matrices

Digital image data can be represented by matrices, so matrix theory and matrix algorithms can be used to analyze and process images. Performing matrix transformation on plaintext images can change pixel values and mask plaintext information, thus encrypting the data. The transformation matrix must be reversible to restore a ciphertext image to a plaintext one. Reference [6] provided a good method to generate transformation matrices.

Assuming x is a positive integer and $a(x)_{i,j} = C_{x+j-1}^{i-1}, 1 \le i,j \le n$, we can further define the matrix A as follows:

$$A = \begin{bmatrix} a(x)_{1,1} & \dots & a(x)_{1,n} \\ \vdots & \vdots & \vdots \\ a(x)_{n,1} & \dots & a(x)_{n,n} \end{bmatrix}. \quad (1)$$

A is an integer type matrix, and its inverse matrix B is

$$B = \begin{bmatrix} b(x)_{1,1} & \dots & b(x)_{1,n} \\ \vdots & \vdots & \vdots \\ b(x)_{n,1} & \dots & b(x)_{n,n} \end{bmatrix}. \quad (2)$$

where $b(x)_{i,j} = (-1)^{i+j} \sum_{l=0}^{n-j} C_{x+l-1}^{l} C_{l+j-1}^{i-1}$.

According to the above matrix generation method, we can generate different transformation matrices for image encryption.

2.2 Zigzag Transform

Zigzag transform is a classic method for scanning matrix elements and can be used for image scrambling [24]. The matrix elements are scanned in a "Z" shape order, and then the scanned elements are sequentially stored in a vector. Finally, the vector is transformed back into a two-dimensional matrix. A specific process of the Zigzag transform is shown in Figure 1.

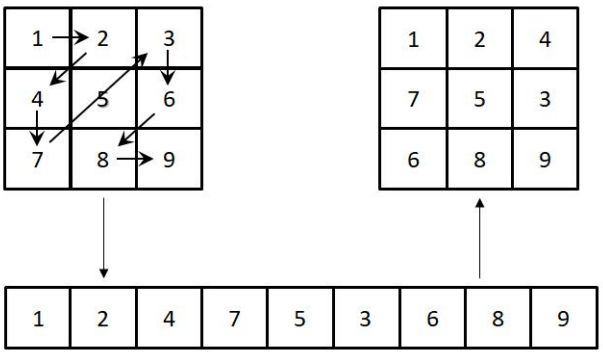


Figure 1: Zigzag Transform.

2.3 Chaotic Systems

In 2022, Sahoo presented a modified three dimensional Chen chaotic system with multi-wings attractors, which is described by the following equation [15]:

$$\begin{cases} \dot{x_1} = a(x_2 - x_1) \\ \dot{x_2} = cx_2 - x_1x_3(1 - ksin(k_1x_3)) + (c-a)x_1 \\ \dot{x_3} = -bx_3 + x_1x_2 \end{cases} \quad (3)$$

where x_1, x_2, x_3 are state variables, and a, b, c, k, k_1 are system parameters. When the system parameters are $a = 35, b = 3, c = 28, k = 0.5, k_1 = 1$, the chaotic system (3) exhibits complex chaotic behavior. In addition, it has a higher value of the largest Lyapunov exponent than the original Chen chaotic system. The state space plots for system (3) are shown in Figure 2.

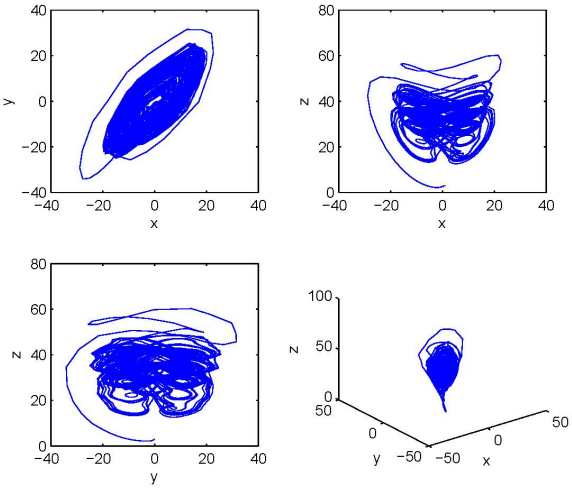


Figure 2: Typical dynamical behaviors of the chaotic system.

3 The Encryption Method

Assume that the size of the color plain image P_0 is $M \times N \times 3$, where M and N are the height and width of the image, respectively. Denote the color components of red, green and blue of P_0 as P_R, P_G and P_B, respectively. The specific steps of the proposed encryption algorithm are described as follows:

3.1 Generation of the Chaotic Encryption Sequences

Calculate The initial values x_0, y_0, z_0 of the chaotic system (3) utilizing the following equations:

$$\begin{cases} x_0 = \frac{\sum_{ij} P_{Rij}}{255MN} + 0.01 \\ y_0 = \frac{\sum_{ij} P_{Gij}}{255MN} + 0.02 \\ z_0 = \frac{\sum_{ij} P_{Bij}}{255MN} + 0.03 \end{cases} \quad (4)$$

and choose the system parameters a, b, c, k, k_1.

Do the iteration system (3) $L + 2000$ times with the parameters x_0, y_0, z_0 and then remove the former 2000 values so that three chaotic sequences x_s, y_s, z_s of length L are gotten, where $L = M \times N$. Calculate four sequences m_1, m_2, m_3 and T which will be used in the following encryption process with x_s, y_s, z_s by

$$\begin{cases} m_1 = |x_s| \times 10^{15} \mod 256 \\ m_2 = |y_s| \times 10^{15} \mod 256 \\ m_3 = |z_s| \times 10^{15} \mod 256 \\ T = |x_s| \times 10^{13} \mod 3 + 1 \end{cases} \tag{5}$$

3.2 Pixel Scrambling

The main purpose of pixel scrambling is to disrupt the spatial correlation between adjacent pixels and conceal the original visual information of an image. The steps for pixel scrambling are as follows:

1) The three components R, G, and B of a plain image are arranged in a row from top to bottom and from left to right. They are then recombined into a matrix P_1 with a size of $3 \times L$.

2) The Zigzag transformation method is applied H times to P_1 and then we can obtain the corresponding scrambled matrix P_2, where H is a positive integer.

Through pixel scrambling, not only the correlation among pixels is disrupted, but the effective fusion of the elements from the different R, G, and B components is also achieved.

3.3 Matrix Transformation

For each column of the matrix P_2, a transformation matrix is selected and applied based on the random number sequence T. Specifically, for the $i-th$ column of matrix P_2, we construct a transformation matrix A using the parameter $T(i)$ and perform a matrix transformation on it to obtain the new elements of the $i-th$ column:

$$P_3(:, i) = A(T(i))P_2(:, i) \ mod \ 256 \tag{6}$$

where $1 \leq i \leq L$.

After applying the matrix transformation to all columns, we obtain a new matrix P_3. Since the sequence T consists of three elements, there are a total of three different transformation matrices. On the other hand, the sequence T is generated from plaintext image and chaotic sequence so that it possesses a certain level of randomness. Therefore, the dynamic matrix transformation method proposed in this article can enhance the encryption effect of the matrix.

3.4 Pixel Diffusion

We perform a diffusion operation on P_3 to further alter the pixel values. The specific method is to perform XOR operations between the 1st, 2nd, and 3rd rows of P_3 with the sequences m_1, m_2, and m_3 respectively as follows:

$$\begin{cases} C_R = P_3(1, i) \oplus m_1(i) \\ C_G = P_3(2, i) \oplus m_2(i) \\ C_B = P_3(3, i) \oplus m_3(i) \end{cases} \tag{7}$$

Transform the R, G, B components C_R, C_G, C_B into matrix forms and combine them to demonstrate the encrypted color image C.

The decryption process is the inverse process of encryption and is omitted here for the sake of simplicity.

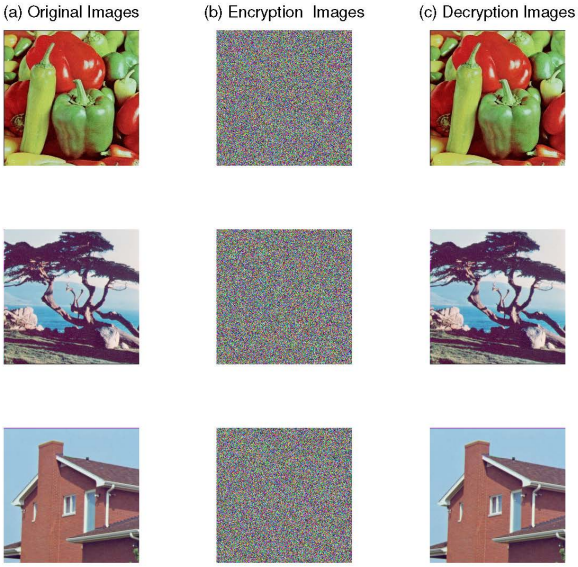


Figure 3: The experimental results of the encrypted images. (a) The original images. (b) The encryption images. (c) The decryption images.

4 Tests and Analysis of the Proposed Scheme

In the experiment, three standard color images Peppers, Pine, and House from the USC-SIPI database are selected as test images, all of which are of the same size $256 \times 256 \times 3$. The encrypted and decrypted images obtained are shown in Figure 3. From Figure 3, it can be observed that the encrypted images appear as disordered snowflake-like noise, indicating that the encrypted images perfectly hide the information of the original images. Additionally, the decrypted images appear identical to the original images, confirming the effectiveness of the decryption algorithm.

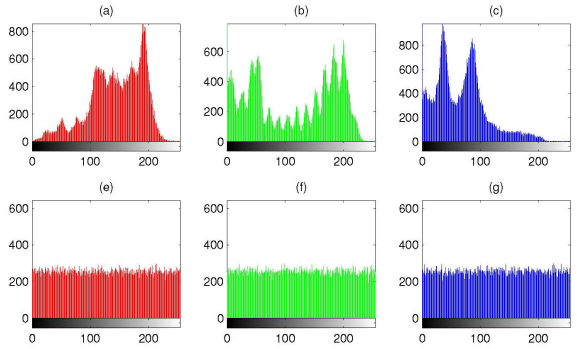


Figure 4: Histograms of Peppers in red, green, and blue. Histograms of original and corresponding encrypted images are displayed in rows 1 and 2, respectively.

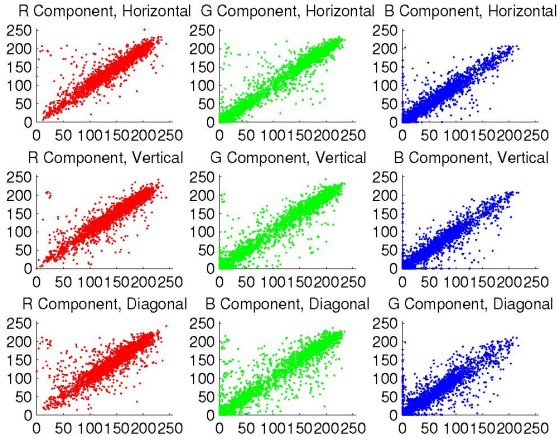


Figure 5: Correlation distributions of plaintext Lena image in each direction.

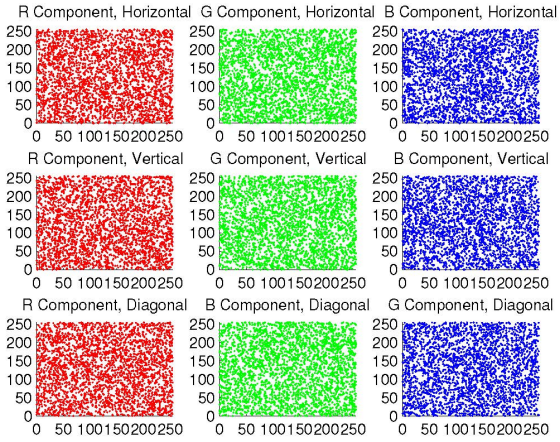


Figure 6: Correlation distributions of encrypted Lena image in each direction.

4.1 Key Space Analysis

In the algorithm used in this paper, the key space consists of five parameters $(x0, y0, z0, a, b, c, k, k1)$, which are the initial values of the chaotic system parameters and the control parameter of the chaotic system. Assuming that each parameter is represented with double precision accuracy up to 15 decimal places and considering the number of pixel scrambling H, the key space is more than 10^{120}, which is large enough to resist violent attacks.

4.2 Histogram Analysis

The histogram reflects the distribution of pixel values in digital images. A good encryption algorithm should have a uniformly distributed histogram of the ciphertext image, so that attackers cannot obtain the information about the plaint image through the histogram. Figure 4 shows the histograms of Peppers. As shown in Figure 4, the histograms of the encrypted images are evenly distributed, which can effectively mask the distribution information of the plain image, thus resisting histogram attacks.

4.3 Correlation Analysis

There is a significant correlation among adjacent pixels in a plain image. After the image is scrambled and encrypted, it is necessary to reduce the pixel correlation to the point where adjacent pixel values have no discernible pattern. The calculation formula for the correlation among adjacent pixels in an image is as follows [19]:

$$r_{xy} = \frac{\sum_{i=1}^{N}((x_i - E(x))(y_i - E(y)))}{\sqrt{(\sum_{i=1}^{N}(x_i - E(x))^2)(\sum_{i=1}^{N}(y_i - E(y))^2)}} \quad (8)$$

$$E(x) = \sum_{i=1}^{N} x_i \quad (9)$$

$$E(y) = \sum_{i=1}^{N} y_i \quad (10)$$

where x_i and y_i are gray-level values of the selected adjacent pixels, and N is the number of sample pixels.

Randomly selecting 3000 adjacent pixels from the standard test image Peppers and its encrypted image, the Equation (8) are used for correlation analysis. The selected pixels are paired, with the grayscale value of the previous point as the x-coordinate and the grayscale value of the subsequent point as the y-coordinate, resulting in Figure 5 and Figure 6. It can be observed that the original image exhibits high correlation and is mostly concentrated near a line with a slope of 1, while the encrypted image is distributed uniformly within a rectangular region. Quantitative analysis of their correlation using Equation (8) is shown in Table 1 and Table 2. From

Table 1: Correlation coefficients of the R,G and B components of the plaintext color image of Peppers.

Component	Horizontal	Vertical	Diagonal
R component	0.9319	0.9261	0.8835
G component	0.9664	0.9662	0.9427
B component	0.9298	0.9314	0.8900

Table 2: Correlation coefficients of the R,G and B components of the encrypted color image of Peppers.

Component	Horizontal	Vertical	Diagonal
R component	-0.0245	-0.0256	-0.0011
G component	-0.0037	0.0194	-0.0228
B component	0.0057	-0.0087	-0.0079

Tables 1 and 2 and Figures 5 and 6, it can be seen that adjacent pixels in the plaintext image exhibit high correlation, while the adjacent elements in the encrypted image are nearly uncorrelated, indicating that the algorithm effectively destroys the correlation of the plaintext information.

4.4 Information Entropy Analysis

Information entropy reflects the randomness and disorderliness of information. The higher the entropy of the encrypted image, the greater the randomness and the higher the security. The formula for information entropy is [8]:

$$H(m) = -\sum_{i=0}^{255} P(m_i) \log_2 P(m_i) \tag{11}$$

where m_i is the $i-th$ gray level for the digital image and $P(m_i)$ represents the probability of m_i.

For a color image, we can calculate the information entropies of the R, G and B components respectively. The information entropies of R, G and B components of the original Peppers image are 7.3920, 7.6150 and 7.1738. The information entropies of R, G and B components of the encrypted Peppers image are 7.9974, 7.9970 and 7.9971, which are all very close to the ideal value 8, indicating that the algorithm exhibits good randomness and security.

4.5 Analysis of Differential Attack Resistance

The resistance to differential attacks is an important indicator for evaluating the effectiveness of image encryption. It measures the impact of making small changes to the pixel values of a plaint image on the resulting encrypted cipher images. The greater the impact, the higher the sensitivity of the encryption algorithm to changes in the plain image, and the higher its resistance to differential attacks. Typically, two metrics, Number of Pixels Change Rate (NPCR) and Unified Average Changing Intensity (UACI), are used to assess it. The NPCR and UACI values can be calculated by [11]:

$$NPCR = \frac{\sum_{ij} D_{ij}}{W \times H} \times 100\% \tag{12}$$

$$UACI = \frac{1}{W \times H} \frac{\sum_{ij} (C_1(i,j) - C_2(i,j))}{255} \times 100\% \tag{13}$$

where $C_1(i,j)$ and $C_2(i,j)$ are the encrypted images for the plain images and D_{ij} is defined by

$$D_{ij} = \begin{cases} 0 & \text{if } C_1(i,j) = C_2(i,j) \\ 1 & \text{if } C_1(i,j) \neq C_2(i,j) \end{cases} \tag{14}$$

The theoretical value of NPCR is 99.609375%, and the theoretical value of UACI is 33.463541%.

Table 3: The NPCR and UACI values

Component	R	G	B
NPCR	99.54%	99.58%	99.22%
UACI	33.46%	33.56%	33.56%

According to Equations (12) and (13), the NPCR and UACI values of R, G and B components of the Peppers image are given in Table 3. It is evident that they are very close to the theoretical values, indicating that this algorithm has stronger resistance to differential attacks.

5 Conclusions

In this work, we proposed a novel image encryption algorithm based on chaotic systems and dynamic transformation matrices. The presented technique's security testing demonstrated its security and effectiveness. The main advantages of the paper are as follows:

1) Dynamic matrix transformation can make an encryption algorithm more random and enhance the encryption effect;

2) Both Zigzag transformation and dynamic matrix transformation can fully mix the R, G, B components of a color image;

3) Both dynamic matrix transformation and XOR diffusion operations can change the value of image pixels to hide image information.

Therefore, dynamic transformation matrices based image encryptions have high research prospects in the field of information protection and worthy of further research.

Acknowledgments

The authors gratefully acknowledge the anonymous reviewers for their valuable comments. This work is partially supported by the National Natural Science Foundation of China (No.11871417) and the Fundamental Science (Natural Science) Foundation of the Jiangsu Higher Education Institutions of China (Grant No.23KJA120004).

References

[1] M. J. Aqel, Z. ALQadi, A. A. Abdullah, "RGB color image encryption-decryption using image segmentation and matrix multiplication," *International Journal of Engineering and Technology*, vol.7, No. 3, pp.104-107, 2018.

[2] J. Blackledge, W. Govere and D. Sibanda, "Phase-Only Digital Encryption," IAENG International Journal of Applied Mathematics, vol.49, No. 1, pp.3131-3136, 2023.

[3] W. Diffie, M. E. Hellman, "Exhaustive Cryptanalysis of the NBS Data Encryption Standard"," *IEEE Computer*, vol.10, no.6, pp.74-84, 1977.

[4] C. Dong, "Dynamic Analysis of a Novel 3D Chaotic System with Hidden and Coexisting Attractors: Offset Boosting, Synchronization, and Circuit Realization," *Fractal and Fractional*, vol.6, 547, 2022.

[5] H. Dong, E. Bai, X. Q. Jiang, "Color image compression-encryption using fractional-order hyperchaotic system and DNA coding," *IEEE Access*, vol.8, pp. 163524-163540, 2020.

[6] Y.S. Dong, "Proof of a method for finding the inverse of integer matrices," *Journal of Changchun Normal University*, vol.4, no.2, pp.4-5, 2007.

[7] S. Han, S. Yang, "An asymmetric image encryption based on matrix transformation," *IEEE International Symposium on Communications and Information Technologies*, Sapporo, Japan, October 26, 2006.

[8] K.M. Hosny, S.T. Kamal, M.M. Darwish, "Novel encryption for color images using fractional-order hyperchaotic system," *Journal of Ambient Intelligence and Humanized Computing*, vol.13, no.2, pp. 973-988, 2022.

[9] S. Kanwal, S. Inam, "Mohamed Tahar Ben Othman. An Effective Color Image Encryption Based on Henon Map, Tent Chaotic Map, and Orthogonal Matrices," *Chaos, Solitons and Fractals*, vol.22, No. 12, 4359, 2022.

[10] 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.

[11] M.G.A. Malik, Z. Bashir, N. Iqbal, and Md. A. Imtiaz, "Color image encryption algorithm based on hyper-chaos and DNA computing," *IEEE Access*, vol. 8, pp. 88093-88107, 2020.

[12] M. Naim, A. A. Pacha, "A novel image encryption algorithm based on advanced hill cipher and 6D hyperchaotic system," *International Journal of Network Security*, Vol. 25, No. 5, pp. 829-840, 2023.

[13] G. Qu, X. Meng, Y. Yin, "Optical color image encryption based on Hadamard single-pixel imaging and Arnold transform," *Optics and Lasers in Engineering*, vol.137, no.20, 106392, 2021.

[14] R. L. Rivest, A. Shamir, L. M. Adleman, "A method for obtaining digital signatures and public-key cryptosystems," *Communications of the ACM*, vol.21, no.2, pp.120-126, 1978.

[15] S. Sahoo, B. K. Roy, "Design of multi-wing chaotic systems with higher largest Lyapunov exponent," *Chaos, Solitons and Fractals*, Vol.157, 111926, 2022.

[16] N. P. Smart, "The discrete logarithm problem on elliptic curves of trace one," *Journal of Cryptology*, vol.17, no.3, pp.143-145, 2004.

[17] D. R. Stinson, "Cryptanalysis of the AES: A Brief Survey," *Journal of Cryptology*, vol.15, no.2, pp.143-158, 2002.

[18] Y. Tao, W. H. Cui, Z. J. Ming, Z. Zhao and T.W. Shi, "An Image Encryption Algorithm Based on Hopfield Neural Network and Lorenz HyperChaotic System", IAENG International Journal of Computer Science, vol.49, No. 1, pp.3131-3136, 2022.

[19] J. Wang, X. Zhi, X. Chai, Y. Lu, "Chaos-based image encryption strategy based on random number embedding and DNA-level self-adaptive permutation and diffusion," *Multimedia Tools And Applications*, vol. 80, no.01, pp. 16087-122, 2021.

[20] L. Wang, "Image Encryption Based on Hyperchaotic Systems And DNA Encoding," *International Journal of Network Security*, Vol. 25, No. 3, pp. 515-521 2023.

[21] X. Wang, Y. Su, "Image encryption based on compressed sensing and DNA encoding," *Signal Processing Image Communication*, vol.12, 116246, 2021.

[22] X.Y. Wang, X.L. Wang, L. Teng, D.H. Jiang, and Y. Xian, "Lossless embedding: A visually meaningful image encryption algorithm based on hyperchaos and compressive sensing," *Chinese Physics B*, vol.32, no.2, 020503, 2023.

[23] G.Q. Xiong, Z.C. Cai, S.F. Zhao, "A bit-plane encryption algorithm for RGB image based on modulo negabinary code and chaotic system," *Digital Signal Processing*, vol.141, no.9, 104153, 2023.

[24] X. Xu, J. Feng, "Research and implementation of image encryption algorithm based on zigzag transformation and inner product polarization vector," *IEEE International Conference on Granular Computing*, San Jose, USA, August 14, 2010.

[25] X. Zhang, L. Wang, Y. Wang, Y. Niu, Y. Li, "An image encryption algorithm based on hyperchaotic system and variable-step josephus problem," *International Journal of Optics*, vol.4, pp.1-15, 2020.

[26] S. Zhou, Y. Qiu, X. Wang, "Novel image cryptosystem based on new 2d hyperchaotic map and dynamical chaotic s-box," *Nonlinear dynamics*, vol.111, pp.9571-9589, 2023.

Biography

Chunming Xu is an associate professor at the mathematics and statistics from Yancheng Teachers University, China. His main research interests include image processing and artificial intelligence.

Yong Zhang is a professor at the mathematics and

statistics from Yancheng Teachers University, China. His main research interests include cryptography and optimization.

Cost-Effective EHR Management: Image Compression and Blockchain

Faheem Ullah[1], Jingsha He[1], Nafei Zhu[1], Ahsan Wajahat[1], Ahsan Nazir[1], Siraj uddin Qureshi[1], and Hasan Shahzad[2]

(Corresponding author: Faheem Ullah)

School of Information Technology and Software Engineering, Beijing Univeristy of Technology[1]

Email: fahim.ullah@yahoo.com

School of Materials and Manufacturing, Beijing Univeristy of Technology[2]

Beijing, China

(Received Nov. 14, 2023; Revised and Accepted May 9, 2024; First Online Aug. 17, 2024)

Abstract

Effectively managing Electronic Health Records (EHRs) poses a growing challenge in today's healthcare landscape, particularly with the exponential increase in medical image data. The surge in such data not only escalates storage costs but also accentuates the immediate need for pragmatic and innovative solutions. We present a framework designed to tackle this challenge from two vital angles. Firstly, we propose a Purpose-Based Access Control (PBAC) system enforced through smart contracts, facilitating controlled access to Covid-CT scan medical images. Secondly, we delve into the domain of image compression for Covid-CT scan images, targeting a reduction in the storage expenses associated with managing large volumes of EHR data. Through the application of Run-Length Encoding (RLE) image compression techniques, we effectively minimize the storage footprint of EHR data. This dual-pronged approach showcases the potential to significantly enhance the overall efficiency and cost-effectiveness of secure EHR management, highlighting the seamless integration of PBAC and Blockchain technologies.

Keywords: EHR; EHR Security and Privacy; Image Compression; Purpose Based Access Control; Smart Contract

1 Introduction

EHRs have revolutionized healthcare by digitizing patient information, providing efficient data management, and improving patient care [1, 26, 27]. EHR systems encompass a wide range of healthcare data, including textual medical records, laboratory results, and medical images. Among these, medical images, such as X-rays, MRIs, and CT scans are pivotal for diagnosis, treatment planning, and research. However, the substantial volume of medical image data generated in modern healthcare facilities poses challenges in terms of storage, transmission, and efficient utilization [2].

Effective management and secure access control to EHRs, including medical images, are critical to ensure patient privacy and data integrity [3]. PBAC policies mechanisms have emerged as a powerful tool for governing access to healthcare data, allowing organizations to enforce fine-grained access policies based on user roles and specific purposes for access [4, 5]. The implementation of PBAC policies in healthcare systems, particularly in the context of medical image data, requires innovative approaches to ensure both security and operational efficiency. Moreover, addressing the storage and transmission challenges of medical images is paramount. Image compression techniques have been instrumental in reducing the storage requirements of medical images while preserving diagnostic quality [6]. By achieving a balance between data size reduction and image fidelity, healthcare institutions can optimize their storage resources without compromising patient care [7].

1.1 Our Contribution

This study aims to bridge the gap between secure access control, efficient data management, and image compression in the context of EHRs, focusing on medical image data. We propose a framework that combines PBAC policies enforced through smart contracts with advanced image compression techniques. Through case studies and empirical evaluations, we demonstrate the effectiveness of our approach in reducing storage requirements while maintaining the quality of medical images.

1) The paper introduced smart contract for purpose based EHR access.

2) The paper presents a comprehensive analysis of case studies showcasing the effectiveness of a proposed PBAC policy for medical image data.

3) Image processing techniques are applied to reduce the storage size of medical image data while maintaining image quality and efficiency.

The remaining sections of this article follow the following organization: Section 2 provides a brief literature review, summarizing the relevant existing research. Section 3 discusses the fundamental concepts and technologies that serve as the foundation of our approach. Section 4 presents the system architecture, outlining its components and their interactions, as well as describing the implemented access control policies. Finally, Section 5 presents detailed information about our experimental setup, including the results obtained, accompanied by an in-depth analysis of those results.

2 Literature Review

In the evolving landscape of healthcare data management, the integration of smart contract-based access control and PBAC has emerged as a central strategy to ensure data security and efficient access management [29]. These access control mechanisms, when combined, offer a robust framework that not only champions the cause of data security and privacy but also significantly enhances the efficiency quotient related to Electronic Health Record (EHR) management and storage.

Efficient management of medical image data is a paramount concern in healthcare, given the substantial volume generated daily. Addressing this challenge, a variety of image compression techniques have been explored to reduce storage requirements without compromising diagnostic quality [15]. For instance, Jayasankar *et al.* conducted a study focusing on lossless compression methods such as Huffman coding, Run-Length Encoding (RLE), and Greffier in the context of medical images, assessing their effectiveness in preserving image fidelity while achieving data reduction [9] [10].

Furthermore, researchers have examined JPEG-XT based image compression in radiology, highlighting its efficiency in reducing file sizes without introducing artifacts in radiological images [12]. Wavelet-based compression techniques have gained prominence in preserving diagnostic quality while achieving data reduction in medical imaging. Lundervold *et al.* explored wavelet-based compression in MRI and CT scan images, illustrating its potential for efficient data compression while preserving clinical usefulness. This aligns with the broader theme of image processing in medical imaging discussed in [13].

The Digital Imaging and Communications in Medicine (DICOM) standard has provisions for image compression, allowing for interoperability and efficient data sharing [24]. Pervan et.al presents MIDOM, a DICOM-based medical image communication system with custom lossless compression methods while reducing network latency and presenting efficient medical data sharing, especially in underdeveloped areas [14]. Monika et.al introduces Coefficient Mixed Thresholding-based Adaptive Block Compressed Sensing (CMT-ABCS) for efficient medical image compression while focusing on improving image quality measures. The proposed method eliminates blocking artifacts, reduces storage requirements, and minimizes runtime complexity [15].

In recent years, deep learning has emerged as a promising approach to image compression [18]. Wuzhen *et al.* investigated the application of convolutional neural networks (CNNs) for image compression, showcasing the potential of deep learning techniques to outperform traditional compression methods [17]. Additionally, researchers, such as Ishware *et al.* and Serge et.al have compared the trade-offs between lossy and lossless compression methods in medical imaging applications, weighing factors such as compression efficiency and image quality preservation [20] [21]. Moreover, established compression standards like JPEG 2000 have been extensively studied for their application in various medical imaging modalities, including CT, MRI, and ultrasound. Collectively, these studies underscore the diversity of image compression techniques applied in healthcare, each tailored to meet the unique requirements of specific applications, reflecting ongoing efforts to enhance data management efficiency while preserving clinical utility. EHRs have significantly reshaped healthcare data management, particularly with the inclusion of medical images crucial for diagnosis and patient care [25]. Managing the ever-expanding volume of medical image data requires effective solutions for storage, transmission, and retrieval.

3 System Model Architecture

We provide a brief overview of fundamental concepts and technologies that are essential for understanding the subsequent discussions in the paper as shown in Figure 1.

3.1 Preliminaries

The components and stakeholders involved in the proposed scheme are given below.

3.1.1 Purpose Based Access Control

PBAC is a finely-tuned access control model tailored to EHR storage reduction through image compression. It regulates access based on predefined purposes, safeguarding patient EHR while allowing access to authorized entities only for specified actions and duration, making it a potent security measure.

3.1.2 IPFS

IPFS, integrated into the context of EHR storage reduction through image compression, is a decentralized and scalable file storage system. It uniquely identifies and securely stores compressed EHR images using hash keys. This integration optimizes resource utilization, eliminates

redundant storage, and enhances storage efficiency for reduced EHR data size.

3.1.3 Smart Contract

In the context of EHR storage reduction through image compression, smart contracts automate agreements between parties like DRs and DCs. These contracts facilitate the enforcement of EHR storage reduction policies and the efficient management of compressed EHR data, all within a secure and transparent blockchain framework.

3.1.4 Image Compression

Image compression techniques, tailored for EHR storage reduction, are essential for managing and storing the substantial volume of medical images generated in healthcare. These techniques ensure that EHR data can be efficiently stored in a compressed format, reducing storage requirements without compromising the quality and integrity of the medical images.

3.2 System Architecture

The DR requires access to specific EHR for specific purpose, and the DC agrees via smart contract to provide the EHR after being made aware of the specific purpose in step 1 and 2 of Figure 3. In this study the DR needs EHR data for the purpose to reduce image storage size by applying image processing techniques in step 4 of Figure 3.

3.2.1 Smart Contract

To ensure that the access to the EHR is done for the intended purpose, PBAC policies are defined and deployed using smart contracts. There are three smarts contracts, Registration Smart Contract, AccessRequest Contract and Revoke contract for secure and transparent means of ensuring that the researcher only accesses the EHR for the agreed-upon purpose and that the DC's rights and interests are protected.

The Registration Smart Contract defines a registration system where EHR owner can sign up as a "DataOwner" and responsible for adding EHR data details, while EHR requester can sign up as a "Requester" request for EHR data. The enum Purpose is used to specify the purpose of the EHR access being requested. It currently only includes two options, "Trade Like" and "Research". The contract also includes enum to define two types of users and their purpose. After registration, DR made request for specific type of EHR and for the specific purpose via AccessRequest Contract. The contract includes a "requestAccess" function that allows users to request access to the EHR.

In instances where specific EHR are in high demand, the DC may request payment. Upon successful payment, access to the specified EHR is granted to the DR through the contract, marked by the emission of the AccessGranted event.

The Revoke contract empowers the DC to revoke access in response to illegitimate use. Additionally, the DR can report and seek payment revocation for unauthorized EHR access, ensuring stringent control over EHR data integrity.

3.2.2 RLE for Lossless Compression

There are two main compression techniques used in medical imaging modalities, namely lossless and lossy compression [22]. In this study we will use loseless compression technique by applying RLE. RLE is a lossless compression technique that works by identifying runs of consecutive pixels with the same intensity value and replacing them with a code that represents the length of the run and the intensity value.

In RLE, the image is scanned row by row, and each row is compressed separately. A run is defined as a sequence of consecutive pixels with the same intensity value, starting from the first pixel in the sequence and ending with the last pixel in the sequence.

To compress a run, the length of the run and the intensity value are encoded using a pair of numbers. The length of the run is encoded using a fixed number of bits, and the intensity value is encoded using a variable number of bits, depending on the range of intensity values in the image.

Let I be the original image with dimensions M x N and let i,j be the indices, of the pixel in the image I. Let R be a run of consecutive pixels with the same intensity value, represented as a pair of values: the length of the run L and the intensity value v. Let E be the encoded data generated by the RLE compression. Then, the RLE compression equation can be written as:

$$E = (L_1, v_1)(L_2, v_2)\dots(L_k, v_k) \tag{1}$$

where$k$ is the number of runs in the image, and L_i and v_i are the length and intensity value of the $i-th$ run, respectively. The length of the run L_i is the number of consecutive pixels with the same intensity value, and can be calculated as:

$$L_i = \max(j) - \min(j) + 1 \tag{2}$$

where $max(j)$ and $min(j)$ are the indices of the last and first pixel in the run, respectively. The intensity value v_i is the value of the pixels in the run, and can be calculated as:

$$v_i = I(i, j) \tag{3}$$

where i and j are the indices of any pixel in the run.

The number of bits used to encode each value can vary depending on the range of intensity values in the image, but can be fixed for all runs in the image. In practice, the length of the run L_i can be encoded using a fixed number of bits, while the intensity value v_i can be encoded using a variable number of bits, depending on the range of

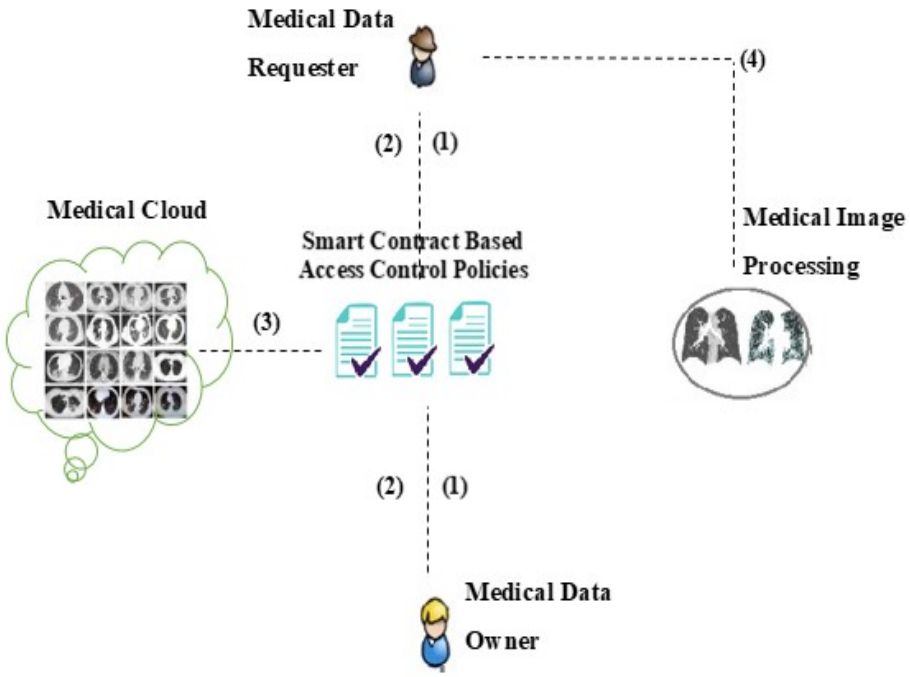


Figure 1: Proposed Technique's Workflow

intensity values in the image. After medical image compression, there are several parameters that can be used to assess the quality of the compressed medical images. Some commonly used parameters are:

Mean Square Error (MSE): This is a measure of the average squared difference between the original and compressed images. A lower MSE value indicates a higher quality image.

$$MSE = \frac{1}{N}\sum\sum(x_{ij} - y_{ij})^2 \quad (4)$$

where N is the total number of pixels in the image, xij is the value of the original image at pixel (i, j) and yij is the value of the compressed image at pixel (i, j).

Peak Signal-to-Noise Ratio (PSNR): This is a measure of the difference between the original and compressed images. A higher PSNR value indicates a higher quality image.

$$PSNR = 20 \times \log_{10}(MAX) - 10 \times \log_{10}(MSE) \quad (5)$$

where MAX is the maximum pixel value in the image and MSE is the mean squared error between the original and compressed images.

Structural Similarity Index (SSIM): This is a measure of the similarity between the original and compressed images in terms of luminance, contrast, and structure. A higher SSIM value indicates a higher quality image.

$$SSIM = \frac{(2\mu_x\mu_y + C1)(2\sigma_{xy} + C2)}{(\mu_x^2 + \mu_y^2 + C1)(\sigma_x^2 + \sigma_y^2 + C2)} \quad (6)$$

where μ_x and μ_y are the mean values of the original and compressed images, σ_x and σ_y are their standard deviations, σ_{xy} is their covariance, and $C1$ and $C2$ are constants to avoid division by zero.

Compression Ratio (CR): This is the ratio of the size of the compressed image to the size of the original image. A higher compression ratio indicates a higher level of compression, but may also result in lower image quality.

$$CR = \frac{\text{Size of compressed image}}{\text{Size of original image}} \quad (7)$$

Visual inspection: It is also important to visually inspect the compressed images to ensure that important diagnostic information is not lost or distorted during compression. These equations provide a quantitative way to assess the quality of compressed medical images.

4 Results and Analysis

4.0.1 Access Control

After registration by Registration Contract, the DR request for EHR via AccessRequest smart contract. In response to the heightened demand for specific EHR, a necessary tradeoff is introduced. Healthcare Data Users (HDU) are obligated to compensate the Healthcare Data Custodian (DC) for accessing such data following their request. The Registration contract, implemented on the

Binance Smart Chain (BSC), facilitates this process, utilizing Binance Coin (BNB) in wei denomination, as illustrated in Figure 2.

Upon successful payment, the AccessRequest contract grants the DR access to the specified EHR at a designated location for a predefined session time, illustrated in Figure 3.

The Revoke contract functions as a safeguard, protecting the DC's rights when there is inappropriate use of EHR data by the DR. In such instances, the Revoke contract denies access to the EHR data, ensuring the confidentiality of the records.

In scenarios where DRs obtain illegitimate EHR data, such as low-quality medical image records, the Revoke contract promptly initiates. Additionally, DRs can report instances of illegitimate EHR data. Figure 4 indicates the revocation of payment to DRs who have paid but received illegitimate EHR, as validated by Table 3 and Equations (1) $\sim$ (7). The BigNumber in Figure 4 represents the amount of Binance Coin in wei, which is by default large in value for testing purpose.

Utilizing the Revoke contract, DCs guarantee that their EHR data remains untainted by improper or unlawful access, while DRs receive access only to legitimate EHR data.

Figure 5 visualizes the response times of system under different PBAC policies implemented through smart contracts. As the number of access requests to the EHR system increases, the graph illustrates how response times vary for policies 10, 50, and 100.

Table 1 lists gas costs of smart contracts which is important to know the computational expenses of deploying and executing contracts. Registration Contract requires 0.002651 Eth (approximately 1.93 US$), AccessRequest Contract 0.00262 Eth (about 1.19 US$), and Revoke Contract 0.002632 Eth (approximately 0.8 US$).

Table 1: Gas Costs Used in Contracts

Contract Contract	**Gas Cost in Eth**	**Gas Cost $ in US$**
Registration Contract	0.002651	1.93
AccessRequest Contract	0.00262	1.19
Revoke Contract	0.002632	0.8

4.1 Image Processing Compression Technique

We applied the RLE compression technique to the CT-Scan images of lungs which where taken from dataset [27] and [28] to conduct experiments. Three kinds of three CT-Scan images of lungs were chosen : a healthy lung image, a moderately COVID-19 affected lung image, and a low-quality blurry image as illegitimate EHR data. The objective is to reduce image storage size while maintaining image quality so accordingly, we adjusted the compression ratio based on the quality of each image.

4.1.1 Applying RLE Compression Technique to Image of Healthy Lungs

Run-Length Encoding (RLE) is a simple yet effective lossless compression technique that capitalizes on the presence of consecutive repeated values in data. In the context of medical imaging, especially CT-Scan images of lungs, the application of RLE can be particularly beneficial given the repetitive patterns often found in such images.

For our experiment, we selected a high-resolution CT-Scan image of a healthy lung as a representative sample. This image, characterized by its clear and distinct patterns, serves as an ideal candidate to demonstrate the efficiency of the RLE compression technique.

Upon applying RLE to the original healthy lung image, we obtained compressed versions of the image, which are visually represented in Figure 6 Original and in Figure 7 compressed respectively. The primary goal of this compression was not only to reduce the storage size but also to ensure that the diagnostic quality of the image remains uncompromised.

A comparative analysis between the original and compressed images was conducted to evaluate the effectiveness of the RLE compression. The results of this analysis are tabulated in Table 2. As observed, the Mean Squared Error (MSE), Peak Signal-to-Noise Ratio (PSNR), Root Mean Square Error (RMSE), Artifact Power, Multi-Scale Structural Similarity Index (MSSIM), Compression Ratio (CR), and Image Size parameters provide a comprehensive insight into the quality and efficiency of the compression.

The data suggests that while there is a noticeable reduction in image size (from 72 KB to 25 KB), the image quality metrics, especially PSNR and MSSIM, indicate that the integrity and clarity of the image have been largely preserved. This underscores the potential of RLE as a viable compression technique for medical images, balancing both storage efficiency and image quality.

4.1.2 Applying RLE Compression Technique to CT-Scan Image of Moderately COVID-19-Affected Lung

In the realm of medical imaging, CT-Scan images of lungs affected by COVID-19 present unique challenges due to the varying patterns of lung damage caused by the infection. For this experiment, a CT-Scan image showcasing a moderate level of COVID-19 infection was chosen. This image, with its distinct patterns of infection, offers a unique opportunity to evaluate the efficacy of RLE compression in preserving diagnostic details.

After applying RLE to the original image, the compressed versions were obtained, as depicted in Figure 8 original) and in Figure 9 (compressed) respectively. A detailed analysis, summarized in Table 3 was conducted to assess the quality of the compressed image. The metrics, including MSE, PSNR, RMSE, Artifact Power, MSSIM, CR, and Image Size, suggest that while the image size

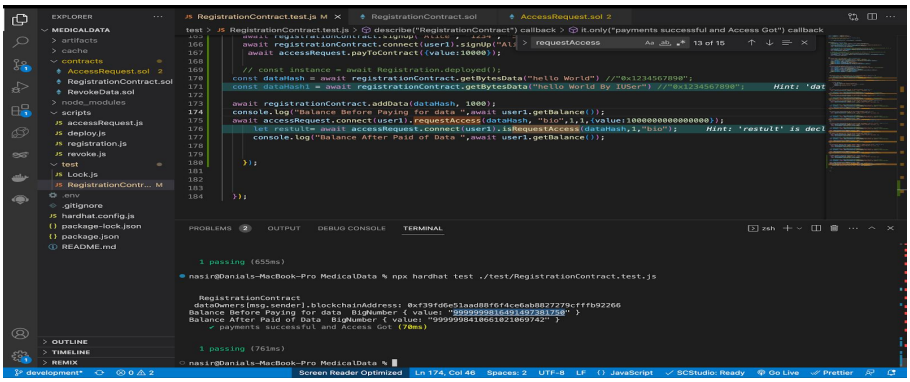


Figure 2: Successful Payment for EHR Access

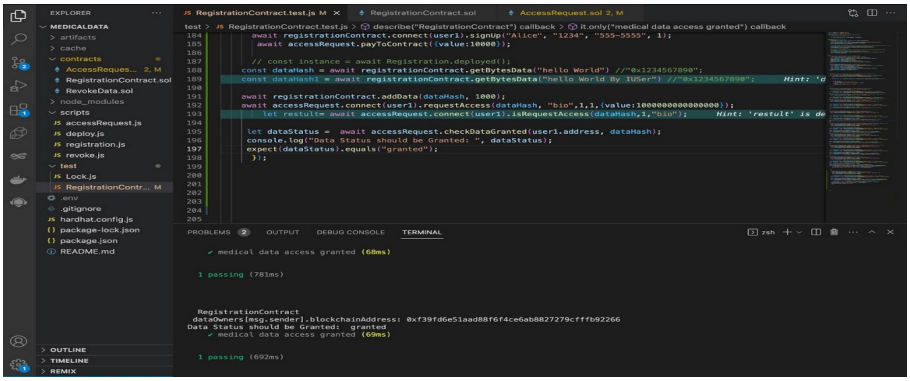


Figure 3: EHR Access Granted

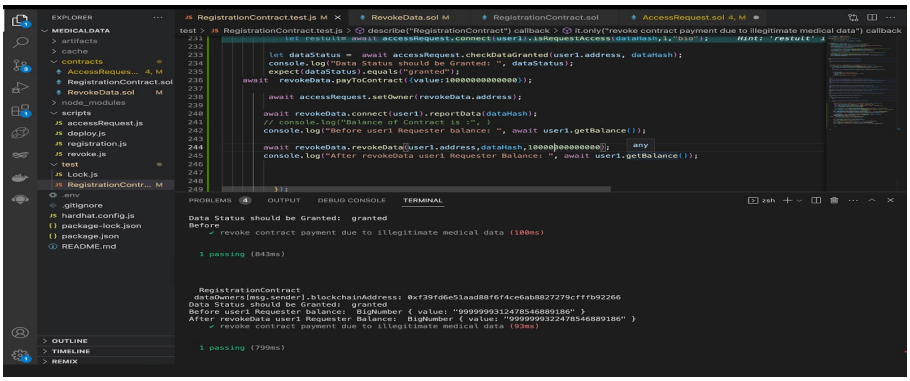


Figure 4: Revoke Payment After Receiving Illegitimate EHR

Table 2: Quality Performance Parameter of Healthy Lung

Test Image	MSE	PSNR	RMSE	Artifact Power	MSSIM	CR	Image Size (KB)
Original	311.84	29.24	27	0.03	0.49	87	72
Compressed	395	28	29.99	0.032	0.043	88	25

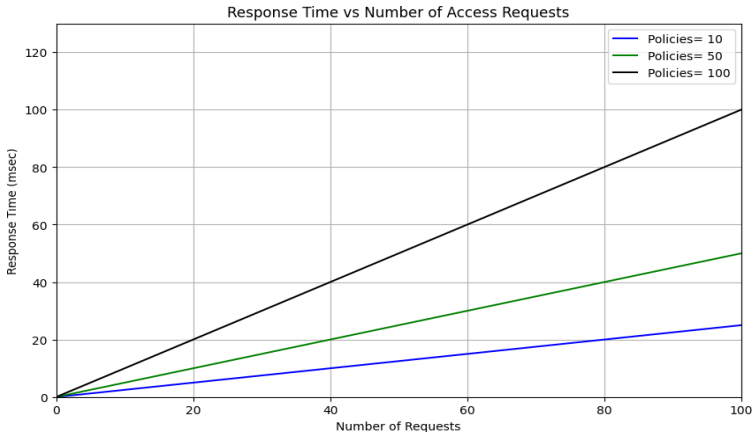


Figure 5: Response Time for Varying Policies

Figure 6: Original Image of Healthy Lungs

Figure 7: Compressed Image of Healthy Lungs

was significantly reduced (from 74 KB to 22 KB), the diagnostic quality remained largely intact, making RLE a promising technique for compressing CT-Scan images of COVID-19 affected lungs.

4.1.3 Applying RLE Compression Technique to CT-Scan Lung Image of Low Quality

Low-quality medical images, especially those that are blurry or lack clarity, pose significant challenges for compression techniques due to the inherent lack of distinct patterns. The inherent lack of distinct patterns can make the compression process less effective and, in some cases, further degrade the image quality.

In this context, a low-quality CT-Scan lung image was subjected to RLE compression. The resulting compressed images are illustrated in Figure 10 . Note that as shown in Figure 10 failed to yield satisfactory results for quality performance parameters as shown in Table 4, which was further confirmed through visual inspection. The RLE compression did not yield satisfactory results for this particular image. The degradation in quality was evident, suggesting that RLE might not be suitable for compressing low-quality medical images due to significant degradation in image quality.

Furthermore, such unsatisfactory results can have implications in the medical diagnosis process. As discussed in section 4.0.1 and as illustrated in Figure 10. DR can report these inadequacies and potentially request a revocation of payment due to the compromised/ illegitimate image quality. This emphasizes the importance of ensuring that compression techniques are carefully chosen based on the quality and characteristics of the medical images.

A sample of 50 lung images, comprising both normal and images affected by COVID-19, was extracted from dataset [27] and [28] to conduct an empirical evaluation of the RLE image compression technique. The experimen-

Table 3: Quality Performance Parameter of Moderately COVID-19 Affected Lung Image

Test Image	MSE	PSNR	RMSE	Artifact Power	MSSIM	CR	Image Size (KB)
Original	311.84	29.24	27	0.03	0.49	87	74
Compressed	395	28	29.99	0.032	0.043	88	22

Table 4: Quality Performance Parameter of COVID-19 Lung Image

Test Image	MSE	PSNR	RMSE	Artifact Power	MSSIM	CR	Size (KB)
Original	1.8	45	1.34	1222901	0.99	10	45

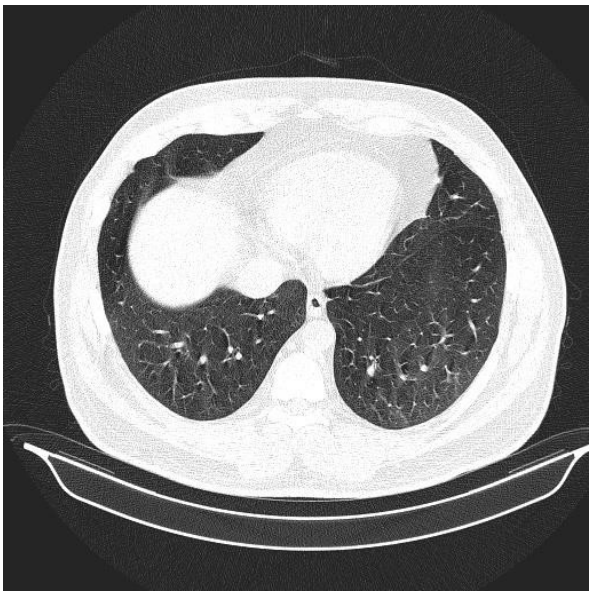


Figure 8: Original Image of Covid 19-Effected Lungs

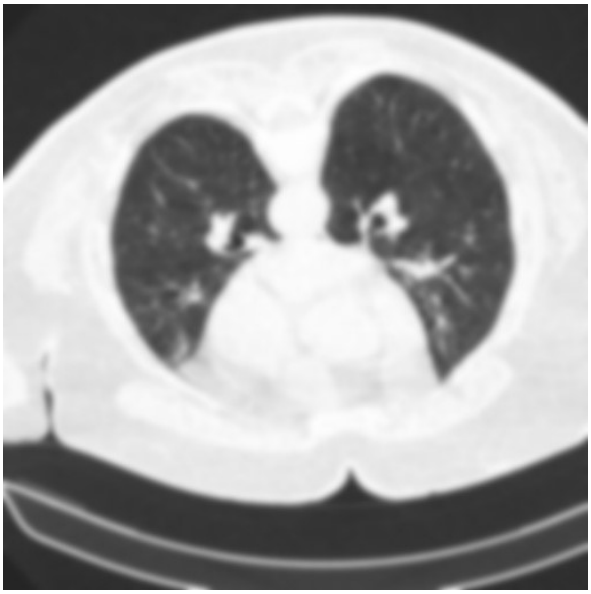


Figure 10: Illegitimate Image

Figure 9: Compressed Image of Covid-19 Effected Lungs

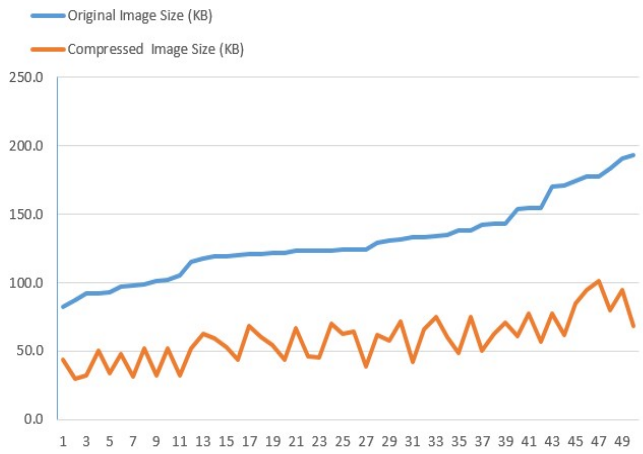


Figure 11: Image Size Before and After Compression

tal results demonstrated a substantial reduction in image size, with compression ratios ranging from 45% to 68%, as depicted in the Figure 11.

4.1.4 Limitations

This technique is primarily recommended for deployment in contexts prioritizing cost efficiency, where minor compromises to image quality are acceptable.However, its effectiveness may be limited in scenarios where high-fidelity image reproduction is crucial for detailed image analysis and diagnostic evaluation. For instance, its suitability for subnodular applications may be limited.

4.1.5 The Experiment Setup Information

The experiments of this scheme were conducted on macos ventura version 13.2.1 with Apple M1 Pro i9 @3.20 GHz and 16GB RAM. We used solidity 0.8.18 latest version with hardhat also used hardhat optimizer enabled with 100 and deployed it on hardhat local network, ganache network and BSC testnetwork.to verify all code on BSC explorer to show transparently all codes.

Our image processing setup utilized a high-performance workstation with a dedicated GPU and high-speed SSDs for efficient data handling. The software suite included MATLAB, Python with OpenCV, Adobe Photoshop, and ImageJ, catering to various image enhancement and analysis needs. Equipment like high-resolution scanners tools like GIMP and NIH Image facilitated additional manipulation and statistical analysis of images.

5 Conclusion and Future Work

In this study, we introduced an innovative approach that integrates PBAC policies with blockchain's smart contracts to fortify the security and streamline the management of EHR data. By synergizing PBAC's robust access control mechanisms with the immutable and transparent nature of smart contracts, we've crafted a formidable defense against potential breaches, ensuring utmost privacy and security of EHRs.

Furthermore, the incorporation of image compression techniques not only optimizes storage but also enhances the efficiency of EHR data management. Looking ahead, we envision refining our model with advanced cryptographic techniques and exploring the potential of AI-driven analytics to further enhance the integrity and utility of EHR systems.

References

[1] M. Paul, L. Maglaras, M. A. Ferrag, and I. AlMomani, *Digitization of healthcare sector: A study on privacy and security concerns*, ICT Express, Elsevier, 2023.
@articlepaul2023digitization, title=Digitization of healthcare sector: A study on privacy and security concerns, author=Paul, Metty and Maglaras, Leandros and Ferrag, Mohamed Amine and AlMomani, Iman, journal=ICT Express, year=2023, publisher=Elsevier

[2] Mengfang Li, Yuanyuan Jiang, Yanzhou Zhang, and Haisheng Zhu, *Medical image analysis using deep learning algorithms*, *Frontiers in Public Health*, vol. 11, pp. 1273253, 2023, Frontiers Media SA.

[3] M. J. H. Faruk, H. Shahriar, B. Saha, and A. Barek, *Security in Electronic Health Records System: Blockchain-Based Framework to Protect Data Integrity*, in *Blockchain for Cybersecurity in Cyber-Physical Systems*, pp. 125–137, Springer, 2022.

[4] G. Wu, S. Wang, Z. Ning, et al., *Blockchain-enabled privacy-preserving access control for data publishing and sharing in the internet of medical things*, IEEE Internet of Things Journal, vol. 9, no. 11, pp. 8091–8104, IEEE, 2021.

[5] G. Wu, S. Wang, Z. Ning, and B. Zhu, *Privacy-preserved electronic medical record exchanging and sharing: A blockchain-based smart healthcare system*, IEEE Journal of Biomedical and Health Informatics, vol. 26, no. 5, pp. 1917–1927, IEEE, 2021.

[6] R. Sekaran, V. K. M. Nagaraju, V. Jagadeesan, M. Ramachandran, and A. Kumar, *Medical Data Compression for Lossless Data Transmission and Archival*, Internet of Medical Things: Remote Healthcare Systems and Applications, pp. 55–74, Springer, 2021.

[7] M. Hartmann, U. S. Hashmi, and A. Imran, *Edge computing in smart health care systems: Review, challenges, and research directions*, Transactions on Emerging Telecommunications Technologies, vol. 33, no. 3, pp. e3710, Wiley Online Library, 2022.

[8] R. Monika and Samiappan Dhanalakshmi, *An efficient medical image compression technique for telemedicine systems*, *Biomedical Signal Processing and Control*, vol. 80, pp. 104404, 2023, Elsevier

[9] U. Jayasankar, V. Thirumal, and P. Dhavachelvan, *A survey on data compression techniques: From the perspective of data quality, coding schemes, data type and applications*, Journal of King Saud University-Computer and Information Sciences, vol. 33, no. 2, pp. 119–140, Elsevier, 2021.

[10] J. Greffier, A. Hamard, F. Pereira, C. Barrau, H. Pasquier, J. P. Beregi, and J. Frandon, *Image quality and dose reduction opportunity of deep learning image reconstruction algorithm for CT: a phantom study*, European radiology, vol. 30, pp. 3951–3959, Springer, 2020.

[11] Z. Li, A. Ramos, Z. Li, M. L. Osborn, X. Li, Y. Li, S. Yao, and J. Xu, *An optimized JPEG-XT-based algorithm for the lossy and lossless compression of 16-bit depth medical image*, Biomedical Signal Processing and Control, vol. 64, pp. 102306, Elsevier, 2021.

[12] Z. Li, A. Ramos, Z. Li, M. L. Osborn, X. Li, Y. Li, S. Yao, and J. Xu, *An optimized JPEG-XT-based algorithm for the lossy and lossless compression of 16-bit depth medical image*, Biomedical Signal Processing and Control, vol. 64, pp. 102306, Elsevier, 2021.

[13] Vlad-Ilie Ungureanu, Paul Negirla, and Adrian Korodi, *Image-Compression Techniques: Classical and "Region-of-Interest-Based" Approaches Presented in Recent Papers*, *Sensors*, vol. 24, no. 3, pp. 791, 2024, MDPI.

[14] B. Pervan, S. Tomic, H. Ivandic, and J. Knezovic, *MIDOM—A DICOM-Based Medical Image Communication System*, Applied Sciences, vol. 13, no. 10, pp. 6075, MDPI, 2023.

[15] R. Monika and S. Dhanalakshmi, *An efficient medical image compression technique for telemedicine systems*, Biomedical Signal Processing and Control, vol. 80, pp. 104404, Elsevier, 2023.

[16] N. Krishnaraj, M. Elhoseny, M. Thenmozhi, M. M. Selim, and K. Shankar, *Deep learning model for real-time image compression in Internet of Underwater Things (IoUT)*, Journal of Real-Time Image Processing, vol. 17, pp. 2097–2111, Springer, 2020.

[17] Ehsaneddin Jalilian, Heinz Hofbauer, and Andreas Uhl, *Iris image compression using deep convolutional neural networks*, *Sensors*, vol. 22, no. 7, pp. 2698, 2022, MDPI.

[18] A. Murat Tekalp, Michele Covell, Radu Timofte, and Chao Dong, *Introduction to the issue on deep learning for image/video restoration and compression*, *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 2, pp. 157–161, 2021, IEEE.

[19] V. S. Guntuboina, *Efficient Image Data Compression Techniques: A Comprehensive Review and Comparative Study*.

[20] T. Ishware and S. Metkar, *Comparative Analysis of Various Standards for Medical Image Compression*, in *International Symposium on Intelligent Informatics*, pp. 351–363, Springer, 2022.

[21] Sergey Krivenko, Vladimir Lukin, Olha Krylova, Liudmyla Kryvenko, and Karen Egiazarian, *A fast method of visually lossless compression of dental images*, *Applied Sciences*, vol. 11, no. 1, pp. 135, 2020, MDPI.

[22] Yaghoub Pourasad and Fausto Cavallaro, *A novel image processing approach to enhancement and compression of X-ray images*, *International Journal of Environmental Research and Public Health*, vol. 18, no. 13, pp. 6724, 2021, MDPI.

[23] S. Kaur, G. Chaudhary, J. D. Kumar, M. S. Pillai, Y. Gupta, M. Khari, V. García-Díaz, and J. Parra Fuente, *Optimizing fast fourier transform (FFT) image compression using intelligent water drop (IWD) algorithm*, International Journal of Interactive Multimedia and Artificial Intelligence . . . , 2022.

[24] Michele Larobina, *Thirty years of the DICOM standard*, *Tomography*, vol. 9, no. 5, pp. 1829–1838, 2023, MDPI.

[25] T. Xue, Y. Wen, B. Luo, G. Li, Y. Li, B. Zhang, Y. Zheng, Y. Hu, and D. Meng, *SparkAC: Fine-Grained Access Control in Spark for Secure Data Sharing and Analytics*, IEEE Transactions on Dependable and Secure Computing, vol. 20, no. 2, pp. 1104–1123, IEEE, 2022.

[26] Statista, *EHR access among adults U.S. 2018*, Sept. 4, 2023. (`https://www.statista.com/statistics/829500/electronic-health-record-access-us/`)

[27] Office of the National Coordinator for Health Information Technology, *Adoption of Electronic Health Records by Hospital Service Type 2019-2021*, April 2022. (`https://www.healthit.gov/data/quickstats/adoption-electronic-health-records-hospital-service-type-2019-2021`)

[28] I. Abedi, M. Vali, B. O. Shahreza, and H. Bolhasani, *HRCTv1-COVID-19: A High Resolution Chest CT Scan Image Dataset for COVID-19 Diagnosis and Differentiation*, Mendeley Data, vol. V2, 2022.

[29] Mpyana Mwamba Merlec and Hoh Peter In, *SCCAAC: A Smart Contract-Based Context-Aware Access Control Scheme for Blockchain-Enabled IoT Systems*, *IEEE Internet of Things Journal*, 2024, IEEE.

Biography

Faheem Ullah received M.S degrees from the Xian Jiaotong University, China,in 2017. He is currently pursuing a Ph.D. degree at the Beijing University of Technology, Beijing, China. His research interests include information security, Blockchain, Access Control and EHR.

Jingsha He received a bachelor's degree in computer science from Xi'an Jiaotong University, China, and the mas- ter's and Ph.D. degrees in computer engineering from the University of Maryland, College Park, MD, USA. He worked for several multinational companies in USA, including IBM Corp., MCI Communications Corp., and Fujitsu Laboratories. He is currently a Professor with the Faculty of Information Technology, Beijing University of Technology(BJUT), Beijing. He has published more than ten articles. He holds 12 U.S. patents. Since August 2003, he has been published over 300 papers in scholarly journals and international conferences. He also holds over 84 patents and 57 software copyrights in China and authored nine books. He was a principal investigator of more than 40 research and development projects. His research interests include information security, wireless networks, and digital forensics.

Nafei Zhu received the B.S. and M.S. degrees from Cen- tral South University, China, in 2003 and 2006, respectively, and the Ph.D. degree in computer science and technology from the Beijing University of Technology, Beijing, China, in 2012. She was a Postdoctoral Research

Fellow with the State Key Laboratory of Computer Science, Institute of Software, Chinese Academy of Sciences, Beijing, from 2015 to 2017. She is currently an Associate Professor at the Faculty of Information Technology, Beijing University of Technology. She has published over 20 research papers in scholarly journals and international conferences. Her research interests include information security and privacy, wireless communications, and network measurement.

Ahsan Wajahat received the B.S. and M.S degrees in information technology from the Sindh Agriculture University, Pakistan in 2012 and 2016, respectively. He is currently pur- suing a Ph.D. degree at the Beijing University of Technology, Beijing, China. His research interests include machine learn- ing, information security, forensic networks and data mining.

Ahsan Nazir has received his M.Sc degree from the University of Engineering and Technology Lahore in 2016. From September 2015 to August 2018 he worked as Software Engineer at Dunya Media group Lahore since September 2018 he is doing Ph.D. in Software Engineering from Beijing University of Technology, Beijing China . He has published more than 10 journals and conference papers .His area of research include eGovernment,IoT, Software Engineering and Machine learning applications.

Siraj uddin Qureshi received the B.S. and M.S degrees in information technology from the Sindh Agriculture University, Pakistan in 2012 and 2016, respectively. He is currently pursuing a Ph.D. degree at the Beijing University of Technology, Beijing, China. His research interests include machine learning, information security, forensic networks and data mining. Currently, he is pursuing PhD in Information Technology at the Beijing University of Technology, China. He has nine research publications to his credit as main author and co-author, which featured national and international journals and conferences. Sirajuddin research areas include but are not limited to Network Forensics Analysis, Digital Forensics, Cyber security, Computer Networks and Network Security.

Hasan Shahzad is currently a Postdoctoral Researcher at Dongguan University of Technology in Dongguan, China, a position he has held since September 2023. He earned his Doctorate from Beijing University of Technology, where he conducted research from 2018 to 2023. Hasan's research focuses on advanced thermal and fluid dynamics, particularly in the study of nanofluids, magnetohydrodynamics, and fluid- structure interactions. His work often employs cutting-edge computational methods, including finite element methods, ma- chine learning, and artificial neural networks.

Detecting IoT Botnet Attacks through Ensemble and Meta Ensemble Approaches

Xiangjun Ma, Jingsha He, Ahsan Nazir, Nafei Zhu, Xiao Hu, Faheem Ullah, Ahsan Wajahat, Yehong Luo, And Sirajuddin Qureshi

(Corresponding authors: Ahsan Nazir)

Faculty of Information Technology, Beijing University of Technology
Beijing 100124, China
Email: ahsan_ravian@hotmail.com

(Received Nov. 28, 2023; Revised and Accepted May 9, 2024; First Online Aug. 17, 2024)

Abstract

The escalating growth of the Internet of Things (IoT) has precipitated heightened concerns regarding security, particularly with the proliferation of IoT botnets orchestrating malicious activities. This study addresses the critical imperative for the precise detection of IoT botnet attacks by introducing a pioneering approach based on ensemble and meta-ensemble techniques. Drawing upon the N-BaIoT dataset, this research explores the efficacy of ensemble strategies and innovative meta-ensemble methods to rectify existing gaps in IoT security measures. The results exhibit exceptional accuracy, with the ensemble model achieving a peak rate of 99.93%. The meta-ensemble method further enhances overall detection accuracy to an impressive 99.95%. These findings underscore the robust performance of the proposed approach across diverse IoT scenarios, signifying a significant advancement in IoT security. By offering insights into effective botnet attack detection and fortifying the integrity of IoT networks, this research contributes substantially to the ongoing discourse in cybersecurity.

Keywords: Ensemble Learning; IoT Botnet; IoT Security; Internet of Things; Machine Learning

1 Introduction

The transformative expansion of the Internet of Things (IoT) has reshaped various sectors, introducing unprecedented connectivity and automation in domains such as healthcare, transportation, and smart homes [1, 2]. As this exponential growth unfolds, it brings with it significant challenges, notably in the realm of cybersecurity. With billions of interconnected devices, the escalating potential for malicious activity and cyberattacks underscores the critical need to secure IoT systems [3, 4]. Among the myriad threats, IoT botnets have emerged as a major concern, orchestrating large-scale distributed denial of service (DDoS) attacks, unauthorized data access, and malware propagation [5, 6]. Detecting and mitigating these botnet attacks is imperative for safeguarding the security and integrity of IoT networks [7].

In response to the limitations of existing IoT botnet detection methods, this study proposes an innovative machine learning-based approach. Traditional methods face challenges in terms of accuracy, adaptability, and resilience [8, 9]. Leveraging Ensemble and Meta Ensemble, which harness the collective intelligence of multiple machine learning classifiers, our study aims to provide a more robust solution. The N-BaIoT dataset, capturing real-world IoT network traffic, is employed to gain insights into botnet attack patterns [10]. This dataset provides us with valuable insights into the unique characteristics and patterns of botnet attacks in the IoT environment. Table 1 presents a comprehensive overview of devices along with their corresponding Benign and Malicious data. The provided device IDs will be referenced throughout the research article to uniquely identify each device. The table includes a total of 9 real IoT devices which make up the N-BaIoT collection. Notably, each separate device corresponds to a distinct dataset, equating to a total of nine independent datasets. We use the most advanced machine learning algorithms, including Naive Bayes, Decision Trees, and Random Forests, to develop individual classifiers that can effectively identify botnet attacks. These classifiers are trained and evaluated using the N-BaIoT dataset, allowing us to evaluate their performance and identify their strengths and weaknesses. We apply all classifiers on each Device's dataset and this gives us a handful of information about IoT Botnets.

Despite significant advancements in IoT security measures, the detection and mitigation of botnet attacks remain key challenges. Existing methodologies often fall short in detecting innovative botnet activities amidst legitimate IoT traffic, primarily due to their limited robustness and adaptability. Also, the dynamic nature of IoT environments exacerbates the complexity of detecting anomalous behaviors indicative of botnet incursions. Traditional approaches falter in keeping pace with the

evolving tactics employed by malicious actors, thereby leaving IoT networks vulnerable to exploitation. Addressing these challenges requires the development of innovative detection techniques that can accurately identify and mitigate IoT botnet threats. By leveraging advanced machine learning techniques such as Ensemble and Meta Ensemble, our research aims to bridge this gap by providing a robust and adaptive solution capable of enhancing the security of IoT systems.

The research methodology revolves around the collective intelligence of diverse machine learning algorithms to construct a robust botnet detection framework. Initially, we employ a Voting Classifier in the Ensemble phase, which amalgamates the predictions of multiple base classifiers, including Naive Bayes, Decision Trees, and Random Forests, to yield a consensual decision. This ensemble approach harnesses the diverse strengths of individual classifiers, thereby enhancing the overall detection accuracy and resilience. Subsequently, in the Meta Ensemble phase, we employ a Gradient gradient-boosting classifier, which iteratively trains weak learners to form a strong learner capable of discerning intricate patterns inherent in botnet activities. This meta-ensemble strategy further refines the detection capabilities by leveraging the feedback loop between base classifiers, enabling the system to adapt dynamically to evolving threats. By integrating these advanced techniques with the comprehensive analysis of the N-BaIoT dataset, our methodology facilitates a holistic understanding of botnet behaviors and empowers the detection framework with the agility required to mitigate emerging threats effectively.

1.1 Our Contribution

This research represents a significant advancement in the field of IoT security by introducing an ensemble and meta-ensemble approach tailored for IoT botnet detection. By surpassing the limitations of conventional methodologies, the study lays the foundation for a robust and adaptive detection framework capable of mitigating the escalating threat landscape posed by IoT botnets. The utilization of ensemble strategies, coupled with meta-ensemble techniques, offers a paradigm shift in enhancing detection accuracy and resilience. Through rigorous experimentation and analysis, the efficacy of the approach in discerning complex botnet activities amidst legitimate IoT traffic is demonstrated, thereby fortifying the security posture of IoT networks. Moreover, the research contributes to the broader discourse on cybersecurity by showcasing the potential of advanced machine learning techniques in addressing emergent threats in IoT ecosystems. The insights obtained from this study pave the way for the development of proactive security measures aimed at safeguarding the integrity and confidentiality of IoT systems against malicious intrusions.

- Development of Ensemble and Meta-Ensemble Models: We propose innovative ensemble and meta-ensemble models tailored for IoT botnet attack detection, leveraging the collective strength of multiple classifiers to enhance detection accuracy and robustness.

- Comprehensive Evaluation Across Diverse Datasets: Our study conducts a thorough evaluation of the proposed models across nine diverse IoT device datasets, demonstrating their adaptability and generalization across various sub-datasets. This comprehensive assessment underscores the models' suitability for real-world IoT environments.

- Superior Performance and Strategic Technique Integration: Our proposed models consistently outperform existing state-of-the-art studies, showcasing superior accuracy, precision, recall, and F1 score. By strategically integrating ensemble and meta-ensemble methods, our approach addresses the complex challenges of IoT botnet attack detection, providing valuable insights for future research in IoT security.

2 Literature Review

In this section, we present a comprehensive review of the existing literature and research concerning IoT botnet attack detection and machine learning methodologies. Various methodologies, techniques, and algorithms proposed in prior studies are explored, shedding light on their strengths, limitations, and the gaps they leave to be addressed. The primary aim of this literature review is to contextualize our research and underscore the novelty and contribution of our work in the field.

Hostiadi *et al.* (2016) utilized the benchmark BoTNet dataset and demonstrated that a randomized data partitioned learning model-based approach achieve an accuracy of 99.98% in just 21.38 seconds of training time [11]. Similarly, Chowdhury *et al.* (2016) emphasized the necessity of employing two machine learning algorithms, achieving a false positive rate of 0.09%, to reliably detect network traffic incursions [12]. Researchers have also developed strategies for identifying questionable texts utilizing a range of unsupervised machine learning approaches, yielding accuracy rates as high as 90% [13]. Indre *et al.* (2016) proposed connecting digital enterprises to the supply chain to mitigate security concerns associated with IoT infrastructure [14].

The proliferation of IoT botnets as a significant threat to IoT system security has spurred extensive research into effective detection methods. Anomaly-based detection techniques, leveraging statistical models and machine learning algorithms, have been widely investigated to identify abnormal behaviors in network traffic. Studies by Asgharzadeh *et al.* (2023) and Bhavsar *et al.* (2023) have proposed anomaly detection approaches utilizing features extracted from IoT network traffic, demonstrating promising results in detecting botnet activities [15,16].

Table 1: Dataset Information for 9 Real IoT Devices in N-BaIoT

Device ID	Device Name	Benign Rows	Botnets Rows
1	Danmini_Doorbell	49548	886768
2	Ecobee_Thermostat	13113	735395
3	Ennio_Doorbell	39100	846306
4	Philips_B120N10_Baby_Monitor	175240	775156
5	Provision_PT_737E_Security_Camera	62154	766106
6	Provision_PT_838_Security_Camera	98514	738377
7	Samsung_SNH_1011_N_Webcam	52150	839269
8	SimpleHome_XCS7_1002_WHT_Security_Camera	46585	856870
9	SimpleHome_XCS7_1003_WHT_Security_Camera	19528	831298

Signature-based detection techniques, relying on predefined attack patterns, have also been explored [17]. However, these techniques are limited in their ability to detect novel IoT botnet attacks.

Research in IoT security has seen significant advancements, with studies exploring various techniques for botnet detection. Sakthipriya *et al.* (2023) conducted a comparative analysis of dimensionality reduction methods, revealing the superiority of the autoencoder algorithm in achieving a remarkable accuracy of 95.02% on the N-BaIoT dataset [18]. Meanwhile, Umair *et al.* (2023) investigated the efficacy of spiking neural networks (SNNs) for malware classification, achieving 71% accuracy and highlighting the potential of SNNs in event-driven classification tasks [19]. AL-Akhras *et al.* (2023) delved into intrusion detection systems (IDS) for IoT environments, showcasing the effectiveness of noise filtering algorithms like RENN and DROP5 in improving accuracy on datasets including N-BaIoT. These studies collectively contribute to bolstering the security framework of IoT systems through innovative detection methodologies [20].

Machine learning has garnered significant attention for IoT botnet attack detection due to its ability to assess intricate patterns and adapt to new threats. Popular machine learning algorithms employed in this context include Naive Bayes, Decision Trees, and Random Forests. Studies by Garg *et al.* (2021) and Ahmed *et al.* (2022) have examined the effectiveness of Naive Bayes in detecting botnet attacks by analyzing network traffic features [21, 22]. Decision Tree algorithms, as investigated in the study by Saif *et al.* (2023), have shown promising results in capturing decision rules for distinguishing botnet traffic from normal traffic [23]. Similarly, Random Forest, an ensemble method, has been employed in various studies, showcasing its ability to improve detection accuracy by combining multiple decision trees [24, 25]. Ensemble methods, such as the Voting Classifier, have been proposed to enhance the performance of individual classifiers by combining their predictions, thereby improving the detection accuracy and robustness of IoT botnet attack detection systems [26].

The Internet of Things (IoT) botnet attack detection and prevention studies are comprehensively summarized in Table 2. The many machine learning and deep learning models utilized in each study are highlighted, along with the datasets used and the model accuracy. The significant results of each study have been put together to show improvements in IoT botnet intrusion detection and prevention approaches. These studies address growing security issues that involve IoT networks and gadgets, which have grown in prominence as targets for illicit activity on the Internet. Researchers developed useful ways for identifying botnet attacks by comparing and analyzing the performance of several models, enhancing accuracy, precision, recall, and F1-score. The insights provided in this table will contribute to the literature review of our research article, providing a valuable reference for selecting the most suitable model for intelligent botnet attack detection in the IoT and identifying the gaps and opportunities for further research in this field. Our research intends to contribute to the development of more robust and efficient IoT botnet attack detection systems by addressing recognized research gaps and limitations.

3 Methodology

In this section, we elucidate the methodology employed to detect botnet assaults within IoT networks, utilizing a diverse ensemble and meta-ensemble approach. The study leverages base classifiers including Naive Bayes, Decision Trees, and Random Forest, integrated within Ensemble and Meta-Ensemble frameworks. Anchored on the N-BaIoT dataset, which comprises network traffic data from nine real-world IoT devices, each device housing a distinct dataset, rigorous preprocessing was undertaken to streamline the dataset for effective machine learning application. The methodology entails training these base classifiers on the individual datasets representing network traffic from each IoT device, encompassing both benign and botnet activity. Notably, within each dataset, all malicious classes have been combined into one, labeled as "Malicious," while benign activities are categorized as "Benign." Subsequently, predictions from these base classifiers are aggregated to construct an Ensemble Model utilizing the Voting Classifier technique. Additionally, a meta-ensemble model is deployed to further refine detection capabilities. Cross-validation techniques are utilized to ensure impartial evaluation and comparison of

Table 2: Related Studies on IoT botnet attack detection

Study	Models Applied	Dataset Used	Contribution to the Field
Asgharzadeh *et al.* [15]	CNN	NSL-KDD, TON-IoT	Enhances anomaly detection in IoT using DL
Bhavsar *et al.* (2023) [16]	Logistic Regression, K Nearest Neighbour, CART, Support Vector Machine	NSL-KDD, CICIDS-2017, and IOTID20	Proposed Deep Learning based IDS outperforms traditional methods
Suthar *et al.* (2022) [17]	Signature-based mechanism	Emotet	Proposed Snort Based Botnet Detection System
Garg *et al.* (2021) [21]	Naïve Bayes, Random Forest, Support Vector Machine	UNSW NB15	Proposed an intelligence-based system that can investigate or detect the intrusion in the IoT botnet
Ahmed *et al.* (2022) [22]	Multiple ML Models	Ton-IoT Dataset	Analysis of supervised ML algorithms on Ton-IoT Dataset
Saif *et al.* (2023) [23]	RF, Naïve Bayes, DT, SVM, Logistic Regression	N-BaIoT Dataset	Advances botnet attack detection in IoT
Padmashree *et al.* (2022) [24]	Naive Bayes, SVM, Gradient Boosting, Logistic Regression	KDD CUP99 dataset	Proposed ML based IoT Botnet Detection system
Motylinski *et al.* (2022) [25]	RF,KNN,SVM, Logistic R.	IoT Bot dataset	Reduces training and estimation times for large-scale systems
Sakthipriya *et al.* (2023) [18]	Adaptive Voting Classifier	BoT IoT dataset	Improves detection of all attack categories in highly imbalanced datasets
Sakthipriya *et al.* (2023) [18]	Autoencoder, PCA	N-BaIoT Dataset	Demonstrates the effectiveness of dimensionality reduction methods for IoT botnet detection
Umair *et al.* (2023) [19]	Spiking Neural Networks (SNNs)	N-BaIoT Dataset	Highlights the potential of SNNs for IoT botnet detection
AL-Akhras *et al.* (2023) [20]	RENN, Explore, DROP5	IoTID20, N-BaIoT, MedBIoT Datasets	Shows the effectiveness of noise filtering algorithms in improving IDS accuracy

model performance across the nine sub-datasets. This section offers a comprehensive overview of the methodology adopted to address the research objectives effectively, focusing on ensemble and meta-ensemble techniques applied to experimentation on multiple IoT datasets. Figure 1 depicts the processes involved in this research Methodology, as well as the ensemble and meta Ensemble methodologies used. The figure also shows the primary dataset, which consists of 9 IoT devices. Generic steps to be taken to conduct this research are presented in algorithm 1.

3.1 Dataset

In this research, we conducted an extensive analysis of a comprehensive dataset named the N-BaIoT dataset [6], which encompassed network traffic data from nine distinct IoT devices. The IoT devices used in the dataset are real-world devices commonly found in homes, making the dataset more representative of actual IoT network scenarios. The dataset consisted of both benign and malicious traffic classes, rendering valuable insights into the security aspects of these devices. Our primary objective was to comprehend the unique characteristics of each device and its corresponding traffic patterns. Upon meticulous examination of the dataset, we made notable observations. Each device exhibited varying quantities of benign traffic and was subject to different types of Gafgyt and Mirai attacks. For instance, Device 1 recorded 49,548 instances of benign traffic, along with significant occurrences of Gafgyt Combo (59,718), Gafgyt Junk (29,068), Gafgyt Scan (29,849), Gafgyt TCP (92,141), Gafgyt UDP (105,874), Mirai ACK (102,195), Mirai Scan (107,685), Mirai SYN (122,573), Mirai UDP (237,665), and Mirai UDP Plain (81,982) instances. Similarly, the remaining

devices displayed distinct patterns. Device 4 manifested a substantial number of benign instances (175,240) and various Gafgyt and Mirai attack instances. On the other hand, Device 9 exhibited a lower count of benign instances (19,528) but had higher occurrences of Gafgyt Combo (59,398) and Mirai UDP (157,084) instances. The table 1 shows the summary of the total counts of Benign and Malicious instances for each device.

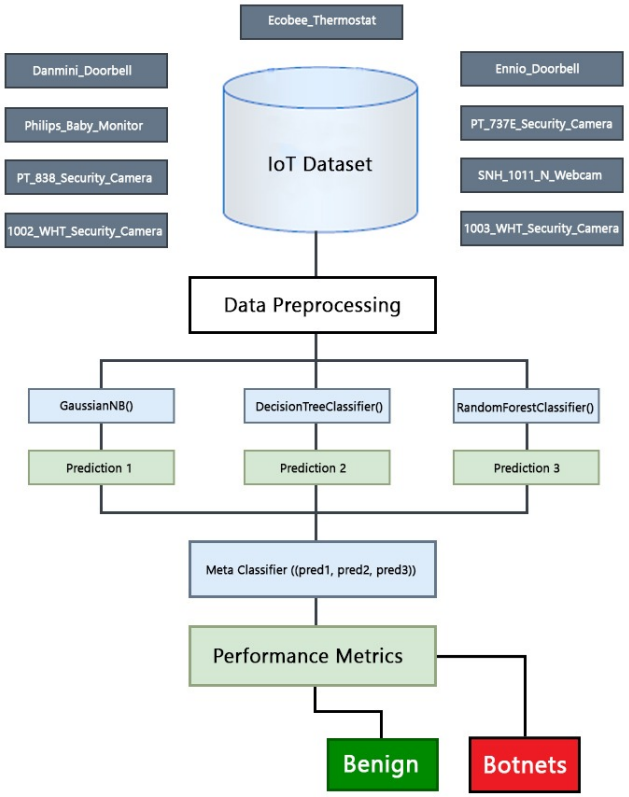


Figure 1: Graphical Overview of Dataset, Preprocessing Steps, and Adopted Methodologies

The N-BaIoT dataset is a key resource for our research activities in this study. Each dataset contains an array of Botnet Attacks, spanning a variety of approaches to attack, and is made up of data from nine unique IoT devices. These many assault types, however, have been grouped into two primary groups for the investigation's purposes: botnets and normal information. This consolidation makes it possible to analyse the dataset extensively, enabling an intensive investigation of the effectiveness and reliability of our recommended approach for recognising and decreasing the cumulative threats posed by Botnets.

The analysis of this dataset enabled us to gain valuable insights into the traffic patterns exhibited by different IoT devices and the prevalence of specific attack types. These insights are pivotal in devising effective security measures and enhancing the resilience of IoT systems against potential threats. To facilitate better comprehension, Figure 2 presents a bar chart that illustrates the distribution of traffic instances across the nine devices in the dataset. Each device is represented by a distinct color, and the height of the bars corresponds to the number of instances. This visual representation enables an overview of the overall traffic patterns observed in the dataset. Additionally, Figure 3 showcases a 3D scatter plot that provides a three-dimensional representation of the dataset, highlighting data from all nine devices.

3.1.1 Data Preprocessing

During this research study, we conducted essential data preprocessing on the N-BaIoT Dataset prior to model training. The Data Preprocessing phase serves as a fundamental groundwork. This section provides a comprehensive elucidation of the procedures employed on the data to render it suitable for model training. This preprocessing phase encompasses crucial steps including class balancing, data segmentation, and feature selection.

IoT botnet detection demands pristine data quality, making data standardization imperative. With each feature x_i having its own mean μ_i and standard deviation σ_i, the transformation to a standardized feature $x_{\text{scaled},i}$ is executed according to Equation 1.

$$x_{\text{scaled},i} = \frac{x_i - \mu_i}{\sigma_i} \tag{1}$$

This standardization endeavors to establish a uniform scale across all features, thereby facilitating the convergence of model training processes.

Addressing the challenge of class imbalance is a pivotal concern. To tackle this, strategic resampling is employed, culminating in a balanced dataset $X_{\text{balanced}}, y_{\text{balanced}}$. The resampling methodology encompasses the manipulation of either the minority or majority class, with the procedure mathematically encapsulated as Equation 2:

$$X_{\text{bal.}}, y_{\text{bal.}} = \text{resam.}(X_{\text{scaled}}, y, \text{stratify} = y, \text{ran.\_state} = 42) \tag{2}$$

Where
bal.=balanced
ran.=random
resam.= resample

This meticulous resampling mitigates the pitfalls of class imbalance, thus guaranteeing a representative training dataset.

Data partitioning plays a pivotal role in our data preprocessing strategy, strategically allocating the balanced dataset into subsets for training, testing, and validation purposes. This process ensures that our model is developed and assessed on distinct data samples, reflecting real-world scenarios. In this phase, we partition our dataset into three sets: training, testing, and validation, each with its respective purpose.

To achieve this, we leverage the proportions $\text{Proportion}_{\text{train}}$, $\text{Proportion}_{\text{test}}$, and $\text{Proportion}_{\text{val}}$, representing the ratios of samples assigned to each subset

Algorithm 1 IoT Botnet Attack Detection using Ensemble and Meta Ensemble ML Techniques

Require: Dataset D, Split Ratio SR, Thresholds *threshold_malicious*, *threshold_benign*

procedure IoT_Botnet_Detection(D, SR, *threshold_malicious*, *threshold_benign*)
  Load and Pre-process Dataset:
    Perform necessary pre-processing steps on D (e.g., cleaning, normalization).
  Split Dataset:
    Randomly shuffle D.
    Calculate $split\_index = \lfloor SR \times \text{size}(D) \rfloor$.
    $D_{\text{train}} = D[0 : split\_index]$
    $D_{\text{test}} = D[split\_index + 1 : \text{end}]$
  Train Ensemble Classifier:
    Initialize Voting Classifier with Naive Bayes, Decision Tree, and Random Forest classifiers.
    Train the Voting Classifier on D_{train}.
  Train Meta Ensemble Classifier:
    Initialize GB Classifier as the Meta Ensemble model.
    Train the Meta Ensemble model on the predictions of base classifiers from D_{train}.
  Evaluate Model Performance:
    Initialize evaluation metrics variables.
    for each testing sample (x_j, y_j) in D_{test} **do**
      Extract features from x_j.
      $y_{\text{hat\_ensemble}} = \text{VotingClassifier.predict}(x_j)$
      $y_{\text{hat\_meta}} = \text{MetaClassifier.predict}(x_j)$
      Update evaluation metrics based on predicted and actual labels.
  Malicious/Benign Detection:
    for each testing sample (x_k, y_k) in D_{test} **do**
      Calculate $p_{\text{malicious\_ensemble}} = \text{VotingClassifier.predict\_proba}(x_k)$
      Calculate $p_{\text{malicious\_meta}} = \text{MetaClassifier.predict\_proba}(x_k)$
      if $p_{\text{malicious\_ensemble}} > threshold\_malicious$ **or** $p_{\text{malicious\_meta}} > threshold\_malicious$ **then**
        Generate alert/notification for IoT botnet attack.
        Identify the attacking device or node using network analysis techniques.
        Block the device or node to mitigate the attack.
      if $p_{\text{malicious\_ensemble}} < threshold\_benign$ **and** $p_{\text{malicious\_meta}} < threshold\_benign$ **then**
        Proceed to the next phase or normal system operation.
  Output: Evaluation metrics, alerts/notifications, and blocked devices/nodes.
end procedure

relative to the total number of samples in the balanced dataset. These proportions are determined using the following equations:

$$\text{Proportion}_{\text{train}} = \frac{N_{\text{train}}}{N} \tag{3}$$

$$\text{Proportion}_{\text{test}} = \frac{N_{\text{test}}}{N} \tag{4}$$

$$\text{Proportion}_{\text{val}} = \frac{N_{\text{val}}}{N} \tag{5}$$

In Equations 4, 5, and 5, N_{train}, N_{test}, and N_{val} represent the respective numbers of samples in the training, testing, and validation subsets, and N signifies the total number of samples in the balanced dataset X_{balanced}.

This division ensures that the class distribution is preserved in each subset, maintaining the inherent characteristics of the original dataset. By adhering to these optimal proportions, we mitigate biases and ensure that our model's performance evaluation remains faithful to real-world conditions.

The process of feature selection is a pivotal aspect of preprocessing, wielding the potential to profoundly impact model performance. Feature relevance is meticulously assessed through the prism of Analysis of Variance (ANOVA). The F-statistic for a given feature x_i is outlined by:

$$F_i = \frac{\text{between-group variance}_i}{\text{within-group variance}_i}$$

The resultant F-statistics inform the curation of the top k features, subsequently steering the model training trajectory.

The indices corresponding to the selected features are housed within selected_feature_indices. The attendant feature names are captured by selected_feature_names, a key asset in feature identification during the forthcoming analysis and modeling stages.

In a bid to craft an IoT botnet detection model of exceptional prowess, this intricate data preprocessing serves as the bedrock, priming the dataset for the upcoming stages of model development and evaluation.

3.2 Performance Metrics Used in this Research

Accuracy: Accuracy is a fundamental metric that measures the overall correctness of our model's predictions. It is calculated using the formula:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{6}$$

where: TP represents the number of correctly identified botnet attacks. TN represents the number of correctly identified non-botnet traffic flows. FP represents the number of non-botnet traffic flows incorrectly classified as botnet attacks. FN represents the number of actual botnet attacks incorrectly classified as non-botnet traffic flows.

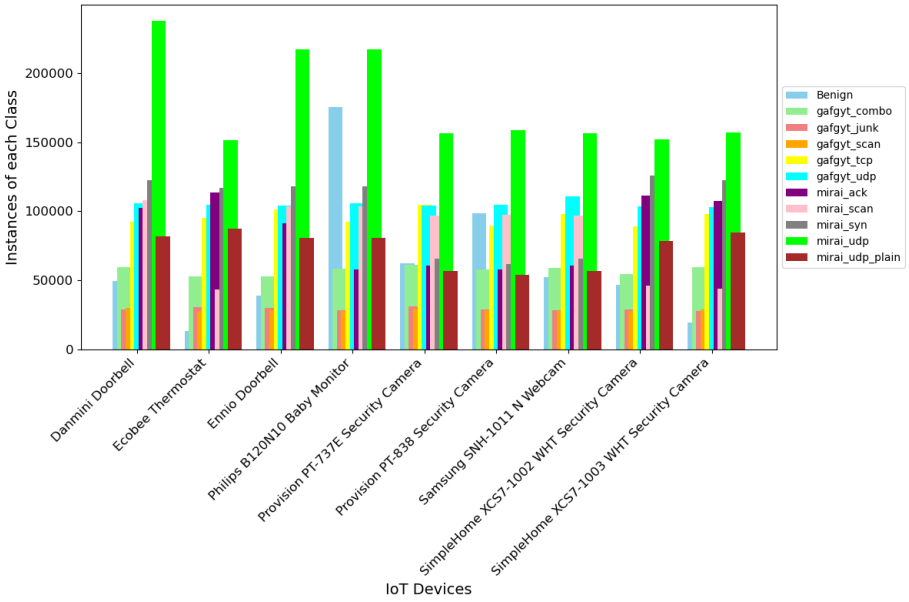


Figure 2: Classification of Nine IoT Devices in the Dataset

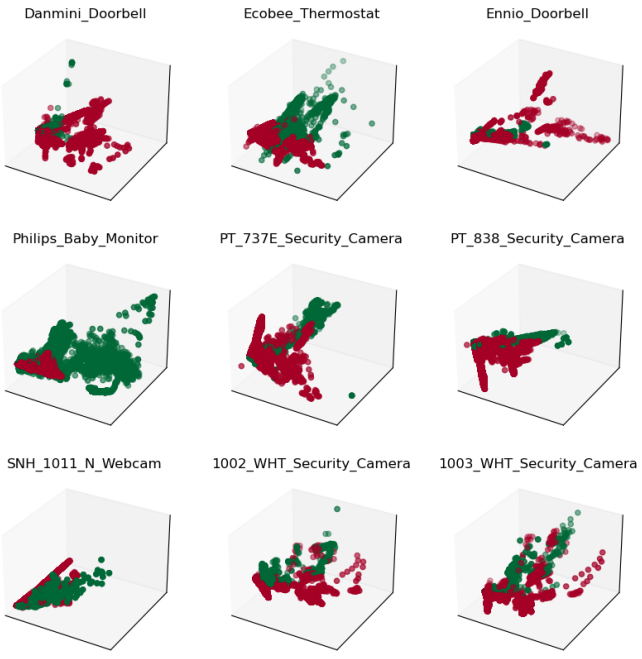


Figure 3: 3D Scatter Plot of Dataset

Precision: Precision measures the accuracy of our model's positive predictions, specifically its ability to correctly identify botnet traffic. It is calculated as:

$$Precision = \frac{TP}{TP + FP} \tag{7}$$

Recall: Recall, also known as True Positive Rate or Sensitivity, assesses our model's capability to identify all instances of botnet attacks in the network traffic. It is computed using the formula:

$$Recall = \frac{TP}{TP + FN} \tag{8}$$

F1 Score: F1 Score is the harmonic mean of Precision and Recall, providing a balanced assessment of both metrics. It takes into account both false positives and false negatives, making it valuable in detecting botnet attacks effectively. The formula for F1 Score is:

$$F1Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{9}$$

These evaluation indicators are crucial for assessing how well our botnet attack detection strategy is performing. We can make sure that our model accurately and consistently identifies botnet attacks while reducing the risk of false positives and false negatives in the IoT network by analyzing Accuracy, Precision, Recall, and F1 Score.

3.3 Voting Classifier in Ensemble Learning

In this research endeavor, we propose an ensemble model that amalgamates the unique advantages of Naive Bayes,

CART (Classification and Regression Trees), and Random Forest algorithms, with the aim of enhancing the accuracy and robustness of classification tasks. Leveraging the versatile capabilities of the Voting Classifier from the scikit-learn library, we unify these base classifiers into a cohesive model, proficiently handling diverse data types and capturing varied patterns.

$$\begin{aligned}\text{classifier\_model} = \text{VotingClassifier}\\ (\text{estimators} = [('nb', \text{GaussianNB}()),\\ ('dt', \text{DecisionTreeClassifier}()),\\ ('rf', \text{RandomForestClassifier}())],\\ \text{voting} =' hard')\end{aligned} \quad (10)$$

The Voting Classifier employs a majority voting strategy to aggregate predictions, where the final class label is determined based on the label with the highest aggregate score. Our ensemble model's effectiveness is demonstrated through extensive experiments on the N-BaIoT dataset, revealing superior classification accuracy and overall performance compared to using a single classifier in isolation. We conduct comprehensive evaluations of the ensemble models, assessing various metrics such as accuracy, precision, recall, and F1 score. These assessments provide valuable insights into the model's classification capabilities, including its ability to cope with different class distributions and imbalances. By fusing the strengths of Naive Bayes, CART, and Random Forest algorithms within the Voting Classifier framework, our proposed ensemble approach offers improved classification performance, resulting in a more reliable and adept model for IoT botnet attack detection. Algorithm 2 elucidates the procedures involved in training and testing our Ensemble Model. It comprehensively outlines the sequential steps undertaken to establish and evaluate the performance of our ensemble model.

Algorithm 2 Voting Classifier with Naive Bayes, Decision Tree, and Random Forest

Require: Dataset D, Number of base classifiers N

procedure VotingClassifier(D, N)
  Load and Pre-process Dataset:
    Perform necessary pre-processing steps on D (e.g., cleaning, normalization).
  Initialize Base Classifiers:
    Initialize Naive Bayes classifier NB, Decision Tree classifier DT, and Random Forest classifier RF.
  Train Base Classifiers:
    for i **in** 1 to N **do**
      Randomly select a subset D_i from D.
      Train NB on D_i, Train DT on D_i, Train RF on D_i.
  Voting Process:
    for each testing sample x_j in D **do**
      Predict class probabilities $P_{NB}(x_j)$, $P_{DT}(x_j)$, $P_{RF}(x_j)$.
      Calculate weighted average probabilities $P_j = \frac{1}{3}(P_{NB}(x_j) + P_{DT}(x_j) + P_{RF}(x_j))$.
  Output: Ensemble predictions P_j for all testing samples in D.
end procedure

3.4 Meta Ensemble Strategy

To implement the Meta Ensemble approach, we meticulously selected three distinct base classifiers: Naive Bayes, Decision Trees, and Random Forest, each contributing its unique algorithmic strengths to the ensemble. These base classifiers were trained on the same dataset, generating independent predictions. Subsequently, we leveraged the gradient-boosted classifier, known for its excellent performance, to train a meta-classifier using the predictions of the base classifiers as features. In the context of our research focused on the detection of IoT botnet attacks using our ensemble model, we have incorporated crucial mathematical equations that underpin our methodology.

$$F(x) = \sum_{i=1}^{M} w_i \cdot f_i(x) \quad (11)$$

Equation 11 captures the core principle of our ensemble model, vital for IoT botnet attack detection. In this equation, M represents the set of base models in the ensemble, w_i signifies the adaptive weight assigned to the ith base model's prediction $f_i(x)$ for the input x. The ensemble prediction $F(x)$ is computed by summing the weighted predictions, collectively forming our model's holistic assessment of the input's classification.

$$\text{argmin}_w \sum_{j=1}^{N} (y_j - F(x_j))^2 \quad (12)$$

Equation 12 embodies the weight learning process integral to our ensemble model. In the context of IoT botnet attack detection, y_j represents the true label of the jth instance in the training dataset. The objective is to minimize the squared difference between the true label and the ensemble prediction $F(x_j)$, thereby iteratively determining optimal weights w that facilitate accurate aggregation of base model predictions.

$$P(Y = c|x) = \sum_{i=1}^{M} w_i \cdot P_i(Y = c|x) \quad (13)$$

Equation 13 outlines the amalgamation of class probabilities for a given input x in the context of our ensemble model's operation. The ith base model's estimated probabilities $P_i(Y = c|x)$ are weighted by w_i to compute the final probability $P(Y = c|x)$ for our ensemble model. This comprehensive probability calculation empowers our model to make well-informed decisions when classifying IoT botnet attacks.

These equations collectively define the underlying mechanics of our Meta Ensemble learning methodology, forming the cornerstone of our ensemble model's efficacy

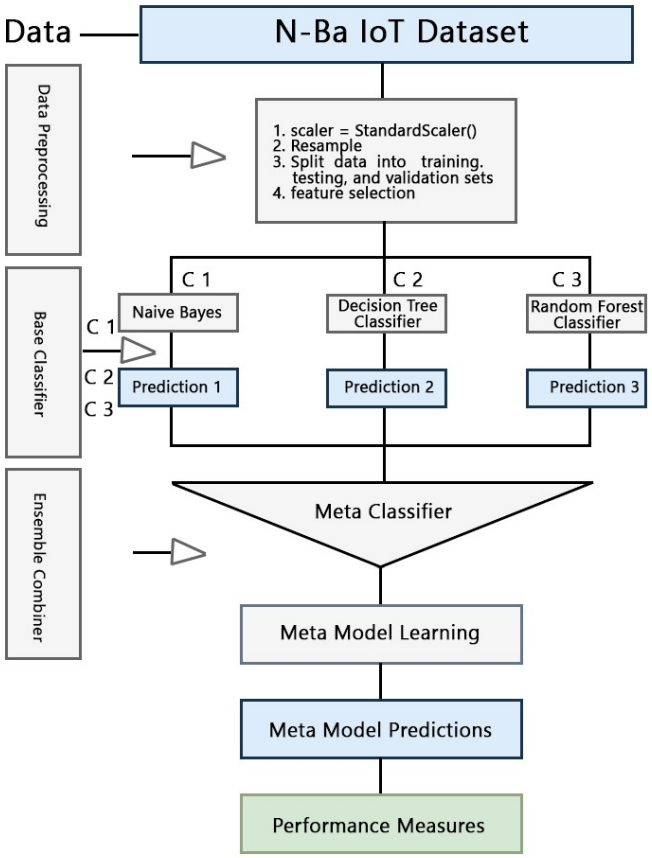


Figure 4: Meta Ensemble Classifier Methodology

in identifying and classifying IoT botnet attacks with unparalleled precision. Through meticulous weight learning and probabilistic aggregation, our model excels in its mission to contribute significantly to the domain of IoT security and the mitigation of botnet attacks.

The meta-classifier adeptly learned how to amalgamate insights from each classifier, resulting in more accurate and reliable predictions as shown in Figure. The Meta Ensemble approach bestowed several advantages to our study. By harnessing the diversity of base classifiers, we adeptly addressed challenges arising from complex IoT botnet detection scenarios. This diversity enabled us to emphasize the strengths of individual classifiers while mitigating their weaknesses, ultimately leading to improved generalization and overall performance. Techniques used for Meta Classifier are shown in Figure 4. Algorithm 3 elucidates the procedures involved in the training and testing of our Meta Ensemble Model. It comprehensively outlines the sequential steps undertaken to establish and evaluate the performance of our Meta ensemble model.

Our methodology includes an extensive strategy that effectively identifies and mitigates IoT botnet attacks via utilizing ensemble and meta-ensemble strategies. We have carefully explained how to apply these innovative techniques, supported by a solid grasp of the nuances of the N-BaIoT dataset. The adaptability and precision of our framework are further enhanced by making use of an array of basic classifiers. The empirical findings that demonstrate the significant contributions and breakthroughs our technique provides to the field of IoT security will be presented as we move into the results section. We illustrate the practical practicality and effectiveness of our suggested ensemble and meta-ensemble solutions in enhancing the security of IoT environments against changing botnet threats through exhaustive research and insightful analyses.

Algorithm 3 Meta Ensemble Classifier with Naive Bayes, Decision Tree, and Random Forest

Require: Dataset D, Number of base classifiers N

procedure MetaEnsembleClassifier(D, N)
  Load and Pre-process Dataset:
    Perform necessary pre-processing steps on D (e.g., cleaning, normalization).
  Initialize Base Classifiers:
    Initialize Naive Bayes classifier NB, Decision Tree classifier DT, and Random Forest classifier RF.
  Train Base Classifiers:
    for i **in** 1 to N **do**
      Randomly select a subset D_i from D.
      Train NB on D_i, Train DT on D_i, Train RF on D_i.
  Collect Base Classifier Predictions:
    Initialize empty arrays $Pred_{NB}$, $Pred_{DT}$, $Pred_{RF}$.
    for each testing sample x_j in D **do**
      Predict class probabilities $P_{NB}(x_j)$, $P_{DT}(x_j)$, $P_{RF}(x_j)$.
      Append $P_{NB}(x_j)$ to $Pred_{NB}$, Append $P_{DT}(x_j)$ to $Pred_{DT}$, Append $P_{RF}(x_j)$ to $Pred_{RF}$.
  Train Meta Classifier:
    Initialize Meta Classifier MC (GradientBoostingClassifier).
    Train MC on $Pred_{NB}$, $Pred_{DT}$, $Pred_{RF}$ to learn the optimal combination.
  Meta Ensemble Prediction:
    for each testing sample x_j in D **do**
      Predict class probabilities $P_{NB}(x_j)$, $P_{DT}(x_j)$, $P_{RF}(x_j)$.
      Append $P_{NB}(x_j)$ to $Pred_{NB}$, Append $P_{DT}(x_j)$ to $Pred_{DT}$, Append $P_{RF}(x_j)$ to $Pred_{RF}$.
    Combine base classifier predictions using MC to obtain the final ensemble prediction.
  Output: Ensemble predictions for all testing samples in D.
end procedure

4 Results and Analysis

The comprehensive evaluation of our proposed ensemble and Meta ensemble models across various IoT device datasets yields insightful outcomes, shedding light on their effectiveness in botnet detection. Among the individual classifiers, the Naive Bayes classifier showcases commendable performance across most datasets, particularly excelling on devices 1, 7, and 9, where it achieves remarkable accuracy, precision, recall, and F1 scores. However, challenges are encountered with certain devices, notably 3 and 8, which exhibit relatively slower performance. Notably, device 3 presents a trade-off between high recall and lower accuracy, indicative of its proficiency in identifying botnet instances while potentially leading to a higher rate of false positives.

The Voting Classifier, amalgamating predictions from Naive Bayes, Decision Trees, and Random Forests, emerges as a robust contender in IoT botnet detection, surpassing individual base classifiers in predictive prowess. Notably, devices 1, 3, 4, 7, and 9 demonstrate exceptional performance across various metrics, with accuracy ranging from 0.999920 to 0.999986, precision from 0.999489 to 1.000000, recall from 0.999524 to 0.999871, and F1 Score from 0.999541 to 0.999936. These devices consistently exhibit near-perfect accuracy, precision, recall, and F1 scores, underscoring the Voting Classifier's efficacy in botnet detection across a diverse range of IoT devices. However, device 5 displays slightly diminished performance metrics, with accuracy at 0.985928 and recall at 0.814753, suggesting potential areas for improvement. Despite this, the Voting Classifier strategically leverages the strengths of diverse classifiers, enhancing overall detection accuracy. Detailed insights into the Voting Classifier's performance metrics are provided in Table 3, offering a comprehensive overview of its performance on each device dataset. These findings are further complemented by visual representations in Figure 6, providing a clear depiction of the Voting Classifier's performance across the evaluated devices. In Figure 5, the learning curve of the proposed Ensemble techniques is displayed. Notably, the model exhibits high performance, with both training scores and cross-validation scores indicating substantial improvement. Moreover, there are no discernible issues of overfitting or underfitting present in the models

Table 3: Performance of Ensemble Classifier

Device	Accuracy	Precision	Recall	F1 Score
1	0.999956	0.999489	0.999592	0.999541
2	0.995741	0.999501	0.738025	0.849089
3	0.999986	1.000000	0.999871	0.999936
4	0.999950	0.999859	0.999831	0.999845
5	0.985928	0.999707	0.814753	0.897803
6	0.982238	0.999880	0.848873	0.918210
7	0.999920	0.999905	0.999524	0.999714
8	0.999867	0.999036	0.998501	0.998768
9	0.999912	0.997688	0.998457	0.998072

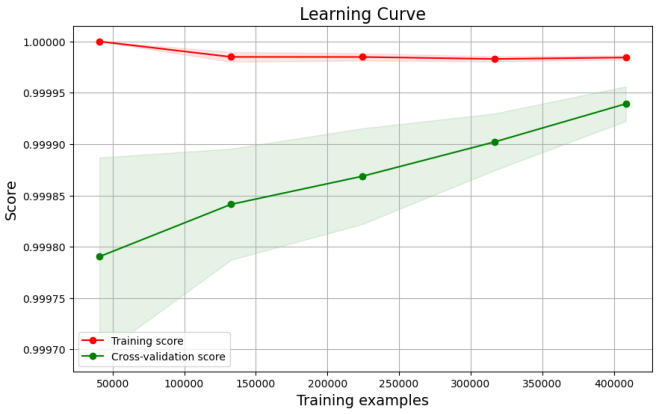


Figure 5: Learning Curve of Proposed Ensemble Techniques

The Meta Ensemble classifier, anchored by the powerful GradientBoostingClassifier, emerges as a standout performer in botnet detection by harnessing predictions from base classifiers. Notably, it achieves remarkable results across all IoT device datasets, attributed to the inclusion of an additional class that elevates overall accuracy. However, the Meta Classifier entails a trade-off between performance enhancement and increased computational demands, manifested in a slightly longer training period. A detailed breakdown of the Meta Classifier's performance metrics is provided in Table 4, accompanied by graphical illustrations in Figure 7. The Meta Ensemble model's prowess in enhancing detection across diverse IoT devices substantiates its significance in the realm of IoT botnet detection.

Table 4: Performance of Meta Ensemble Classifier

Device	Accuracy	Precision	Recall	F1 Score
1	0.999931	0.999285	0.999285	0.999285
2	0.996693	0.998500	0.735814	0.847263
3	0.999859	0.999742	0.998969	0.999356
4	0.999900	0.999634	0.999747	0.999691
5	0.987017	0.998144	0.813241	0.896255
6	0.987078	0.998619	0.846025	0.916010
7	0.999800	0.999049	0.999524	0.999286
8	0.999745	0.996899	0.998394	0.997646
9	0.999888	0.997685	0.997429	0.997557

The culmination of our meticulous classifier evaluations reveals the emergence of the Meta Ensemble classifier as a standout performer. Bolstered by the robust GradientBoostingClassifier, it showcases elevated prowess in botnet detection by integrating predictions derived from foundational classifiers. This astute approach ushers in an era of unparalleled performance across the spectrum of IoT devices constituting the N-BaIoT dataset. Notably, devices 1, 3, 4, 7, and 9 exhibit exceptional results across various metrics. These devices consistently demonstrate high accuracy, precision, recall, and F1 Score, ranging from 0.999800 to 0.999931 for accuracy, 0.999285 to

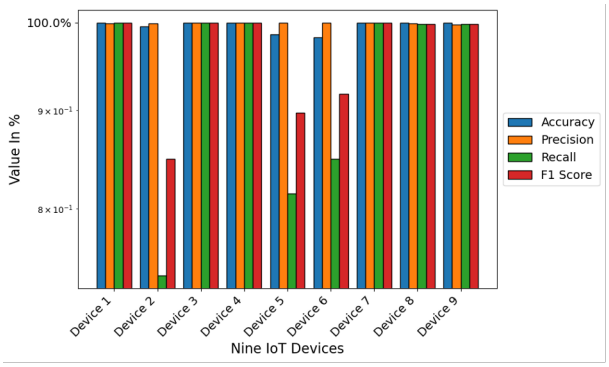


Figure 6: Comparison of Voting Classifier Performance Metrics across all 9 devices

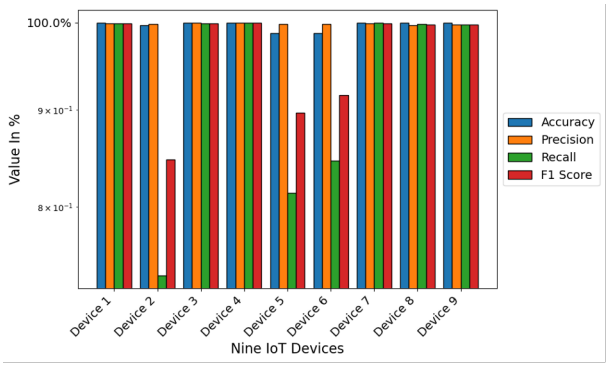


Figure 7: Comparison of Meta Ensemble Classifier Performance Metrics across all 9 devices

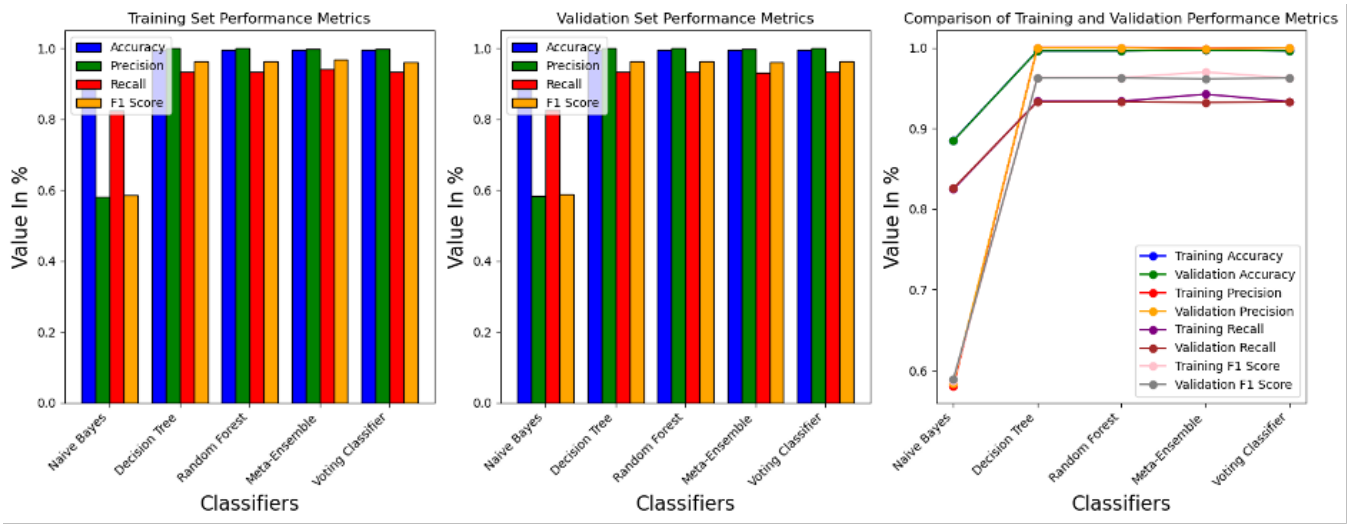


Figure 8: Comparison of Training and Testing Set Performance across all Classifiers

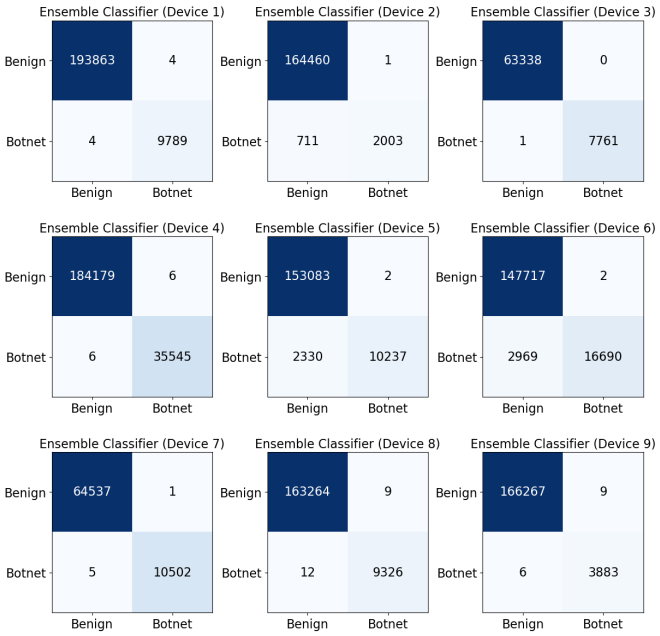


Figure 9: Confusion matrices illustrating the performance of the proposed ensemble classifier across all 9 devices

0.999742 for precision, 0.997429 to 0.999969 for recall, and 0.997557 to 0.999691 for F1 Score.

However, device 5 shows slightly lower performance metrics, with accuracy at 0.987017 and recall at 0.813241, indicating a potential area for improvement. This observation underscores the significance of evaluating the trade-offs between performance enhancement and other factors in practical deployment scenarios. Further analysis of the Meta Ensemble classifier's results highlights its robustness across various IoT devices, emphasizing its potential efficacy in real-world botnet detection applications.

The synthesis of our results, encapsulated in Tables 3 and 4, along with the vivid illustrations found in Figures 6 and 7, underscores the significance of ensemble

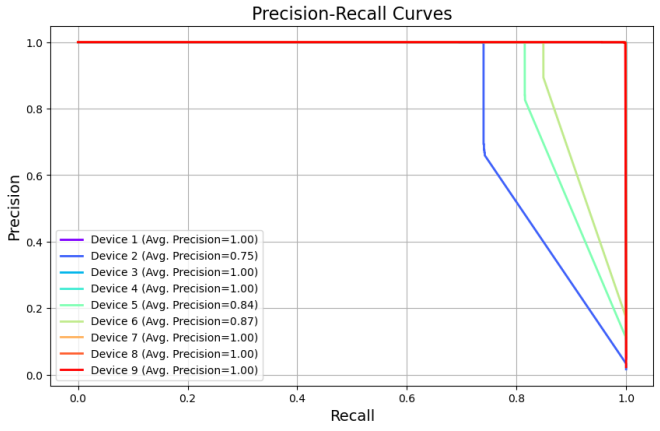


Figure 10: Device-specific precision-recall analysis: Curves for 9 devices with the proposed classifier

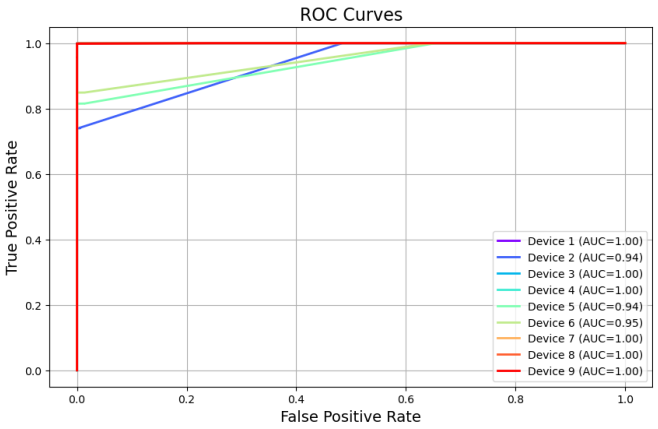


Figure 11: Device-specific ROC analysis: Comparative curves for 9 devices

methodologies in IoT botnet detection, with profound implications for real-world deployment. Our evaluation extends beyond traditional metrics, encompassing precision-recall curves and ROC curves to assess classifier performance comprehensively. Precision-recall curves offer insights into precision and recall trade-offs across different decision thresholds, while ROC curves facilitate comparisons of true positive and false positive rates, aiding in assessing classifier efficacy. The performance on training and testing sets, as depicted in Figure 8, provides a thorough review of generalization potential. Additionally, precision and recall performance across IoT devices (Figure 10) and device-specific ROC analysis (Figure 11) offer nuanced insights into the robustness of our approach across various scenarios, contributing to a comprehensive understanding of classifier efficacy.

The thorough evaluation of our proposed ensemble and Meta ensemble models highlights their efficacy in IoT botnet detection across diverse device datasets. While the Naive Bayes classifier demonstrates commendable performance on specific devices, the Voting Classifier and Meta Ensemble classifier emerge as robust contenders, leveraging diverse base classifiers to enhance detection accuracy. The Meta Ensemble classifier, particularly bolstered by the GradientBoostingClassifier, showcases unparalleled performance, albeit with slightly increased computational demands. Our findings underscore the significance of ensemble methodologies in IoT botnet detection, with implications for real-world deployment. Through comprehensive evaluation metrics and nuanced insights from precision-recall and ROC curves, our study provides a holistic understanding of classifier efficacy, paving the way for future advancements in IoT security.

5 Comparison with State-of-the-Art Studies

In this section, we conduct a thorough comparative analysis to evaluate the performance of our proposed ensemble

and meta-ensemble models for the detection of IoT botnet attacks. The comparison encompasses various contemporary methodologies in the field, utilizing a range of performance metrics to provide a nuanced and comprehensive understanding of our approach's effectiveness.

Table 5 presents a detailed breakdown of the performance metrics, allowing for a meticulous examination of our proposed model against recent research studies. The classification into "Ensemble" and "Meta Ensemble" methods, along with metrics such as accuracy, precision, recall, and F1 score, facilitates a comprehensive evaluation of the detection performance.

Starting with Sakthipriya *et al.* [18], which achieved an accuracy of 90.29% without incorporating ensemble methods, our proposed model showcases a substantial leap in accuracy, reaching an impressive 99.99% through the integration of both ensemble and meta-ensemble strategies. This significant improvement underscores the superior accuracy and heightened detection capabilities of our approach.

Examining Okur *et al.* [27], which achieved a commendable accuracy of 99.92% using ensemble methods but without meta-ensemble techniques, and Alkahtani *et al.* [28], which utilized ensemble methods to achieve an accuracy of 99.95%, we observe that our proposed model consistently outperforms these studies. The incorporation of meta-ensemble methods in our approach contributes to its exceptional accuracy of 99.99%, highlighting the efficacy of the combined strategies.

On the meta-ensemble side, Namoun *et al.* [29] focused on these approaches, achieving an accuracy of 88.77%. While our proposed model maintains higher accuracy, it is essential to note that our incorporation of meta-ensemble techniques contributes to a balanced and robust detection approach.

Moreover, considering Faysal *et al.* [30] and Cao *et al.* [31], both utilizing ensemble methods with accuracies of 99.91% and 99.37%, respectively, our proposed model consistently surpasses these studies in terms of accuracy, precision, recall, and F1 score. The integration of both ensemble and meta-ensemble methods in our model proves to be a distinctive advantage, contributing to its superior performance across all metrics.

In essence, our proposed ensemble and meta-ensemble models demonstrate consistent superiority over recent studies, affirming their robustness in addressing the intricate challenges associated with IoT botnet attack detection. The amalgamation of ensemble and meta-ensemble strategies emerges as a strategic choice, providing an effective and comprehensive solution to enhance IoT security.

6 Conclusion

This study underscores the paramount significance of ensemble and meta-ensemble models in IoT botnet detection. In direct comparison with contemporary studies, our investigation outshines accuracy, precision, and F1 scores. The infusion of machine learning techniques into our proposed model significantly fortifies the resilience and efficacy of botnet detection mechanisms. Our approach excels in achieving accuracy, with the ensemble classifier achieving an average of 99.29% across different devices (Table 3). Precision and recall are also commendable, with precision averaging 99.79% and recall averaging 91.96%. The meta-ensemble classifier performs similarly well, achieving an average accuracy of 99.74%, precision of 99.21%, and recall of 99.51% (Table 4).

These outcomes unequivocally highlight the potential of our approach in fortifying IoT security, providing a robust defense against the identification and mitigation of botnet threats. Despite a slightly increased training time, this trade-off is justified given the substantial enhancement in detection capabilities. In the face of ever-evolving IoT security challenges, this research stands as a pivotal contribution.

7 Limitations and Future Work

While our study demonstrates the effectiveness of ensemble and meta-ensemble models in IoT botnet attack detection, several limitations warrant consideration. Firstly, the performance evaluation relies on specific datasets and may not fully generalize to diverse IoT environments. Future research should explore the robustness of our models across various IoT device types and network configurations. Additionally, our study primarily focuses on supervised learning techniques, neglecting the potential of unsupervised or semi-supervised approaches. Investigating the integration of these methods could enhance the scalability and adaptability of our detection framework.

The computational resources required for training and deploying ensemble and meta-ensemble models may pose challenges in resource-constrained IoT environments. Future work should aim to optimize model architectures and algorithms to minimize resource consumption while maintaining detection efficacy. Additionally, the interpretability of ensemble and meta-ensemble models remains a concern, particularly in critical IoT applications where transparent decision-making is essential. Developing techniques for model explainability and uncertainty quantification could enhance the trustworthiness and adoption of our approach.

Our study primarily evaluates the performance of our models under benign conditions and may not adequately address adversarial attacks or dynamic IoT environments. Future research should explore the resilience of ensemble and meta-ensemble models against adversarial manipulations and evolving attack strategies. Additionally, incorporating real-time monitoring and adaptive learning mechanisms could enhance the responsiveness and adaptability of our detection framework in dynamic IoT environments.

While our study represents a significant advancement in IoT botnet attack detection, there exist several avenues

Table 5: Proposed Model vs. Contemporary Research

Study	Ensemble	Meta Ensemble	Accuracy	Prec.	Recall	F1
Sakthipriya et Al. (2023) [18]	N0	N0	90.29	90.13	91.38	88.67
Okur (2023) et Al. [27]	YES	NO	99.92	99.55	97.23	96.44
Alkahtani et Al. (2021) [28]	YES	NO	99.95	98.00	95.00	95.44
Namoun et Al. (2023) [29]	NO	YES	88.77	87.66	88.60	87.56
Faysal et Al. (2022) [30]	YES	NO	99.91	99.20	97.33	94.94
Cao et Al. (2023) [31]	YES	YES	99.37	95.70	92.51	95.57
Proposed study	YES	YES	99.99	99.99	93.99	96.99

for future research to address the identified limitations. By exploring these areas, researchers can further improve the effectiveness, efficiency, and robustness of ensemble and meta-ensemble models in safeguarding IoT ecosystems against emerging cyber threats.

Acknowledgments

The work presented in this paper has been supported by the Beijing Natural Science Foundation (No. IS23054). The authors gratefully acknowledge the anonymous reviewers for their valuable comments.

References

[1] T. Sugiura, K. Yamamura, Y. Watanabe, S. Yamakiri, and N. Nakano, "Circuits and devices for standalone large-scale integration (lsi) chips and internet of things (iot) applications: A review," *Chip*, p. 100048, 2023.

[2] A. Nazir, J. He, N. Zhu, A. Wajahat, F. Ullah, S. Qureshi, X. Ma, and M. S. Pathan, "Collaborative threat intelligence: Enhancing iot security through blockchain and machine learning integration," *Journal of King Saud University-Computer and Information Sciences*, p. 101939, 2024.

[3] A. H. El-Kady, S. Halim, M. M. El-Halwagi, and F. Khan, "Analysis of safety and security challenges and opportunities related to cyber-physical systems," *Process Safety and Environmental Protection*, 2023.

[4] A. Nazir, J. He, N. Zhu, M. S. Anwar, and M. S. Pathan, "Enhancing IoT security: a collaborative framework integrating federated learning, dense neural networks, and blockchain," *Cluster Computing*, 2024. [Online]. Available: https://doi.org/10.1007/s10586-024-04436-0

[5] A. Nazir, J. He, N. Zhu, A. Wajahat, X. Ma, F. Ullah, S. Qureshi, and M. S. Pathan, "Advancing iot security: A systematic review of machine learning approaches for the detection of iot botnets," *Journal of King Saud University-Computer and Information Sciences*, p. 101820, 2023.

[6] S. Dange and M. Chatterjee, "Iot botnet: The largest threat to the iot network," in *Data Communication and Networks: Proceedings of GUCON 2019*. Springer Singapore, 2019, pp. 137–157.

[7] M. Wazzan, D. Algazzawi, O. Bamasaq, A. Albeshri, and L. Cheng, "Internet of things botnet detection approaches: Analysis and recommendations for future research," *Applied Sciences*, vol. 11, no. 12, p. 5713, 2021.

[8] A. Heidari and M. A. Jabraeil Jamali, "Internet of things intrusion detection systems: A comprehensive review and future directions," *Cluster Computing*, pp. 1–28, 2022.

[9] A. Nazir, J. He, N. Zhu, S. S. Qureshi, S. U. Qureshi, F. Ullah, A. Wajahat, and M. S. Pathan, "A deep learning-based novel hybrid cnn-lstm architecture for efficient detection of threats in the iot ecosystem," *Ain Shams Engineering Journal*, p. 102777, 2024.

[10] Y. Meidan, M. Bohadana, Y. Mathov, Y. Mirsky, D. Breitenbacher, A. Asaf, and A. Shabtai, "Detection of iot botnet attacks (n-baiot)," *UCI Machine Learning Repository*, 2018. [Online]. Available: https://doi.org/10.24432/C5RC8J

[11] D. P. Hostiadi and T. Ahmad, "Hybrid model for bot group activity detection using similarity and correlation approaches based on network traffic flows analysis," *Journal of King Saud University-Computer and Information Sciences*, vol. 34, no. 7, pp. 4219–4232, 2022.

[12] M. N. Chowdhury, K. Ferens, and M. Ferens, "Network intrusion detection using machine learning," in *Proceedings of the International Conference on Security and Management (SAM)*, 2016, p. 30.

[13] J. Lee and H. Kim, "Security and privacy challenges in the internet of things [security and privacy matters]," *IEEE Consumer Electronics Magazine*, vol. 6, no. 3, pp. 134–136, 2017.

[14] I. Indre and C. Lemnaru, "Detection and prevention system against cyber-attacks and botnet malware for information systems and internet of things," in *Proceedings of the 2016 IEEE 12th International Conference on Intelligent Computer Communication and Processing (ICCP)*, 2016, pp. 175–182.

[15] H. Asgharzadeh, A. Ghaffari, M. Masdari, and F. S. Gharehchopogh, "Anomaly-based intrusion detection system in the internet of things using a convolutional neural network and multi-objective enhanced

capuchin search algorithm," *Journal of Parallel and Distributed Computing*, vol. 175, pp. 1–21, 2023.

[16] M. Bhavsar, K. Roy, J. Kelly, and O. Olusola, "Anomaly-based intrusion detection system for iot application," *Discover Internet of Things*, vol. 3, no. 1, p. 5, 2023.

[17] F. Suthar, N. Patel, and S. Khanna, "A signature-based botnet (emotet) detection mechanism," *Int. J. Eng. Trends Technol*, vol. 70, no. 5, pp. 185–193, 2022.

[18] N. Sakthipriya, V. Govindasamy, and V. Akila, "A comparative analysis of various dimensionality reduction techniques on n-baiot dataset for iot botnet detection," in *2023 2nd International Conference on Paradigm Shifts in Communications Embedded Systems, Machine Learning and Signal Processing (PCEMS)*. IEEE, 2023, pp. 1–6.

[19] M. Umair, W.-H. Tan, and Y.-L. Foo, "Efficient malware classification with spiking neural networks: A case study on n-baiot dataset," in *2023 Fourteenth International Conference on Ubiquitous and Future Networks (ICUFN)*. IEEE, 2023, pp. 231–236.

[20] M. AL-Akhras, A. Alshunaybir, H. Omar, and S. Alhazmi, "Botnet attacks detection in iot environment using machine learning techniques," *International Journal of Data and Network Science*, vol. 7, no. 4, pp. 1683–1706, 2023.

[21] U. Garg, V. Kaushik, A. Panwar, and N. Gupta, "Analysis of machine learning algorithms for iot botnet," in *2021 2nd International Conference for Emerging Technology (INCET)*. IEEE, 2021, pp. 1–5.

[22] A. Ahmed, "Machine learning based iot-botnet attack detection using real-time heterogeneous data," in *2022 International Conference on Electrical, Computer and Energy Technologies (ICECET)*. IEEE, 2022, pp. 1–6.

[23] S. Saif, N. Yasmin, and S. Biswas, "Feature engineering based performance analysis of ml and dl algorithms for botnet attack detection in iomt," *International Journal of System Assurance Engineering and Management*, vol. 14, no. Suppl 1, pp. 512–522, 2023.

[24] A. Padmashree and M. Krishnamoorthi, "Comparative analysis of dos attack detection in kdd cup99 using machine learning classifier algorithms," in *2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC)*. IEEE, 2022, pp. 1570–1573.

[25] M. Motylinski, . MacDermott, F. Iqbal, and B. Shah, "A gpu-based machine learning approach for detection of botnet attacks," *Computers and Security*, vol. 123, p. 102918, 2022.

[26] D. Krishnan and P. Babu, "An adaptive weighted voting classifier for botnet detection in internet of things," in *2021 IEEE 18th India Council International Conference (INDICON)*. IEEE, 2021, pp. 1–6.

[27] C. Okur, A. Orman, and M. Dener, "Ddos intrusion detection with machine learning models: N-baiot data set," in *4th International Conference on Artificial Intelligence and Applied Mathematics in Engineering. ICAIAME 2022. Engineering Cyber-Physical Systems and Critical Infrastructures*, vol. 7, 2023, pp. Springer, Cham.

[28] H. Alkahtani and T. H. Aldhyani, "Botnet attack detection by using cnn-lstm model for internet of things applications," *Security and Communication Networks*, pp. 1–23, 2021.

[29] A. Namoun, M. Humayun, O. BenRhouma, B. Hussein, A. Tufail, A. Alshanqiti, and W. Nawaz, "Service selection using an ensemble meta-learning classifier for students with disabilities," *Multimodal Technologies and Interaction*, vol. 7, no. 5, p. 42, 2023.

[30] J. Faysal, S. Mostafa, J. Tamanna, K. Mumenin, M. Arifin, M. Awal, A. Shome, and S. Mostafa, "Xgb-rf: A hybrid machine learning approach for iot intrusion detection," *Telecom*, vol. 3, no. 1, pp. 52–69, 2022.

[31] Y. Cao, Z. Wang, H. Ding, J. Zhang, and B. Li, "An intrusion detection system based on stacked ensemble learning for iot network," *Computers and Electrical Engineering*, vol. 110, p. 108836, 2023.

Biography

Xiangjun Ma completed his Bachelor's degree at Yantai University in 2006 and went on to earn a Master's degree from the College of Applied Mathematics and Physics at Beijing University of Technology in 2009. Currently, he is pursuing a doctoral degree at the School of Information Science and Technology, Beijing University of Technology. His research interests lie in cybersecurity, social network analysis, and IoT security. With a keen interest in exploring the intricacies of network security and addressing emerging challenges in securing the Internet of Things, he is dedicated to advancing knowledge in these areas.

Jingsha He received a bachelor's degree in computer science from Xi'an Jiaotong University, China, and the master's and Ph.D. degrees in computer engineering from the University of Maryland, College Park, MD, USA. He worked for several multinational companies in the USA, including IBM Corp., MCI Communications Corp., and Fujitsu Laboratories. He is currently a Professor at the Faculty of Information Technology, Beijing University of Technology(BJUT), Beijing. He has published more than ten articles. He holds 12 U.S. patents. Since August 2003, he has published over 300 papers in scholarly journals and international conferences. He also holds over 84 patents and 57 software copyrights in China and authored nine books. He was a principal investigator of more than 40 research and development projects. His research interests include information security, wireless

networks, and digital forensics.

Nafei Zhu received the B.S. and M.S. degrees from Central South University, China, in 2003 and 2006, respectively, and the Ph.D. degree in computer science and technology from the Beijing University of Technology, Beijing, China, in 2012. She was a Postdoctoral Research Fellow with the State Key Laboratory of Computer Science, Institute of Software, Chinese Academy of Sciences, Beijing, from 2015 to 2017. She is currently an Associate Professor with the Faculty of Information Technology, Beijing University of Technology. She has published over 20 research papers in scholarly journals and international conferences. Her research interests include information security and privacy, wireless communications, and network measurement.

Ahsan Nazir achieved his PhD in Software Engineering from Beijing University of Technology in 2024. His research interests include Cyber Security, IoT Security, Machine Learning, Deep Learning, Federated Learning, Blockchain, and Metaverse technologies. A prolific scholar, he has authored over 20 articles in prestigious journals and conferences.

Xiao Hu completed her Bachelor's degree in Computer Science and Technology from the Beijing University of Civil Engineering and Architecture in 2021. Subsequently, she pursued a Master's degree in Software Engineering from Beijing University of Technology, graduating in 2024. Her areas of research include Information Security, Blockchain, and Machine Learning..

Ahsan Wajahat received the B.S. and M.S. degrees in information technology from the Sindh Agriculture University, Pakistan, in 2012 and 2016, respectively. He is currently pursuing a Ph.D. degree at the Beijing University of Technology, Beijing, China. His research interests include machine learning, information security, forensic networks, and data mining. He has received awards and honors from the China Scholarship Council (CSC), China.

Faheem Ullah received the M.S degrees from the Xian Jiaotong University, China, in 2017. He is currently pursuing the Ph.D. degree with the Beijing University of Technology, Beijing, China. His research interests include information security, Blockchain and data mining. He has received awards and honors from the China Scholarship Council (CSC), China.

Sirajuddin Qureshi received his bachelor's degree in Computer Sciences from Quaid-e-Awam University of Engineering, Science & Technology, Pakistan. Afterwards, he pursued his Master's in Information Technology from Sindh Agricultural University Tandojam, Pakistan. Currently he is pursuing PhD in Information Technology at Beijing University of Technology, China. He has nine research publications to his credit as main author and co-author, which featured national and international journals and conferences. Sirajuddin's research areas includes but not limited to Network Forensics Analysis, Digital Forensics, Cyber security, Computer Networks and Network Security.

Research on Data Security and Privacy of Smart Grids

Min-Shiang Hwang[1,2], Yung-Ling Chang[3], Ko-Yu Lin[4], Cheng-Ying Yang[5], and Iuon-Chang Lin[4]
(Corresponding author: Iuon-Chang Lin)

Department of Computer Science & Information Engineering, Asia University[1]
Fintech and Blockchain Research Center, Asia University[2]
The Ph.D. Program in Artificial Intelligence, Asia University, Taiwan[3]
500, Lioufeng Rd., Wufeng, Taichung 41354, Taichung, Taiwan (R.O.C.)
Department of Management Information Systems, National Chung Hsing University[4]
Department of Computer Science, University of Taipei, Taipei, Taiwan[5]
Email: cyang@utaipei.edu.tw
(Received Dec. 4, 2023; Revised and Accepted July 21, 2024; First Online Aug. 31, 2024)

Abstract

Developing a smart grid is a trend nowadays and an indispensable basic necessity of life. In recent years, the electricity demand has risen rapidly due to economic development and quality of life improvement. The power grid in significant cities is unstable due to old equipment, and power outages occur frequently. The smart grid can adjust the grid's distribution strategy through the consumer's real-time electricity demand so that the consumer's electricity consumption data is always transmitted to the grid. Failure to protect the privacy and security of such information will expose consumers' habits and even make the grid unnecessarily wasteful. The smart grid's abnormality detection efficiency determines the overall grid's stability. Failure to detect anomalies or power theft in the grid promptly will not only lead to the selection of the wrong distribution strategy for the smart grid but also reduce the transmission efficiency of the grid. This research ensures the privacy and security of transmitted data to enhance the efficiency of abnormality detection in the smart grid. The research issues are divided into privacy of transmitted data, data security authentication, and anomaly detection efficiency.

Keywords: Authenticated Key Exchange; Data Privacy; Smart Grid; Smart Meter

1 Introduction

The power grid is a network system that connects the power supply and demand sides to transmit electricity. The power grid comprises power plants, substations, transmission, and distribution systems [36]. Businesses, government agencies, medical institutions, and households use a lot of electrical equipment every day. The regular operation of these devices depends on the stable power delivery from the power grid. Most countries' power grids lack flexibility, and the equipment is gradually getting older. Rising electricity consumption and uneven power have led to frequent power outages in recent years [2].

In recent years, solar photovoltaics have thrived. In addition to reducing the burden on the environment, they also provide a new outlet for power supply. However, even if there is power, insufficient feeders cannot be properly utilized. How to effectively utilize the feeder capacity will be a driving force—a significant issue in renewable energy power generation.

Smart Grid is the key to solving the above problems. The essential features of Smart Grid are to improve the overall system efficiency, grid resiliency, and self-healing capability. Grid resiliency refers to the ability to respond to power or outages. During this period, the power grid can quickly return to normal and continue to operate. This can be achieved by adding additional distributed power sources and integrating them into the grid when a power outage occurs. Self-healing allows the system to quickly identify grid faults, reduce times of power outages, help the grid recover and continue operating faster, and improve system performance. In traditional power grids, energy losses occur for several reasons, including power station failure or damaged transmission lines. Smart grids improve operations, reducing energy costs and improving system performance by using more efficient ways to transmit power. To this end, many smart devices are distributed throughout the smart grid to effectively manage power generation, transmission, distribution, and consumption. To effectively manage these smart devices, it is crucial to maintain the security and stability of the smart grid [4, 10].

This research takes the data transmitted in smart grids

as the main direction. It studies how to enhance the security and privacy of data to improve the abnormality detection capabilities and speed of smart grids. The problems to be solved are as follows:

1) Privacy issues in data transmission: Since the smart grid will determine the power configuration according to the consumers' usage needs, consumer data will be transmitted in the power grid. If this data is leaked to malicious users, knowing the user's living habits will lead to unnecessary energy losses in the power grid. The first topic of this research is to propose a solution to protect data privacy and use blockchain to design an architecture to ensure privacy.

2) Data exchange and verification correctness issues: Consumers' electricity consumption data will be collected and sent to suppliers through AMI (Advanced Metering Infrastructure), allowing suppliers to charge corresponding electricity bills. Suppose the data is tampered with or forged by malicious users during transmission. In that case, it will not only cause economic losses to the supplier but also make the smart grid information asymmetric and unable to distribute power effectively. This research is to study key exchange and digital signatures in the research issue to enhance data security in smart grids.

3) Data exchange and verification efficiency issues: Consumers' electricity consumption data will be collected and sent to the electricity supplier through AMI so that the supplier can charge the corresponding amount of electricity bills or conduct load analysis. However, when many AMI devices send statistical data to the supplier simultaneously, without an efficient verification mechanism, it is easy to cause network congestion and verification delays, affecting the accuracy of electricity bill calculation and power dispatching. This project is expected to study a batch verification solution for smart grids in the second year to reduce overall grid delays and ensure data accuracy.

4) The problem of the smart grid quickly detecting abnormal points: Since the smart grid is a decentralized power grid structure if abnormal points or power theft cannot be detected and eliminated immediately, it will not only reduce the overall grid transmission efficiency but also cause inconveniences and be a necessary waste of energy. The research is to study the deployment of blockchain smart contracts and deep learning methods to detect electricity theft in smart grids.

2 The Motivation

This research is based on a smart grid ensuring the security and privacy of transmitted data to improve smart grid anomaly detection. The following are the motivation and purpose of the research:

Motivation 1: Carbon emission issues are becoming increasingly important

In response to global climate change, carbon reduction and neutrality are gradually being taken seriously, with net zero emissions as the goal to prevent the earth from continuing to heat up. However, the current industry cannot wholly achieve net zero emissions, so it will use carbon neutrality. To control carbon emissions harmoniously, use methods such as planting trees or using renewable energy to offset the carbon emissions generated. Taiwan currently uses coal-fired power generation as a representative of high carbon emissions, but the current power generation profile in Taiwan is still based on coal-fired power generation. Coal and gas-fired power generation are the largest. Restricting their power generation will lead to insufficient power supply in Taiwan. It is essential to increase sustainable and clean renewable energy significantly. Therefore, this project will maximize the use of renewable energy through research on smart grids while also improving the efficiency of power generation and distribution, thereby avoiding energy waste and solving the problem of power shortage.

Motivation 2: Avoid leakage of users electricity consumption data

Since there are many sensors in the smart grid, consumers' power consumption data will be recorded and sent to the power supplier so that the supplier can adjust the configuration of power equipment based on this data. Suppose the privacy of this data is not ensured. In that case, it will lead to the exposure of consumers living habits because a lot of information can be learned from these electricity consumption data. For example, when consumers use electricity, they can roughly estimate the time when they are not at home or learn what consumers are doing through electricity consumption. This information is enough to leak privacy. Therefore, this project is expected to design a framework to ensure data privacy in smart grids and study solutions using blockchain and its privacy framework.

Motivation 3: Protect users' smart meters from attacks

A smart grid is an electric power infrastructure that uses information and communication technology to transmit and distribute electricity. One of its essential purposes is to accurately bill consumers in the power grid to allocate and manage electricity effectively. Smart meters play a role in this task. It plays an important role; its functions are (1) measuring energy consumption, (2) reporting energy consumption data to the control center, and (3) receiving electricity cost or control signals. The data collected through these smart meters can enable the smart grid to make

corresponding power distribution strategies. If these smart meters do not ensure their security, the data they send may be tampered with or even destroyed by malicious users. In addition to placing an additional burden on the power grid network, it will also reduce the overall security of the smart grid. Therefore, this research will study the data in the smart grid can ensure its security without affecting grid transmission efficiency.

Motivation 4: Enable smart grids to have the ability to monitor abnormalities in real-time
Energy losses in smart grids are divided into technical losses and non-technical losses. Usually, some technical losses will occur during electricity's transmission or transformation process. Non-technical losses are also known as energy theft, commonly known as electricity theft. The losses caused by these energy thefts cannot be underestimated. In addition to causing direct economic losses, these electricity thefts will also disrupt smart grids' power generation and management—transmission strategies. Practical and comprehensive energy theft detection becomes very important. In traditional power grids, detecting electricity theft usually requires regular inspections by professionals, which is time-consuming, laborious, and inefficient. However, smart grids can automatically collect energy consumption data from various places, which can not only solve the efficiency of traditional manual execution but also make it more efficient—real-time detection of anomalies [44].

3 Related Works

Traditional grids are no longer a solution for power transmission and distribution due to shortcomings such as prolonged power outages, energy storage issues, and high carbon emissions. A holistic solution is proposed to solve these challenges: the smart grid [49], which comprises advanced metering infrastructure equipment composed of smart meters, communication technology, meter data management systems, etc. [35]. These devices are regarded as a bridge between consumers and energy suppliers. Advanced metering devices collect energy consumption data in real-time. These time series data are transmitted to the meter data management system through commonly used fixed and public networks and then recorded. And analyze smart meter data to enable grid operators to optimize power supply and distribution [8]. This research studies the data security and privacy of smart grids to solve the anomaly detection problems of smart grids. It develops HAN (Home Area Networks) data privacy protection, grid data security certification, and rapid energy theft detection.

3.1 Related Works on Data Privacy for Smart Grids

A lot of electricity consumption data is circulating in the smart grid, most of which is collected from smart meters at the user end. These sensitive smart meter data can reflect the user's living habits and actual electricity consumption time [22, 32, 34], if these data are not handled properly, it may lead to the leakage of user privacy. However, unlike traditional data privacy, smart meter data privacy should consider the complex structure of the power sector, the legacy of closed-system energy technology, and legal and government restrictions [3]. How to provide adequate privacy protection is the current communication process priorities.

Smart meters usually collect the energy consumption of household appliances every 15 minutes. Then, the smart meters will compile the data and send it to the supplier as the basis for electricity billing. If a malicious user reads the energy consumption data from it, it may be possible to infer the consumption. Therefore, energy consumption data must be encrypted to protect security and privacy before the device transmits and leaves the HAN [2]. Each meter reports its energy consumption or uses data aggregation technology to combine energy consumption from the same area. Consumption data is summarized, and reports are sent to the control center in batches. The control center generates monthly bills and shows how much electricity is needed to maintain the electricity meters in the area [39]. The control center will continuously collect electricity consumption data, which must be encrypted or aggregation because if a malicious user changes the transmitted data, it will be easier to see through [29].

3.2 Related Works on Security Certification of Smart Grids

Since this research will ensure data security through key exchange and digital signatures, we will discuss these two parts below.

If the data in the smart grid is tampered with or manipulated by malicious users, it will impose additional load on the network [51]. Therefore, one of the critical requirements for smart grid security is data encryption, which requires mutual authentication and key exchange [33]. Tsai & Lo proposed an authentication scheme based on bilinear pairing [41]. Although it can hide the identity of smart meters to achieve anonymity, the bilinear pairing has a high computational cost and is unsuitable for smart grids with limited equipment resources. Mahmood *et al.* [31] proposed a key exchange scheme using elliptic curve encryption [7, 16]. Although it is more suitable for smart grids, it cannot satisfy anonymity. Although Abbasinezhad-Mood & Nikooghadam [1] allow smart meters to register with a trusted organization to provide anonymity, the registration phase requires a lot of computation. It is not practical to encrypt smart meter identities. Kumar *et al.* [24] proposed a lightweight authenti-

cation and key exchange protocol between smart meters and NAN gateways in smart grids. Their scheme used an elliptic curve encryption algorithm and AES to achieve encryption. These calculations all rely on the NAN gateway. The NAN gateway will require much memory space to complete this function. Although Kumar *et al.* believe that their solution is anonymous, it can be tracked by tracking transmissions between other meters in the NAN. The identity identification code is used to determine the transmission pipeline of a specific meter, so the anonymity effect cannot be achieved.

To reduce the computational burden and delay caused by signatures, the concept of execution first and verification later has been developed. This concept divides signature generation into two stages: the offline stage before the signed message is given and the offline stage for execution. They are calculated so that messages can be signed faster online. For example, the elliptic curve digital signature algorithm [19, 21, 45], Fiat-Shamir [12], Schnorr [37], RSA [5, 6, 17, 46], El-Gamal [9, 18, 25, 27, 30], and digital signature standards can all be calculated without knowing the message to be signed. Similarly, online/offline signing [11,38] can convert any ordinary signing scheme into a pre-computed scheme. In these solutions, the throughput is still affected by the high offline computing cost, and their end-to-end latency is also affected by the expensive computing cost in the online certification stage, resulting in their inability to be appropriately applied to smart grids.

3.3 Related Works on the Batch Verification of Smart Grids

Using the tree data structure as the verification scheme, Fouda *et al.* designed a lightweight message verification protocol based on Diffie-Hellman [13] to solve the mutual verification problem between the home area network Gateway and the building area network Gateway. Liu *et al.* [28] constructed a key management scheme based on a crucial graph for AMI to achieve secure communication. However, Wan *et al.* [42] found this solution vulnerable to denial of service attacks. So, they proposed an improved solution that combined an identity cryptosystem and an efficient key tree to achieve secure communication between smart meters and meter data management systems. Li *et al.* [26] designed a lightweight smart grid verification protocol based on the Merkel hash tree, considering smart meters with limited computing resources. However, the protocol failed to resist replay attacks.

In-vehicle networks similar to smart grid environments have a higher frequency of transmitting messages between sensors, so there are also several verification-related studies. Jiang *et al.* [20] proposed a secondary method for vehicle-to-infrastructure communication. Binary Authentication Tree reduces computational complexity by reducing the entire batch verification bottleneck. Wang *et al.* [43] found that Jiang *et al.*'s BAT scheme cannot resist forgery attacks, and an attacker can forge the entire batch of verifications and signatures of other vehicles. Shim [40] also pointed out that Jiang *et al.*'s BAT scheme cannot resist replay and witch attacks.

3.4 Related Works on Anomaly Detection Efficiency of Smart Grids

State-based energy theft detection uses additional resources such as wireless sensor networks, observation instruments, and RFID to identify energy theft. Han & Xiao [15] proposed using observation instruments to calculate energy losses. This observation instrument tracks the energy delivered to consumers in the smart grid NAN (Neighbor Area Network). Use the approximate difference between actual energy consumption and billing data to distinguish tampered meters from healthy meters. Zheng *et al.* [50] proposed an anomaly detection framework by combining the maximum information coefficient and clustering technology to quickly search and explore density peaks to detect anomalies through smart observation instruments between consumers and DNOs (Distributed Network Operator). DNOs and energy retailers collect data from smart meters and sensors to identify malicious behavior. Kim *et al.* [23] proposed a non-technical loss detection algorithm that uses the Intermediate Monitor Meter to analyze the power consumption of each consumer to detect whether electricity is stolen. However, the above methods may cause the system to produce misjudgment results due to sudden changes in consumers' electricity consumption habits.

Energy theft detection based on machine learning uses data from smart grids and is trained to find anomalies automatically. Guntur & Sarkar [14] used consumers' historical energy consumption patterns to improve the efficiency of energy theft detection. The experimental results show that the random forest set accurately detects energy theft. Still, achieving robustness and stability for highly inconsistent energy consumption models is challenging. Yip *et al.* [48] used multiple linear regression to identify the location of faulty smart meters using consumer-based energy consumption patterns. This mechanism does not perform well with unstable energy consumption characteristics. Yao *et al.* [47] used a Convolutional Neural Network to propose a smart grid energy theft detection scheme to detect whether abnormal behavior occurs in meter measurement data. They are using the Paillier algorithm to protect the power consumption data transmitted by users.

4 Research Issues

In this article, we will propose three research issues on smart grid data security and privacy. The issues to be discussed and solved in this research are as follows: In the first research topic, a data privacy protection framework for Home Area Networks (HAN) will be designed using blockchain. The second topic is to study lightweight key exchange to achieve mutual authentication, exchange ses-

sion keys to enhance security, and use digital signatures to ensure integrity. The third research topic combines the research results of the first and second topics to design a smart grid energy theft system. While ensuring data privacy and security, we improve detection efficiency, use blockchain to design smart contracts to detect users stealing electricity, and combine deep learning models to detect abnormal points in the power grid to maintain the intelligence and high availability of the power grid. Integrating three research topics will result in a smart grid that transmits data with privacy and security and detects abnormalities in real time.

4.1 Research Issue on the Data Privacy for Smart Grids

Smart grids need to transmit the electricity consumption data of many users. These data represent each user's living habits or methods. If the privacy of this data is not guaranteed, the user's privacy will be leaked, allowing other malicious users to use it. Anyone can take advantage of the opportunity. Ensuring data privacy protects users makes smart grid data challenging to forge and avoids unnecessary energy consumption in the grid. This research hopes to propose a smart grid architecture that guarantees privacy, protects the data privacy of users and smart grids, and makes the grid less susceptible to damage. The following problems to be solved are described separately:

1) Users' living habits can be inferred from electricity consumption data:
   Today's society attaches a great sense of privacy, and protecting privacy has become a research and discussion direction for many people. Most of the data transmitted in smart grids are users' electricity consumption. This data includes which users used how much electricity at a certain point in time. If information such as this is intercepted and deciphered by malicious users, the smart grid will not be able to protect the user's privacy. This research intended to study the privacy protection methods of smart grids for home area networks (HAN) and design a privacy protection framework to solve this problem so that related facilities can be deployed more quickly.

2) Ensuring data privacy can make power grid data less susceptible to forgery:
   In addition to users' data privacy, the data privacy of the smart grids needs to be taken seriously. The smart grid has many different sensors, and the data transmitted by these sensors often plays a significant role in maintaining the grid's stability. Privacy is not adequately protected when the total amount is transmitted to the local control center for distribution. The total amount of electricity in the area is revealed; it may be possible to infer information such as the area's total population data. The power allocated to the area is far lower than the demand, leaving users without power. This research studies the privacy of smart grid data transmission. It solves this problem by combining the blockchain tool hyperledger so that the data of users and the grid will not be disclosed while also allowing the smart grid to maintain its stability.

To summarize the above, the first research topic studies the data privacy issues of smart grids, the use of blockchain to design a privacy protection mechanism, and the design of a data privacy protection framework for HAN to ensure the privacy of users and smart grids and also It can improve the efficiency of future subordinate systems.

4.2 Research Issue on the Security Certification of Smart Grids

In the second year of this project, it is expected to solve the security problem of data transmitted in the smart grid. The reason why the smart grid can minimize energy loss is partly dependent on the transmitted data. The transmitted data has various nodes in the power grid. The status data is more of the user's power consumption data. However, if the security of this data is not guaranteed, it can easily be tampered with by malicious users. It can easily lead to incorrect data in the power grid, causing the power grid to have a negative impact on energy errors in judgment. This project is expected to study key exchange and digital signature methods to enhance the security of data in the power grid to solve this problem. The description is as follows:

1) Prevent user information from being tampered with or destroyed:
   The smart grid will determine its distribution strategy based on the current level of electricity usage. Figure 1 shows the data transmission process of the smart grid. The electricity usage data will be measured through smart meters and the data will be sent to the smart grid. In the power grid, if the user's data is tampered with or destroyed during the transmission process, not only will the data between the meter and the power grid be asymmetric, but it will also cause errors in the power grid's strategic selection. This project will use lightweight key exchange to ensure security, and use a low-latency digital signature mode to authenticate the integrity of these data, improving the credibility of the data in the power grid.

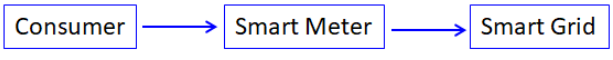


Figure 1: Smart grid data transmission process

2) Protect data transmitted in the smart grid:
   The data transmitted in the smart grid will be used as a reference to determine the grid's current power

strategy. If these data are tampered with and destroyed, the smart grid will need to repeatedly verify their correctness, which will increase the network load and further degrade the grid. Operations will be inefficient or even unable to provide services. Therefore, this plan ensures the security of data transmitted to the power grid and reduces the transmission cost of the power grid.

To summarize the above, the second-year research plan is expected to study the data security issues of smart grids and use key exchange and digital signatures to complete identity authentication with lower latency and computational costs to enhance data security and integrity.

4.3 Research Issue on the Batch Verification of Smart Grids

The smart grid needs to transmit the power consumption data of many users. This data is continuously transmitted according to the power supplier's measurement cycle, which may be once an hour or even 15 minutes. Many AMI devices simultaneously send statistical data to the power supplier for verification. For example, Without an efficient verification mechanism, it can easily cause delays in verification operations. The research will explore improving the existing verification algorithm to achieve batch verification of power consumption data in smart grids. This problem can be solved by verifying all data in the batch simultaneously to ensure that the data is correct. The problem is described as follows:

1) There are many smart grid users and a vast number of verifications:
   The smart grid consists of many Neighborhood Area Networks (NANs), as shown in Figure 2. Each NAN represents a floor, building, or community and is equipped with many smart meters and other equipment to record users' power consumption data. Due to the large number of users, the electricity consumption data generated will be extensive. Suppose the traditional transaction-by-transaction verification mechanism is still used. In that case, it will cause severe delays in the verification operation and make it difficult to achieve real-time power consumption monitoring and electricity bill calculation. This project will study how to use high-efficiency power consumption data verification algorithms to enable smart grids to have the ability to verify massive amounts of transmitted data in batches to achieve real-time energy utilization.

2) Ensuring data verification efficiency can reduce the delay of the power grid and achieve immediate results:
   Smart grid operation relies on the accuracy of a large amount of power consumption data to determine the best grid dispatch strategy. However, the efficiency of verifying a large amount of user data one by one is slow, leading to delays in power grid policy judgment and making it impossible to check whether user data has been maliciously tampered with. This will not only reduce the operating efficiency of the power grid but also fail to meet the needs of real-time monitoring and processing of power consumption data. This research will propose a batch data verification mechanism suitable for smart grids to accelerate verification efficiency.

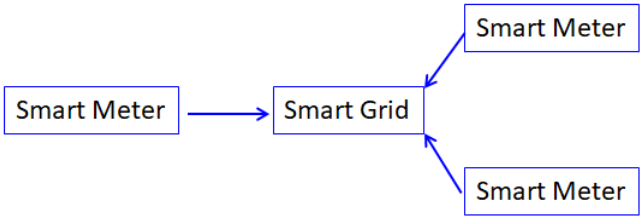


Figure 2: The data transmission from multiple AMIs

To summarize the above, the research studies the smart grid's batch verification scheme and explores its efficiency concerning the reputation score structure. Reducing the power grid's computing cost can significantly increase the verification speed and ensure the power grid strategy. The immediacy and correctness of judgment can achieve the effect of real-time monitoring and processing of electricity consumption data.

4.4 Research Issue on the Anomaly Detection Efficiency of Smart Grids

In the third research issue, it solves the problem of energy theft in smart grids. Since the smart grid will automatically collect power consumption data from various places, from this data, the smart grid can check whether there are any abnormalities in the data to achieve real-time monitoring. If the power in the power grid is stolen but no alarm is issued, it will not only cause serious economic losses to the power supplier but will even cause the power grid to be unable to operate normally due to unequal power generation and power consumption. Based on the first and second research topics, this research topic uses private and secure data to design the power grid to accurately detect abnormal points with better efficiency and improve the availability of the power grid. The issues to be addressed are detailed below.

1) Smart grid requires high anomaly detection efficiency:
   The smart grid has many sensors and nodes, and the data generated by these sensors or nodes will affect its power decisions. If abnormal points cannot be detected in time and repaired immediately, they will not only reduce the overall transmission efficiency of the smart grid but also cause unnecessary energy waste.

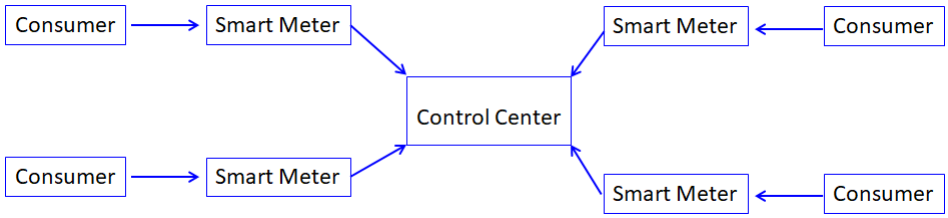


Figure 3: Schematic diagram of smart grid anomaly detection

2) Electricity theft will lead to information asymmetry in the overall power grid:
Smart grids determine power distribution strategies based on the amount of electricity suppliers generate. Figure 3 is a schematic diagram of smart grid anomaly detection. When someone steals electricity but is not detected, the power generation and power consumption will not match, and there will be no excess power in the energy storage device, thus affecting the power strategy of the smart grid. In addition to hindering the interests of energy suppliers, it will also put the entire power grid in an unsafe state. This research studies how to strengthen smart grid energy theft detection and improve grid availability.

5 Conclusion

We have proposed four research issues. These research issues will be able to achieve the following goals:

1) Ensure the privacy of consumers' consumption data in smart grids: This research issue proposes a solution to protect the privacy of consumers' data and proposes an efficient and practical framework.

2) Strengthen the security certification of smart grid transmission data: This research considers the security of data transmitted in the smart grid, using key exchange and digital signature methods to enable each node of the smart grid to establish a secure communication channel and to sign the data so that intruders cannot forge the data.

3) Improve the efficiency of smart grid verification data: This research proposes a data verification efficiency of smart grids. The proposed used a whole batch verification scheme to verify the signatures of the entire area at one time so that the smart grid has high-efficiency and low-latency data flow and achieves the effect of real-time power control of the grid.

4) Break through the abnormality detection efficiency and accuracy of smart grids: This research proposes a system with data transmission privacy and security, rapid and accurate detection of smart grid energy theft, and data circulation in each node. This will not only allow the smart grid to properly utilize its power and make the best power distribution strategy, but it can also solve imminent energy problems and improve the availability of smart grid applications in the future.

Acknowledgments

The National Science and Technology Council partially supported this research, Taiwan (ROC), under contract no.: NSTC 112-2221-E-507 -005 -MY3, MOST 109-2221-E-468-011-MY3, and NSTC 112-2221-E-468-007.

References

[1] D. Abbasinezhad-Mood and M. Nikooghadam, "Design and hardware implementation of a security enhanced elliptic curve cryptography based lightweight authentication scheme for smart grid communications," *Future Generation Computer Systems*, vol. 84, pp. 47-57, 2018.

[2] W. Ali, I. U. Din, A. Almogren, B. S. Kim, "A novel privacy preserving scheme for smart grid-based home area networks," *Sensors*, vol. 22, p.2269, 2022.

[3] M. R. Asghar, G. Dán, D. Miorandi and I. Chlamtac, "Smart meter data privacy: A survey," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 4, pp. 2820-2835, 2017.

[4] S. H. Baghestani, F. Moazami, M. Tahavori, "Lightweight authenticated key agreement for smart metering in smart grid," *IEEE Systems Journal*, vol. 16, no. 3, pp. 4983-4991, 2022.

[5] F. Bao, C. C. Lee, M. S. Hwang, "Cryptanalysis and improvement on batch verifying multiple RSA digital signatures", *Applied Mathematics and Computation*, vol. 172, no. 2, pp. 1195-1200, 2006.

[6] C. C. Chang, M. S. Hwang, "Parallel computation of the generating keys for RSA cryptosystems", *Electronics Letters*, vol. 32, no. 15, pp. 1365–1366, 1996.

[7] H. Debiao, C. Jianhua, H. Jin, "An ID-based client authentication with key agreement protocol for mobile client–server environment on ECC with provable security," *Information Fusion*, vol. 13, no. 3, pp. 223-230, 2012.

[8] R. Dong, S. Hao, T. H. Yang, Z. Tang, Y. Yan, J. Chen, "Recent advances in smart meter: Data analysis, privacy preservation and applications," in

Big Data and Security (ICBDS'21), Communications in Computer and Information Science, vol 1563, Springer, 2022.

[9] T. ElGamal, "A public key cryptosystem and a signature scheme based on discrete logarithms," *IEEE Transactions on Information Theory*, vol. 31, no. 4, pp. 469-472, Jul. 1985.

[10] E. Esiner, U. Tefek, H. S. M. Erol, D. Mashima, B. Chen, Y. C. Hu, Z. Kalbarczyk, D. M. Nicol, "LoMoS: Less-online/more-offline signatures for extremely time-critical systems," *IEEE Transactions on Smart Grid*, vol. 13, no. 4, pp. 3214-3226, 2022.

[11] S. Even, O. Goldreich, S. Micali, "On-line/off-line digital signatures," in *Advances in Cryptology (CRYPTO'89)*, Lecture Notes in Computer Science, vol 435, pp. 263-275, Springer, 1989.

[12] A. Fiat, A. Shamir, "How to prove yourself: Practical solutions to identification and signature problems," in *Advances in Cryptology (CRYPTO'86)*, pp. 186-194, Springr, 1986.

[13] M. Fouda, Z. Fadlullah, N. Kato, R. Lu, X. Shen, "A lightweight message authentication scheme for smart grid communications," *IEEE Transactions on Smart Grid*, vol. 2, no. 4, pp.675-685, 2011.

[14] S. K. Gunturi, D. Sarkar, "Ensemble machine learning models for the detection of energy theft," *Electric Power Systems Researc*, vol. 192, p. 106904, 2021.

[15] W. Han, Y. Xiao, "NFD: A practical scheme to detect non-technical loss fraud in smart grid," in *International Conference on Communications (ICC'14)*, pp. 605-609, 2014.

[16] D. Hankerson, A. J. Menezes, S. Vanstone, *Guide to Elliptic Curve Cryptography*, Secaucus, NJ, USA: Springer-Verlag, 2003.

[17] M. S. Hwang, C. C. Lee, Y. C. Lai, "Traceability on RSA-based partially signature with low computation", *Applied Mathematics and Computation*, vol. 145, no. 2-3, pp. 465-468, Dec. 2003.

[18] M. S. Hwang, C. C. Lee, J. L. Lu, "Cryptanalysis of the Batch Verifying Multiple DSA-type Digital Signatures", *Pakistan Journal of Applied Sciences*, vol. 1, no. 3, pp. 287-288, 2001.

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

[20] Y. Jiang, M. Shi, X. Shen, C. Lin, "BAT: A robust signature scheme for vehicular networks using binary authentication tree," *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, pp. 1974-1983, 2009.

[21] D. Johnson, A. Menezes, S. Vanstone, "The elliptic curve digital signature algorithm (ECDSA)," *International Journal of Information Security*, vol. 1, no. 1, pp. 36-63, 2001.

[22] G. Kalogridis, "Elecprivacy: Evaluating the privacy protection of electricity management algorithms," *IEEE Transactions on Smart Grid*, vol. 2, no. 4, pp. 750–758, 2011.

[23] J. Y. Kim, Y. M. Hwang, Y. G. Sun, I. Sim, D. I. Kim, X. Wang, "Detection for non-technical loss by smart energy theft with intermediate monitor meter in smart grid," *IEEE Access*, vol. 7, pp. 129043-129053, 2019.

[24] P. Kumar, A. Gurtov, M. Sain, A. Martin, P. H. Ha, "Lightweight authentication and key agreement for smart metering in smart energy networks," *IEEE Transactions on Smart Grid*, vol. 10, no. 4, pp. 4349-4359, 2019.

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

[26] H. Li, R. Lu, L. Zhou, B. Yang, X. Shen, "An efficient merkle-tree-based authentication scheme for smart grid," *IEEE Systems Journal*, vol. 8, no. 2, pp. 655-663, 2014.

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

[28] N. Liu, J. Chen, L. Zhu, J. Zhang, Y. He, "A key management scheme for secure communications of advanced metering infrastructure in smart grid," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 10, pp.4746-4756, 2013.

[29] Y. Liu, W. Guo, C. I. Fan, L. Chang, C. Cheng, "A practical privacy-preserving data aggregation (3PDA) scheme for smart grid," *IEEE Transactions on Information Theory*, vol. 15, pp. 1767–1774, 2018.

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

[31] K. Mahmood, S. A. Chaudhry, H. Naqvi, S. Kumari, X. Li, A. K. Sangaiah, "An elliptic curve cryptography based lightweight authentication scheme for smart grid communication," *Future Generation Computer Systems*, vol. 81, pp. 557-565, 2018.

[32] P. McDaniel, S. McLaughlin, "Security and privacy challenges in the smart grid," *IEEE Security & Privacy*, vol. 7, no. 3, pp. 75–77, 2009.

[33] A. Metke, R. Ekl, "Security technology for smart grid networks," *IEEE Transactions on Smart Grid*, vol. 1, no. 1, pp. 99-107, 2010.

[34] A. Molina-Markham, "Private memories of a smart meter," in *Proceedings of the ACM Workshop Embedded Sensors System Efficiency Building*, pp. 61–66, 2010.

[35] J. A. Momoh, "Smart grid design for efficient and flexible power networks operation and control," in *IEEE/PES Power Systems Conference and Exposition*, pp. 1–8, 2009.

[36] A. Muzumdar, C. Modi, C. Vyjayanthi, "Designing a blockchain-enabled privacy-preserving energy theft detection system for smart grid neighborhood area network," *Electric Power Systems Research*, vol. 207, p. 107884, 2022.

[37] C. P. Schnorr, "Efficient identification and signatures for smart cards," in *Advances in Cryptology (CRYPTO'89)*, pp. 239-252, 1989.
[38] A. Shamir, Y. Tauman, "Improved online/offline signature schemes," in *Advances in Cryptology (CRYPTO'01)*, pp. 355-367, 2001.
[39] H. Shen, Y. Liu, Z. Xia, M. Zhang, "An efficient aggregation scheme resisting on malicious data mining attacks for smart grid," *Information Sciences*, vol. 526, pp. 289–300, 2020.
[40] K. A. Shim, "Reconstruction of a secure authentication scheme for vehicular ad hoc networks using a binary authentication tree," *IEEE Transactions on Wireless Communications*, vol. 12, no. 11, pp. 5386-5393, 2013.
[41] J. L. Tsai, N. W. Lo, "Secure anonymous key distribution scheme for smart grid," *IEEE Transactions on Smart Grid*, vol. 7, no. 2, pp. 906-914, 2016.
[42] Z. Wan, G. Wang, Y. Yang, S. Shi, "SKM: Scalable key management for advanced metering infrastructure in smart grid," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 12, pp. 7055-7066, 2014.
[43] H. Wang, B. Qin, J. Domingo-Ferrer, "An improved binary authentication tree algorithm for vehicular networks," in *Fourth International Conference on Intelligent Networking and Collaborative Systems*, pp. 206-213, 2012.
[44] M. Wen, R. Xie, K. Lu, L. Wang, K. Zhang, "FedDetect: A novel privacy-preserving federated learning framework for energy theft detection in smart grid," *IEEE Internet of Things Journal*, vol. 9, no. 8, pp. 6069-6080, 2022.
[45] C. C. Yang, T. Y. Chang, M. S. Hwang, "A new anonymous conference key distribution system based on the elliptic curve discrete logarithm problem", *Computer Standards & Interfaces*, vol. 25, no. 2, pp. 141-145, 2003.
[46] C. C. Yang, T. Y. Chang, M. S. Hwang, "A new group signature scheme based on RSA assumption", *Information Technology and Control*, vo1. 42, no. 1, pp. 61-66, 2013.
[47] D. Yao, M. Wen, X. Liang, Z. Fu, K. Zhang, B. Yang, "Energy theft detection with energy privacy preservation in the smart grid," *IEEE Internet Things Journal*, vol. 6, no. 5, pp. 7659-7669, 2019.
[48] S. C. Yip, W. N. Tan, C. Tan, M. T. Gan, K. Wong, "An anomaly detection framework for identifying energy theft and defective meters in smart grids," *International Journal of Electrical Power & Energy Systems*, vol. 101, pp. 189-203, 2018.
[49] Y. Yoldaş, A. Önen, S. M. Muyeen, A. V. Vasilakos, İ. Alan, "Enhancing smart grid with microgrids: Challenges and opportunities," *Renewable and Sustainable Energy Reviews*, vol. 72, pp. 205-214, 2017.
[50] K. Zheng, Q. Chen, Y. Wang, C. Kang, Q. Xia, "A novel combined data-driven approach for electricity theft detection," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 3, pp. 1809-1819, 2019.
[51] L. Zhu, L. Zhu, M. Li, Z. Zhang, C. Xu, R. Zhang, X. Du, N. Guizani, "Privacy-Preserving authentication and data aggregation for fog-based smart grid," *IEEE Communications Magazine*, vol. 57, no. 6, pp. 80-85, 2019.

Biography

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.

Yung-Ling Chang received her M.S. degree in Management from Huafan University in 2021. She is currently pursuing her Ph.D. in the Artificial Intelligence program at Asia University. Her research interests include big data analytics and database management security.

Ko-Yu Lin received the M.S. degree in Management Information Systems, National Chung Hsing University in 2024. His research interests include smart grids, security in sensors, and blockchain.

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

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.

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.